Appendix 3. Reviewer Comments

Reviewer Comment	Rand Response
Title is not very informative. Should include something about the conditions under study. Example: Efficacy and Safety of Ephedra for Weight Management and Athletic Performance Enhancement.	Change Made
Since the stated overall objective is "to assess he efficacy of herbal ephedra and synthetic ephedrine on weight loss and athletic performance ……" and since there is stated too few studies and data available to conduct an analysis of herbal ephedra on athletic performance, should not the title of the study be altered or the reported at least noted to reflect this limitation?	We think this is more appropriate for the text, and the title reflects the uses for which we attempted to find evidence.
I think you did an excellent job. Having reviewed this subject in more superficial fashion in the past, I can appreciate, more than most, what fine job you have done.	No Response
The overall purpose of the evaluation, including the questions, methods, findings and conclusions are clearly and succinctly written and easy to understand.	No Response
The search for relevant data appears to have been thorough and encompassed a broad range of literature resources.	No Response
The study selection appears to be appropriate for an evidence-based review of this type.	No Response
Data collection and data synthesis appear to be reasonable.	No Response
This is an excellent comprehensive review, and it will make an important contribution to the literature. Strong points are a clear description of review criteria, rigorous assessment methods, and straightforward data presentation. The questions formulated are relevant and appropriate, search strategies seem reasonable, and study selection is well justified. The meta-analyses are useful.	No Response
The Evidence Report utilizes modern methods of meta-analysis of clinical trials. However, it ignores a great deal of scientific evidence that can augment the interpretation of data from the clinical trials and has a major structural flaw and several weaknesses that are discussed below.	No Response. A specific response to the "great deal of scientific evidence ignored" is presented where such evidence is specifically referred to.
In my area of expertise (clinical studies of obesity), the findings were consistent with my understanding of the literature.	No Response
The overall evaluation is clear, and the purpose of the report is well stated.	No Response
Overall I found the report well-researched and written.	No Response
The questions were adequately formulated and easily understood.	No Response

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
Reasons for inclusion or exclusion of studies were clear.	No Response
In evaluating the obesity weight loss clinical studies, the report acknowledges some of the problems [small numbers of subjects, short durations of treatment, etc.] and states that long-term assessments of effectiveness are lacking. It would be useful to put these statements in the context of current knowledge in this area: that weight loss generally ceases after about 6 months irrespective of the treatment and any weight lost is generally regained. Current recommendations for appropriate clinical trials in this area include a much longer duration of treatment [1 – 2 years] and an evaluation of what happens after the agent is withdrawn . Both of these are very important in evaluating the efficacy and the risk to benefit ratio of a particular substance. Although ephedrine plus caffeine combinations [pharmaceutical and dietary supplement sources] are being compared to certain prescription drugs, to date no ephedrine plus caffeine product has undergone the equivalent types of efficacy and safety studies that are required prior to marketing of a prescription drug in the US.	This information was added to the limitations.
The purpose of the study and the means for arriving at its conclusions were clear and relatively easy to follow. The Meta analysis approach was appropriate and the criteria well defined. I believe some discussion should be given to the purported mechanisms of action (i.e. anoretic versus thermogenic) behind the " statistically significant "weight loss attributed to synthetic ephedrine/caffeine/ or ephedra-containing dietary supplements. The impression given by the meta analysis results is that, while statistically significant, these types of products also provide clinically relevant weight reductions. Given the results of the case report analyses, I don't believe the benefit of minimal weight loss (e.g. 1 to 3 pounds per month) outweighs the potential risk of serious adverse health effects exemplified by the case report analysis. Despite the study's inability to assign causality to most, if not all, of the serious adverse events, the authors, in their conclusion, seem to downplay the "potential" risks associated with these products.	This communicates a value judgment about the balance of evidence that is beyond the scope of the EPC. The concern about the report “downplaying” the potential risks is, as later peer review comments will indicate, shared by some other reviewers, but directly contradicted by others.
The appraisal of ephedra studies for weight loss could include a stronger statement about the unusually high attrition rates as compared to many drug studies. Although this is mentioned in the Limitations section, it also might be included in the results section where the data is interpreted. Can you expand on whether attrition rates differed between treatments a placebo groups? In my view, this is a major weakness of the recent efficacy studies involving ephedra.	Attrition rates did not differ between treatment and placebo groups. This has now been added to the results.
It is stated under Findings [p4] and elsewhere “that in aggregate the clinical trials only enrolled a sufficient number of patients to detect a serious adverse event rate of one per one thousand” or “three per thousand “ in the case of	Change Made

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
botanical sources of ephedrine It would be useful to put these numbers in the context of the frequencies of adverse events [common, infrequent, rare, etc.] Using commonly accepted definitions , all of the current clinical trials in aggregate, irrespective of source, lack the power to detect any rare adverse event [defined as greater than 1 per 1000 rate or frequency].	
Throughout the report, reference is made to "synthetic ephedrine". I suggest deleting "synthetic", since ephedrine is ephedrine. Some is extracted from plants and some is synthesized.	As these contradictory comments indicate, there is no agreement among experts about standardized terminology. In this report, for simplicity's sake, we use the term "ephedra" to mean the herb or herb abstract, and "ephedrine" to mean the chemical, regardless of source.
The term "synthetic ephedrine" is ambiguous due to the meaning of the terms "natural" and "synthetic" with respect to natural products chemistry. What could be meant are synthetically derived ephedrine alkaloids because these are natural products by virtue of their existence as naturally occurring compounds regardless of how they are produced. Ephedrine is by definition always a natural product unless one is referring to the racemate that is produced during some synthetic production processes because the specific optical isomer that is identical to naturally occurring ephedrine is itself is itself often synthesized through chiral specific processes. The fact of the matter is that what the draft means when referring to "synthetic ephedrine" could be either naturally or synthetically derived. It may be preferable then, in the interest of clarity throughout the document, to use some consistent terminology, such as: "ephedra" as the name of the crude raw material (with parenthetical identification of the pinyin name: ma huang one time, but not as a substitute common name) which consist of the dried stem of the plant; "ephedra extract" when referring to raw materials or ingredients that are processed extracts of ephedra; "ephedrine " when referring specifically to those one alkaloids as found in the plant or wherever the term "synthetic ephedrine" now occurs in the draft.	
I would place "synthetic" in front of all mention of ephedrine, or ephedrine alkaloids; for policy experts and others it is important to make the distinction between herbal and synthetic. I would use "herbal" ephedra when possible. I would also state more directly and more often why the synthetic ephedrine use is not reviewed as part of the AERs.	
Also, contrary to the phytochemical section of the report, ephedra is know to contain (-)-norephedrine but not (+)-norephedrine. Phenylpropanolamine consists of (-/+)-norephedrine, while ephedra does not contain (+)-norephedrine. The parenthetical identification of norephedrine as phenylpropanolamine should therefore be removed.	Change made

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
A minor point in phytochemistry (page 13) is that only (-)-norephedrine occurs naturally in ephedra, whereas the synthetic drug, phelypropanolamine is the racemic mixture of (+/-) -norephedrine. So, it is more precise to state that ephedra contains norephedrine, as opposed to containing PPA.	
After reading the RAND report, my first impression is the following: What are we evaluating – ephedrine or herb ephedra? The latter is not a single-chemical entity and cannot be assumed to be ephedrine. Even assuming the herb ephedra in the literature is defined to contain specific dosage levels of 'ephedrine,' what efforts were made to ascertain that this 'ephedrine' is indeed ephedrine and not a mixture of ephedrine-type alkaloids, or, worse, different types of alkaloids that are also present in ephedra? Any study or report on a natural product (not just a single-chemical compound) must clearly define what the material under study or being reported is. I don't see such a definition in this report. Despite the limited availability of useful data, this report's conclusions regarding the efficacy of ephedrine (the single-chemical drug), in the presence and the absence of caffeine, in short-term weight loss and athletic performance, appears to be sound. However, this cannot be said of the herb ephedra that contains ephedrine but is not equivalent to ephedrine. Hence, the conclusion regarding ephedra's efficacy "Ephedrine, ephedrine + caffeine, and ephedra-containing dietary supplements + herbs containing caffeine all promote modest amounts of weight loss over the short term..." lacks supporting data, unless all the limited number of clinical studies employing "ephedra-containing dietary supplements" had clearly defined ephedra, including amounts of ephedrine and related alkaloids (not just ephedrine and inert herb carrier).	We agree that the lack of specificity is a problem. We have modified the conclusions to be more specific to only these herbal combinations studied. In the RCTs of herbal ephedra included in the efficacy analysis, the dose of ephedrine alkaloid was stated.
The ODS and AHRQ contracted with RAND (Dr. Paul Shekelle as Task Director) to conduct a thorough synthesis of the clinical efficacy and adverse effects of ephedra. It was clear to me that the objective of this contract was met. The review was complete and the researchers used the systematic review/meta-analysis tool to review the published controlled clinical studies on ephedra-containing dietary supplements.	No Response
There was a mention of 157 articles that were case reports of adverse events published in medical journal, however, they are not included in the case report and there are no mention of the finding in the Limitation section on page 110. Would those case reports provide more information than what are available from FDA? Should a statement be made on why those published case reports not included in the analysis (e.g. potential duplication with FDA, time and	These case reports are now included in this revision.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
resources...etc.)?	
The report has been carried out and is free from bias. It is objective. I cannot comment on some areas of the report that are not my expertise.	No response
The draft report is incomplete since it does not include a review of studies of two types: Toxicology in laboratory animals, and published case reports.	Published case reports are now included. Toxicology and animal studies were not included, as this (and most all EPC reports) focus on clinical studies in humans.
This draft emphasizes the subjective judgments of the authors over the objective findings of the clinical studies and therefore appears from the outset to have a slant against the safety of ephedra products.	We disagree that the report is slanted against the safety of ephedra products, and note the peer review comments we received with exactly the opposite opinion (i.e. that we were too conservative in our conclusions regarding possible adverse events from ephedra).
Given the observations and comments above, one is left with the impression that this draft report has a tone or tenor that leans toward an apparently preconceived conclusion that ephedra supplements are not safe. The tone is established in the abstract by reference to the FDA's AERs "related to herbal ephedra" and "available reports of herbal ephedra-related death, myocardial infarction (heart attack) and cerebral vascular accident (stroke)." the abstract goes on to describe "our causality algorithm" and later to use terms "probably causally related" and "possibly casually related". Nowhere in the abstract is it suggested that these purported AERs were looked at objectively and found (to quote page 112) that "definite causality cannot be determined from case reports". The statements in the abstract strike the reader as definite scientific conclusions rather than subjective observations that is not consistent with other objective data. Nowhere in the abstract are the major limitations described, nor is there any mention that "scientific studies (not additional case reports) are necessary" (from page 113). On page 5 it is stated that "Continued analysis of case studies cannot substitute for a properly designed study to assess causality", yet this is precisely what this draft report has done. Additional statements and references point to a lack of objectivity and a bent toward sensationalism. For example, page 5 a comparison is drawn with phenylpropanolamine and it's "reported association and cerebral hemorrhage" without nothing that the report at issues is highly controversial, or that the report found no association between ephedrine as an over-the-counter drug. The mention of cerebral hemorrhage at the conclusion of the abstract is presumably also a result of some unfounded	We have endeavored to keep the language of the report as factual as possible. We note that other reviewers criticized the report for exactly the opposite reason – being "too soft" and "down playing" the risks of ephedra use. We do not think we can revise the report to reconcile these two divergent opinions. With regard to format, this report adheres to EPC format requirements. With regards to phenylpropanolamine, we note other reviewers critiqued us for not making more of possible similarities. In this case, we deleted the phenylpropanolamine sentence.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
conclusion that phenylpropanolamine have been conclusively tied to cerebral hemorrhage when this is not the case. In counterpoint to a description of the extent of present use and to the long history of use in China, references are made to media attention, lawsuits, a citizens' petition, and a ban by the National Football League, and a Canadian Warning. These references are not helpful in a scientific review that should be evidence based, but instead give an impression of the slant toward a view that ephedra products are not safe.	
Although the draft report contains much factual information about both the benefits and risks of ephedra, ephedrine, and combinations of one or the other with caffeine sources, certain critical components of a full analysis are missing. Specifically, there is much proper emphasis on examining the data for evidence of causality, but little or no attention to the dose-response relationship within any possible causal case. This is a critical limitation that prevents the safety component of the report from being fully useful.	A dose response analysis has been added to the RCT analysis. We indicate that we do not feel such an analysis is justified on the case report data.
The strength of this report is that it is not only comprehensive, but also objectively performed. Another strength of this report is defining the areas that need further research. The limitations are those imposed by the data.	No response
It is clear what was done.	No response
The major strength of study was the statistical approach utilized for assessment of efficacy and the incidence of minor adverse effects.	No response
The major limitation was the coupling of conservative causality assessment criteria with limited medical records and toxicology data while interpreting the case reports. While the case reports do not offer mechanism for assessing the incidence of serious adverse events, they shouldn't be dismissed completely owing to an overly conservative set of exclusion criteria. Case control studies are definitely warranted, but it would be especially tragic if their outcome, when determined three of four years from now, confirm what is strongly suspected at the moment.	We acknowledge our criteria are conservative. We note the great deal of discussion among peer reviewers regarding whether a case report analysis was biased toward or against the safety of these products.
There are some nomenclature issues in the draft that should be corrected or clarified. The term "herbal ephedra" contains a redundancy, as by definition, all ephedra is herbal or herbally derived. Also, and this goes beyond nomenclature to ingredient definition, the term ephedra is often used in the draft when in fact what is being discussed is an extract of ephedra with a specified percentage of ephedrine alkaloids. This sort of misrepresentation of material identity leads to confusion between ephedra as a crude botanical, an extract of ephedra (the form most often used in dietary supplements) with a specified percentage of constituent ephedrine alkaloids, (usually 8%), and the	We have endeavored to keep the nomenclature clear.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
ephedrine alkaloids themselves.	
Perhaps it could be made clearer that, overall, a very small number of people have been studied in controlled trials of any duration. This is an issue as regards to safety, rather than efficacy where the studies, though small, are quite consistent.	We emphasize the limited power of the RCTs to assess safety.
This review reflects my perspective as a neurologist and stroke researcher. It is a very valuable collection of data assembled to address clear, relevant clinical questions.	No response
It was clear how the report was developed.	No response
The major strength of this report is its collection of data systematically on one report for review and assessment.	No response
The major limitation is the way in which the conclusion are stated and failure to distinguish for the lay reader the difference in strength of evidence of adverse reports vs. intervention studies.	We have tried to make this distinction clearer in this revision.
This well-done report takes a conservative approach without extrapolating the interpretation beyond the available data . It clearly describes the methods used, limitations of the methodology, and results. The text under Future Research describing identification of gaps in knowledge is particularly useful. The presentation of the analysis of adverse events reports (AERs) might be made clearer by using different terminology or a narrative explanation of the causality designations.	No response, other than causality has been removed from this revision.
Quality of Life. As I view the field of obesity, there are two reasons people want to lose weight. One is for the health-related benefits. For most physicians, of whom I am one, this is often the major focus of our support for efforts to lose weight. However, over the years, I have come to realize that the major reason people want to lose weight is because obesity is a "stigmatized" condition. The fact that 75% or so of the people volunteering for treatment are women, and that obesity carries such a negative social view stimulates people, particularly women, to use over-the-counter medications. Yet there is no mention that I can find of quality of life in this report.	We agree this outcome is important. However, we did not find it reported in the clinical trials we identified.
Body Composition. One of the interesting responses to treatment with ephedrine and caffeine in the reports of Astrup and his colleagues is the increase in lean body mass, or loss of less lean body mass. The implications of this for use of these medications and in the future research is not even mentioned that I can find.	This distinction is not one that was included as an outcome of interest by our TEP. We agree it is a potential area for future study.
Performance. There is quite a literature on caffeine and performance that certainly plays into the ephedra/caffeine use by athletes. Yet none of this literature is dealt with here.	We were not requested to assess the literature on caffeine and performance.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
Drop Outs. The issue of drop-outs is considered with the <20% vs. >20%. From a therapeutic effect, the "completers" in a trial are much more informative to me than using the data on those who drop-out in a last observation carried forward analysis. We are certain that drop-outs are likely to regain weight - We aren't curing obesity and weight gain during the adult life is the "expected". Moreover, if we do not use the LOCF approach, the impact depends strongly on when people drop out. If they drop-out at month 5 of a 6 month trial it has essentially no effect. If they drop out in the first month it has a major effect.	We agree that knowing when dropouts occurred might make it possible to better understand the results of weight loss trials. However, when dropouts leave a study is not routinely reported and hence we did not have access to these data.
On page 3 there is mention that the studies have "particularly high attrition rates." What is considered a high attrition rate? How do these studies compare to other studies on obesity? There is no explanation as to whether there is a particular challenge in all obesity studies or research in general, or whether this attrition rate appears to be unique to the ephedra studies.	The attrition rate issue is explained in more detail on page 27, where 20% is identified as a threshold. A high attrition rate is not unique to studies of ephedra, but regardless of study question a high attrition rate increases the concern regarding bias.
Long-Term Trials. In the Future Research area you call for "longer" term trials. For all reported drugs the maximal weight loss is achieved by 6 months. Continuing treatment usually maintains an effect, but because weight losses of 10% (20 lbs for someone weighing 200 lbs) does not often get them to a satisfactory weight, people drop-out because of perceived "failure" of the medication. I thus have limited enthusiasm for long term studies with agents that don't produce weight losses of more than 10%. On p. 4 you indicate that there are "no long term" studies. As noted above, I think the 6 month studies that reach a plateau tell us about all we can expect from these trials. Do you disagree?	We clarified this to indicate both longer duration of treatment and maintenance of weight loss.
The report should be reorganized to focus on the conclusions about the need for further research. The section on safety should address expected effects at intended doses and comment on adverse effects of higher doses. The transient nature of the events observed in the clinical studies should be discussed. The FDA AER database unfortunately is not of sufficient quality to comment on either of these issues related to safety.	We do not know if the events observed in the clinical studies were all transient and would not characterize them as so. A dose analysis is now included in this revision.
There should be another draft report issued to the TEP to ensure that these issues are addressed to the satisfaction of the TEP before a report is finalized.	This is not EPC practice, and there is no requirement that the TEP be “satisfied” before the report is finalized. We intentionally recruit TEP members holding differing views in order to be made aware of all viewpoints. Trying to get all such people to be “satisfied” with the final report is an impossibility as demonstrated by the wildly diverging comments we received from TEP members regarding the causality analysis.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
Para 1 I would use the word treatment duration, intervention length, or other terms designating the duration for which the participants were on ephedra/ephedrine instead of using "follow up". I would not use follow-up as it denotes a passive time post-intervention for which participants were followed to measure outcomes. e.g. 19 were excluded from pooled analysis because their intervention periods were less than 8 weeks.	Change Made
The term "follow-up" can have a number of meanings in the context of obesity/weight loss trials. In addition to referring to the duration of treatment with a test agent during which a research subject is evaluated, it can also refer to patient evaluation after treatment has been discontinued. From my reading of this report, you are using "follow-up" to only refer to the time during which treatment is administered. It might be useful to clarify this in the text as 8 weeks of treatment, etc so as to avoid any confusion in meaning.	Change Made
I was slightly troubled by the exclusion of studies with less than eight weeks' follow-up. While I agree that studies with less than eight weeks' follow-up are undesirable, if a large number of such studies exist, it does seem unfortunate to exclude them. I would rather have seen them included and have separate analyses for a very short-term weight loss and slightly longer term weight loss. I think that the exclusion of such studies, if there are many, opens up the report to allegations from companies who have done such short term studies that their important data were not included and that the report is biased. I am not stating that I believe the report is biased, but only by any exclusion of such studies if there are many, opens up the report to this allegation. Moreover, while as I said previously, I do not favor studies less than eight weeks' duration, I still think that while such studies exist there is something we can learn from them.	The exclusion of studies less than 8 weeks duration was made by the TEP and not something we can change at this stage. The key question specified "a sustained period of time" for efficacy and this was judged by the TEP to be at least 8 weeks.
Page 2, para 4: It would be useful to give the reason that the Technical Expert Panel (TEP) gave for suggesting that follow-up of less than 8 weeks is insufficient to assess weight loss. It is because the original charge was to assess long-term weight loss and the TEP thought 8 weeks could not be considered long-term?	Yes, and furthermore even short term weight loss would not be useful below 8 weeks. Explanation made in the Methods.
Follow-up of 8 weeks. This term used on p. 2 and then many other places is confusing. As a clinical investigator, follow-up usually means the time after treatment is complete. You appear to be using it only for the treatment period. It would confuse me less if you said "duration of treatment".	Change Made
This review excluded 19-controlled trials that assessed ephedra or ephedrine for weight loss because there was follow up of less than 8 weeks in each of these. This	These studies were included in the safety assessment.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
exclusion is rational from the perspective of evaluating the evidence for efficacy. Information obtained from these trials about short term adverse effects, or the lack thereof, would be valuable however in the overall evaluation of safety. We strongly encourage the inclusion of all such data from these trials.	
The statement, "In order to improve health outcomes, long term weight loss is necessary" is not accurate. Usually in pharmacotheraphy for weight loss, long-term means one year or more. I am not aware of studies that have used time in place of percent body weight loss as the important measure. Because your point is that the studies were short (<=4 months) I would change loss to maintenance because Yanovski et al. (2002) states that most nonsurgical obesity treatments lead to weight loss for the first four to six months followed by regain. It is not only that the ephedra interventions did not extend beyond 4 months but also that there was not sufficient follow-up to determine if individuals were able to maintain their loss. See review by Yanovski et al., 2002 New Engl Journal Med.	Change Made
Rewrite last sentence to say "In order to improve health outcomes and reduce the risk of morbidities associated with being overweight, sufficient weight loss (5 to 10% of body weight) and long term weight maintenance is necessary.	Change Made
See comments on page 3 and 5 regarding use of term follow-up; treatment duration, ephedra intervention, etc.	No Response
Small weight losses (5 to 10%) of body weight reduce the risk of morbidities associated with being overweight (Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults. National Heart, Lung, and Blood Institute (NHLBI), Clinical guidelines for obesity, 1998.)	Reference Added
This first paragraph seems to blend intervention duration and follow-up post intervention. Please rewrite to reflect data. Longest intervention: 4 months (this is not "follow up").	Change Made
You do not address whether individuals lost a certain percent of their pre-ephedra weight. This measure is important when it comes to defining weight loss success.	Percent of weight loss in the treatment group is now included in this revision.
Maybe the key points of DSHEA needs to be stated in the overview or somewhere else to emphasize herbal supplements versus supplements containing synthetic alkaloids. Maybe place a sentence after the "In addition to the questions related to ephedra-…safety. Because synthetic ephedrine alkaloids….	We revised the text to try and improving clarity.
On page 3 the report mentions that an algorithm for assessing causality was developed by the authors. Was the	This algorithm was deleted from this revision

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
algorithm unique to this study, or is there already significant scientific agreement to its accuracy and validity? If it is a new algorithm, who suggested its use? How was "reasonable certainty" determined?	
The Draft identifies question that guided this Report, both in relation to weight loss and energy enhancement, as "Does ephedra have additive effects with other agents?" Specific emphasis was placed on caffeine and caffeine-containing botanicals, but in Table 1 herbal "agents" were listed as "Herbs commonly combined with ephedra," presumably (though not stated) in products marketed for weight loss.	Change Made
It is stated that “the majority of ephedrine (up to 97%) is excreted unchanged by the urine.” The 97% seems too high. The recent paper by Christine Haller et al (Clin Pharmacol Therap 2002;71:421-32) indicates that about 60% of ephedrine is excreted unchanged in the urine. This is important because the other 40% can be metabolized to other pharmacologically active alkaloids.	Change made
In places, particularly the introduction , the report focuses more on ephedrine than ephedra. Since there were only 5 trials assessing ephedra for weight loss (actually 4, since one is reported twice) and many more synthetic ephedrine, the ephedrine trials would seem to have greater weight than the ephedra trials. Not clear how this influences the results.	We present the results stratified by agent. The efficacy results for ephedrine & ephedra were similar.
There are several problems with Table 1. No references were given to inform as to how the herbs included in this Table were identified as "commonly combined with ephedra" and in fact it is our belief that several of the listed herbs are either uncommonly found in products containing ephedra and marketed for weight loss or are not found in the market. For example, although the aloe resin is known to be a cathartic laxative, we are not aware that it exists as an ingredient in any ephedra product (or in any dietary supplement product), and if it does it is certainly not common. Without attempting to be exhaustive, the same is true for at least the following: cocoas, coffee, scotch broom, jalap bark, and mayapple root. In addition, several of the ingredients are at best questionable for the described categories, and follow-up should be undertaken to find references to support that yellow dock root is a cathartic laxatives. These examples are again not exhaustive; references should be given to support each herb in its classification. An additional oversight is that some listings do provide the part of the plant that is purportedly a commonly combined with ephedra, though very nearly 50% do not. The federal law requires that botanical ingredients in dietary supplements identify the plant part and this Table should do the same.	We greatly shortened this Table to include just the caffeine-containing herbs, as suggested by this reviewer.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
Finally, the Table does not appear to provide any information that is useful toward answering any of the questions proposed by the funding agencies or those that guided the Report. While the question of the additive effect of other agents was proposed and reportedly guided the report, there is no attempt in the Report to actually do this, except in the case of caffeine containing herbs. In summary, it might be best to eliminate the Table to reduce it to consist of just the caffeine containing herbs. If the table is maintained, some effort should be made to actually find each of the listed ingredients in one or more products in the market. This is especially true for hers with significant toxicity potential, such as Scotch broom to mayapple as the final report should not communicate that these ingredients are "commonly" sold. Preferably, such market information would be provided in the form of references. The part of the plant that is used should be including for any plant listed in this Table. References should be provided as to how classifications are made if the categories in the Table are maintained. Notwithstanding the above comments the question whether all of these herbs should be included in the Table, there are several spelling errors in the botanical names: Coffea is correct, as in the 1st such listing but Caffea is not; Camellia is correct as in the 2nd such listing, but Camilla is not; the correct spelling of the species name for Mate is paraguariensis; the references species of mayapple is P. peltatum while Rheum palatum us correctly recorded as rhubarb, R tanguticum (misspelled in the Table) is considered to be a variety of R. palmatum (so R. palmatum var.tanguticum) and R. officinale is misspelled in the Table; the correct spelling of the botanical name for flax ends in "m" rather than "n" (so Linum usitatissium);Irish moss is in the genus Chondrus, not Chrondrus; contemporary authorities accept the name of the slippery elm to be Ulmus rubra rather than Ulmus fulva, these corrections may not be exhaustive.	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
The characterization of DSHEA in the Background section of Chapter 1 is inaccurate, biased, unnecessary and badly written!! It should either be removed -It has nothing to do with the assignment-or expanded, to include other elements of the law. For example:" The DSHEA was passed unanimously in 1994 based in part of Congressional displeasure with the federal governments 'adhoc, patchwork regulatory policy on dietary supplements.' Under these regulations, herbal dietary supplements are not necessarily required to be tested for safety prior to marketing, although marketers are required to assure all of their products are free of significant or unreasonable risks. Also, as with over-the-counter drugs, there is not a requirement to report health problems that resulted from their use. The federal regulations that govern this class of goods are different from this that control either foods or drugs, but as with both of these classes, FDA and FTC maintain significant authority to regulate the manufacture, labeling and claims for dietary supplements and to remove unsafe products."	We have included some, but not all, of this additional material when describing the DSHEA.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
The Dietary Supplement Health and Education Act of 1994 (DSHEA) The brief mention of Public Law 103-417 is inadequate. In 1994, Congress passed the Dietary Supplement Health and Education Act (DSHEA) amending the Federal Food, Drug, and Cosmetic Act. In DSHEA, the term "dietary supplement" is defined as: 1. A product other than tobacco intended to supplement the diet that bears or contains one or more of the following dietary ingredients: * a vitamin; * a mineral; * an herb or other botanical; * an amino acid; * a dietary substance for use by man to supplement the diet by increasing the total dietary intake; or * a concentrate, metabolite, constituent, extract, or combination of the above listed dietary ingredients. 2. A product that is intended for ingestion is not represented as food or as a sole item of a meal or diet, and is labeled as a dietary supplement. 3. It includes an article that is approved as a new drug, or licensed as a biologic, and was, prior to such approval, certification, or license, marketed as a dietary supplement or as a food unless the Secretary has issued a regulation, after notice and comment, finding that the article, when used as or in a dietary supplement under the conditions of use and dosages set forth in the labeling for such dietary supplement, is unlawful. 4. It excludes articles that are approved as a new drug, certified as an antibiotic, or licensed as a biologic, or an article authorized for investigation as a new drug, antibiotic, or biological for which substantial clinical investigations have been instituted and for which the existence of such investigations has been made public, which was not before such approval, certification, licensing, or authorization marketed as a dietary supplement or as a food, unless the Secretary, in the Secretary's discretion, has issued a regulation, after notice and comment, finding that the article would be lawfully marketed as a dietary supplement. 5. It deems a dietary supplement to be a food. 6. It excludes a dietary supplement from the definition of the term "food additive." Important safety measures were included in DSHEA. A food could be deemed to be adulterated if it was a dietary supplement or contained a dietary ingredient that: 1. presents a significant or unreasonable risk of injury; 2. is a new dietary ingredient for which there is inadequate information to provide assurance that such ingredient does not present such risk; 3. poses an imminent hazard to public health or safety; or 4. contains an ingredient that renders it adulterated. Important clarifications were included in the law regarding labels and labeling. Section 5 of DSHEA provides that a publication shall not be defined as labeling when used in connection with the sale of dietary supplements when it: 1. is not false or misleading; 2. does not promote a particular manufacturer or brand of supplement; 3. is displayed so as to present a balanced view of the available scientific information; 4. is displayed physically separate from such supplements; and 5. does not have appended to it any information by sticker or other method. 6. places the burden of proof on the United States in establishing that such matter is false or misleading. Additionally DSHEA: 1. Set forth conditions under which nutritional claims may be made with respect to such supplements. 2. Deemed a dietary supplement misbranded unless its labeling meets specified guidelines. 3. Deemed a dietary supplement which contains a new dietary ingredient adulterated unless: A	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
The draft sites in it's Chapter 1 the findings of a 1996 meeting of the FDA's Food Advisory Committee (FAC), stating that "over half of the members recommended removal of dietary supplements containing ephedra on the market" and gives as it's reference Dr. Lori Love's testimony in August 2000 at another meeting. To assure that the findings of this meeting are most accurately reported it would be best to add a statement such as "a finding that was in direct contravention to the recommendation of the Special Working Group of experts that had been empanelled to offer guidance to the FAC. "The transcript of the 1995 meeting of this Special Working Group can be seen at http://www.cfsan.fda.gov/~dms/ds-ephe1.html . The more important factor with regard to this statement, however, is that it is false. The transcript of this meeting is available on the FDA's website in two PDF files (see http://www.fda.gov/ohrms/dockets/ac/cfsan96.htm). Regardless of how Dr. Love characterized the recommendations of the FAC members, the record shows that only 4 of the eleven voting members of the FAC stated that ephedra products should be removed; even when calculating the opinions of all the meeting's participants, well under half made statements to that effect. The statement in the Draft could be corrected either by changing "over half of" to "a minority of" or by reversing the two sentences (At this time over half of the members recommended that the FDA develop rules on use that would help reduce risk over adverse events, a recommendation that trade groups had made two years earlier". Finally, the use of the word "Thus", at the beginning of the next sentence in this section implies a direct relationship between the reported advice of the FAC and FDA's imposed rule. This is a reinvention of the historical facts. FDA stated in its proposal was based on information that included, but was not limited to the opinions of the FAC. More detail should be added to this section if the report is to be an accurate record of facts. If the only limitation accessible about the history of the controversy regarding the use of ephedra in dietary supplements was from the Background in the Draft's Chapter 1, one would conclude that federal health officials, consumer groups and National Football league had been actively attentive to this issue while industry stood by. This is not the case. The Background information should be expanded to include some or all of the facts: that AHPA adopted labeling guidelines in 1994 that were substantially familiar to those later proposed by FDA; AHPA adopted dosage limits (25mg/servind; 100mg/day of ephedra alkaloids) in 1995; AHPA and others specifically requested in public hearings in 1995 and 1996, and in a meeting with FDA in 1999 that the industry policies be adopted by rulemaking; AHPA and others submitted a Citizen petition in October 2000 (prior to the Public Citizen petition identified in the background) to make the same request in a more formal manner.	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
Many scientists would disagree with the statement [page 4, and elsewhere] that “definite causality cannot be determined for case reports when the adverse event is very serious” [or various iterations of this statement]	We note this comment, and also note that many scientists would agree with it. At any rate, we have deleted from this revision the causality assessment.
p.9 Background states “Three billion servings of ephedra containing products were consumed during 1999” This is a misstatement , as in the transcript Mr. McGuffin indicates “servings sold” rather than servings consumed. As a separate comment, it is unclear as to whether the data on the number of servings actually represents servings manufactured by a particular company or some other measure.	We have revised this statement to make clear this is the industry’s contention.
D. Finally, the report should not repeat the industry assertion that three billion servings of ephedra were consumed in 1999, unless this is based on hard facts. It's a self-serving statement that has the effect of diminishing the safety concerns over ephedra by perhaps inflating the frequency of exposure. Is the three billion estimate based on quantities sold? Surely not all dosages were consumed.	Change made to reflect this is an industry assumption.
p.10 FDA concerns about the safety of ephedrine alkaloid containing products sold as supplements preceded the passage of DSHEA, which changed how FDA could deal with safety in the context of supplements.	We revised the text to reflect this.
Two references in the background section should, in my opinion, be changed. On page 11, you state that "weight loss has been associated with decreased morbidity and mortality" and cite ref. 26, the Williamson et al study. Actually, the literature on this point is quite controversial, and despite the Williamson study, much of the literature shows an increase in mortality with weight loss. All of these studies are observational, and subject to serious limitation. This is why NIDDK is undertaking a very large study (Look AHEAD) to answer questions about morbidity/mortality with voluntary weight loss. The DPP does suggest that intentional weight loss in persons at risk can delay or prevent the onset of type 2 diabetes in persons at high risk. I suggest that you state instead that "intentional weight loss in obese persons leads to reductions in risk factors for disease" and cite the NIH guidelines: Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults--The Evidence Report. National Institutes of Health. Obes Res 1998; 6 Suppl 2:51S-209S.	Change made
Also, on page 14, ref. 69--when discussing the role of ephedrine in humans, its role in stimulation of beta three adrenergic receptors in brown fat is noted. There is very little brown fat in adult humans, and I'm unsure that this would play any role in ephedrine's thermogenic effect. The reference cited is an old one (1982). Someone should be sure that this citation represents current thinking on the role	We deleted this comment.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
(if any) of brown fat in the thermogenic effects of ephedra compounds.	
The definition of overweight is >=25-29.9 (not in excess of 25 but also inclusion of 25) and the definition of obesity is >=30 (not greater than 30 but inclusion of 30) See NHLBI, Clinical guidelines for obesity, 1998	Change made
The attempted intentional weight loss data is only for 1996. The 1998 data you reference is a paper that only includes a subset, only 5 states. The latest national data on attempts for weight control is the 2000 data that is in Reference 10. Therefore, you may want to delete reference to the 1996 data and instead use the 2000 or just edit the sentences to say "The same survey when administered in 2000 showed that one third (38.5%) of subjects were actively trying to lose weight and another third (35.9%) were trying to maintain their weight.ref 10 Furthermore, among those who were overweight 45.0% of subjects were actively trying to lose weight and 34.9% were trying to maintain their weight. Among those who were obese, 65.7% of subjects were actively trying to lose weight and 20.8% were trying to maintain their weight" ref 10. I then go on to reference 29 data. The would suggest using the estimates from Reference 29 to determine a denominator for use. I would suggest also using the Michigan data from this paper to support claims that consumers are not aware of the ingredients in their herbal supplements. Ref 29 –"In a population-based study of 14,679 U.S. adults in 5-states using the 1998 BRFSS data, 7% reported using nonprescription weight loss products; 2% reported using PPA and 1% reported using ephedra products from 1996 to 1998. More women used ephedra products than men; 1.6% of women and 0.4% of men reported using weight loss products containing ephedra. Extrapolated nationally, this study estimated that during 1996-1998, 2.5 million Americans used weight loss products containing ephedra. "This study also has data to suggest that many individuals are not aware they are taking weight loss products that contain ephedra. Of the 183 respondents in Michigan who responded no to the questions about using ephedra and reported to have taken "other" nonprescription weight loss products, 33% reported using name-brand products that claim to contain both ephedra products and chromium picolinate. "	Change made
I would rewrite the sentence regarding Harnack et al. (2001) to be the following (inclusion of small n and previous of ephedra specific for weight loss). It is hard to follow what the 12% of the total is when you don't give the original total usage (61.2%); and, it is really 5.3% that used ephedra for weight loss. This is larger than Ref 29, but Ref 29 has	This change was made consistent with previous reviewers comments.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
14679 individuals whereas the Harnack study has only 376. Among 230 (61.2%) of 376 adults in the St. Paul/MSP area who reported using an herbal products during the past 12 months, 44 (19.1%) used ephedra. Of these 44, 20 (45%) used ephedra for weight loss. Therefore, 5.3% of adults (20 of 376) reported using ephedra for weight loss. --Taking these estimates, you find that 20 (5.3%) of 376 individuals used ephedra for weight loss during the past 12 months (1998/1999).	
Some of the herbs mentioned in the last sentence are not listed in Table 1. Many of the latest formulations also contain bitter orange. I would add these to the Table.	This Table has been greatly shortened and this part has been deleted.
On page 9, the following key information is provided: Ephedra has been used for over 5,000 years. Three billion doses have been sold. Even after the FDA's campaign to advertise the AERs and to have more AERs reported, there has been a 65% increase in volume of sales over the previous five years. Even after the FDA's campaign, there are only 1,500 AERs out of 3 billion servings. That calculates to about 1 adverse event in every 2 million servings. By anyone's standards that is very safe.	This is a judgment and not a statement of evidence, which is what the Evidence Report presents.
On page 10, the statement, "Still, the controversy over ephedra continues," and a reference to litigation have no place in a scientific analysis. It is doubtful that such information was garnered from a review of the published scientific literature. Inclusion of this type of information takes away from the science.	We disagree that these sentences take away from the science, we think they are necessary to put the science in context.
Information from the scientific literature on ephedrine (the purified alkaloid) regarding it mechanism of action. There is a fair amount of literature (1910 to 1930) about ephedrine. For example, Chen KK, Schmidt CF. Ephedrine and Related Substances, Medicine volume 9, number 1, 1930.	This section of the report was not intended to be exhaustive, but to provide context for the reader. Many relevant references may not be included.
Page 10 paragraph 1: Obesity, The definition of obesity had changed since 1991. A result of this change was that in the mid-1990's many more people were considered obese than previously. So, although the incidence of obesity had been increasing since 1991, the change in definition makes it seem more dramatic than it actually was. As a result, this statement may need to be qualified.	This is probably true, but by most standards the incidence is increasing. At any rate we did qualify the statement.
Page 12, paragraph 2: It is appropriate to extrapolate figures from 511 subjects attending a gymnasium to the general public? Suggest qualifying this statement.	We indicated this is the authors' extrapolation.
The RAND Corporation has drafted a document entitled "Ephedra: Clinical Efficacy and Side Effects" in order to assess the efficacy of herbal and synthetic ephedrine on weight loss and athletic performance and to assess the safety of herbal ephedrine products through review of adverse events reported in clinical trials and in reports on	No specific response to these general comments. Specific responses to specific comments below.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
file with the U.S. Food and Drug Administration (FDA). This report will focus on the safety assessment in the RAND report. Prior to commenting on this assessment, however, it is important to note that the RAND report includes a formal meta-analysis that concludes that products containing herbal ephedra and caffeine produce significant weight loss over a 4 to 6 month period. This weight loss is similar to that documented from synthetic ephedrine plus caffeine. Given the epidemic of obesity in the United States and the associated morbidity and mortality from obesity, it must be emphasized that weight loss may play a large role in reducing morbidity and mortality. In fact, several studies have shown that weight loss associated with herbal ephedra and synthetic ephedrine are associated with significant reductions in parameters associated with cardiovascular disease among the obese (e.g., reductions in triglycerides, 1, ApoB, 1 and LDL-cholesterol;2 and increases in HDL-cholesterol2).The evidence for weight loss is quite robust because it is derived from controlled randomized trials. The best data for safety would also be derived from randomized trials. RAND reports on adverse events within randomized trials, noting that there were "no serious adverse events (e.g., death, myocardial infarction, stroke) reported in these clinical trials." Because of the limited numbers of subjects studied in these trials, these studies could only detect a serious adverse event rate of one in a thousand. That is the studies can exclude a rate of serious adverse events of greater than one in a thousand. This should not be inferred to mean that the rate is one in a thousand nor that ephedra even causes adverse events. In the absence of additional controlled studies, RAND then turned to adverse event reports (AERs) filed with the FDA up to September 30, 2001. The limitations of AERs in proving causality, especially when viewed in isolation of the totality of evidence, are well known and have been discussed extensively in the literature and basic textbooks of pharmacoepidemiology. Causality cannot be proven by AERs because there is no comparison control group. Authors of other reviews of AERs in the ephedra database have noted that a collection of AERs "does not prove causation, nor does it provide quantitative information with regard to risk."3 There are several reasons for this which will be discussed briefly.	
To what end was this inquiry directed? Most exogenously ingested chemicals that are thought to enhance athletic performance are banned from competitive sports. To what end will the results of such an inquiry be applied?	This report was commissioned by AHRQ & ODS to assess the state of the science regarding ephedra.
Given the most contradictory recommendations of the CANTOX report regarding the safety of the ephedra supplements, why was a member of CANTOX included as a	We believe that every person on the TEP has a bias. Our goal in selecting the TEP was to try and get a balanced set of

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
part of the Technical Expert Panel? Some bias may have been imparted from a member of an organization with such close ties with the ephedra industry.	biases. We judged it important to include in the TEP a member with close ties to the industry who was also scientifically credible.
Of note, no neurologists (and particularly no stroke experts) were included on the Technical Expert Panel.	While true, a neurologist was included in the group assessing the case reports and a neurologist was included in the peer reviewers.
The group was charged with assembling and evaluating the evidence that ephedra and its congeners favorably affected "energy enhancement", affected weight loss and improved athletic performance. "Energy enhancement" is a vague term and it is not clear how it can be measured or tested.	Our TEP defined this for us as indicated in Table 3.
A basic problem in the method adopted for the pooling of studies rests with the false assumption that these herbal preparations have been standardized and are similar enough in constituents, potency and purity that they can be assumed to have sufficient homogeneity to justify pooling of results. In fact, there is much evidence that this is not the case. Yet much of the report rests on the results of the pooling of many under-powered studies of herbs or ephedra where potency and constituents are vaguely described or even unknown. The herbs are mixtures of many chemicals with various actions so it is doubtful merging or pooling such studies represents a scientifically legitimate exercise.	We disagree with this opinion, and point out the chi-square test of heterogeneity did not reject the null hypothesis of no difference in the effects reported in the four ephedra studies.
In sum, this report addresses two questions that are not relevant to the public health: "energy enhancement" and "improved athletic performance". The public health question it does address, obesity or weight loss, is not answered due to the heterogeneity of the products examined and pooled and the lack of long-term follow-up studies. The case reports of adverse events possibly due to ephedra or ephedrine are not well described and the algorithm adopted is too rigid and is being applied to a data collection system that is unable to obtain the data required for causality in the algorithm.	The key questions were given to us by Federal Agencies and defined by our TEP. Causality was removed from this revision.
There is no reason to engage in further research concerning ephedra or ephedrine. Enough is known about its benefits and risks to remove the drug from the market. More research is merely a stalling device to delay the removal of the product from the market.	This is an opinion and not a comment about evidence to which we can respond.
That ephedra has efficacy for weight loss seems to be true, though people can quibble over how much. The key question, as I see it, is: Is there credible evidence that ephedra poses a significant or unreasonable risk of harm, even when taken at recommended dosages? This report so far is inadequate for addressing this question.	No response as there is no specific critique of methods or analysis.
Page 23, paragraph 3: This paragraph is not clearly written.	Change Made

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
What about: To be accepted for pooled analysis, studies were required to be controlled clinical trials according to the following definitions [insert definitions].	
Page 26, first sentence: Not sure about the implications for BMI of assuming an average height 5'8". Where did this number come from? Does it affect the outcome significantly is this number is off?	We chose this number arbitrarily. The results do not change across a range of potential heights.
Because study ref 87 is in adolescents, I question whether transformation of the data using a height of 5' 8" is a good decision. I would suggest contacting the study authors and request the actual individual height data.	A sensitivity analysis using 5'4" made little difference in the results. Therefore we do not feel it necessary to contact the original authors.
Page 27, Paragraph 1: Less than 20% attrition is a commonly accepted threshold below which concerns about bias increase due to loss of follow up. Should this read greater than 20% would be concern for bias.	The reviewer is correct, change made.
Page 28: Meta-Analysis. Will the two Danish trials be included in the final analysis?	Yes.
Update literature searches past December 2001, if appropriate.	Done
The questions guiding the evidence report were relevant, well formulated and easy to understand. The only problem I saw were questions 3, 7 and 12 were the same (Does ephedra have additive effects with other agents?) Most of the questions were related to the herbal ephedra, but much of the data reviewed was based on synthetic ephedrine.	These questions are the same but refer to, respectively weight loss, athletic performance, and safety.
"Of the 517 articles collected, 56 were controlled clinical trials of either synthetic ephedrine or herbal ephedra..."If the number of articles are added together (56+146+84+19+47+4+3+157), there are only 516 articles. According to page 53 of the Evidence Report, there are 48 controlled trials identified. It is unclear if the two Danish trials are included or not. Even if it is, there are still discrepancies with the number.	These numbers have been reconciled.
In reading the objectives, one assume that synthetic ephedrine was also part of the study objectives. However, this is not the original intent (see page 19 on Original Potential Key Questions). Should the changed in objectives be explained in the evidence report (rather than just a statement of agreement by TEP on page 20)?	It is explained in the text why this change was made, and we also changed the title to reflect this.
It will be more corrected to state that, "Forty-eight were controlled trials assessing ephedra/ephedrine for weight loss."	Change Made
Most likely the only detailed analysis of all the trials of ephedrine in the literature. The detailed explanation of the method with tables and graphs are helpful to the reader.	No Response

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
Should the same list of questions applicable to ephedrine as this is listed in the Objectives, or does ephedrine serve only as information (if that is the case, the objectives need to be reworded)?	We listed the question as received from AHRQ, and then describe how these were modified for the task order.
Page 20, paragraph 2, last sentence:… categories of patients: children, adolescents, young athletes, and adults. These are not generally patients, but rather potential consumers of ephedra or ephedrine products.	Change Made.
Page 22: Additional sources of evidence. Readers may think it unusual (as did several of our reviewers) that RAND would place an announcement seeking unpublished studies in Phytomedicine and Herbalgram. They would wonder why such announcements were not put into more mainstream medical journals such as JAMA and Lancet. It might be useful to mention that the intent in choosing Phytomedicine and Herbalgram was to reach individuals who might know of small studies being done on ephedra or ephedrine the TEP may not have been familiar with.	Change Made.
It is quite evident that a concerted effort was put forth by the authors to search all relevant databases and literature sources for clinical studies assessing the efficacy of the ephedrine/caffeine and ephedra containing dietary supplements. I was somewhat surprised that advertisements were placed only in Phytomedicine and HerbalGram. Phytomedicine is a relatively obscure journal while HerbalGram is targeted more toward the layperson. Were other journals considered?	
On page 26, the authors state that when a standard deviation was missing they imputed an average standard deviation from all other available data. They further state that they weighted all other standard deviations equally (IF I understood them correctly). It was unclear to me why they would weight all the standard deviations equally rather than weighting them by same size.	We weighted each study equally in the imputation procedure, i.e. we did not weight each study by its sample size (we assume the reviewer meant “sample size” not “same size”). Neither approach (weighting equally or weighting by sample size) is entirely consistent with our assumed random effects model. The approach we did take is simply applied, and we have found in practice that the results are fairly insensitive to weight choice.
On page 42, the authors indicate that in their reporting form, BMI greater than 27 was defined as obesity. This seems an odd choice given that both the NIH and the World Health Organization have now reached a consensus that a BMI greater than or equal to 30 should represent obesity and this information was available prior to the initiation of the current project.	The form has been correct to read “overweight/obese”.
On page 55 the authors describe some power analyses. It was not crystal clear to me what null hypothesis was under consideration in the power analyses the described. I think	This text was revised to try and increase clarity.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
greater clarity in this section could be achieved.	
In the weight loss category, of the 24 trials listed in Evidence Table 2, 4 were excluded, 2 because of study design, the other 2 because of "insufficient statistics" and lack of "weight loss outcome" (addressed weight gain). However, 4, available only as abstracts, were rated using the Jadad system, scoring 0,1,2, and 2. Of the 20 trials included in this weight loss panel, only 5 used herbal ephedra, the other 15 employing the pure alkaloid ephedrine. Interestingly the highest scores (5) on the Jadad scale were to the two Boozer studies, which combined herbal caffeine (kola and guarana, respectively).	No Response
Regarding the efficacy aspect of ephedra/ ephedrine use, the rejection of 19 of the identified 48 controlled trials on the basis of a lack of 2 month follow-up, appears reasonable, but assessment of the 19 terms of safety indication may add to the pool of data.	These were included in the safety analysis. The text has been changed to reflect this.
The literature search seems to be appropriate, with the relevant publications being identified. The study selection for efficacy analysis seems justified, whereas the selection of studies for safety is not appropriate. This reviewer finds it justified including only the controlled trials with a placebo arm for efficacy analysis. But for safety evaluation is obvious that all trials should be included. The safety information collected during a clinical trial has much better value and validity than the cases received through the FDA. I suggest therefore that the analysis of safety in terms of adverse effect dropouts and side effects should be re-examined with inclusion of all the available trials.	We did include all available trials in the safety analysis and have clarified the text to reflect this.
Could you estimate the average amount of weight loss per month in each of the ephedrine groups and in the placebo weight loss groups. For the lay press and political readers this may mean more to them than the difference in weight loss between the active intervention and the placebo.	This revision now contains the percent weight loss in the treated group.
It is clear there are data gaps with respect to ephedra use and effects which are apparent in this study. However, I did not identify any evidence of bias in the data collection process.	No Response
It appears that the researchers made every effort to reduce bias in the data collection process. The data collection process was systematic and thorough. Problems were identified and explained in the Limitations section of the report. The researchers did acknowledge that missing information did exist and also described this in the Limitations section. It seems as if the researchers did the best they could have done with the literature captured in the meta-analysis.	No Response
Is there a minimal amount of missing information regarding	We conducted sensitivity analysis on

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
outcomes and other variables considered key to the interpretation of results? The fact that the studies that were found has 6 month or less treatment duration. That is too short a period of time to fully analyze safety or efficacy.	attrition rate greater or less than 20%, and the results are reported in the text. We acknowledge in the limitations that the short duration of the identified studies limits the conclusions that can be drawn.
I believe you did a fine job in synthesizing the data. There was one inconsistency that I think might have been a typographical error. In table 15 you state that the pooled monthly weight loss in pounds is 2.7. In the text on page 55 and 56, you state that the same monthly weight loss is 2.1 pounds. You my want to check this apparent conflict.	This discrepancy has been corrected.
Reasonable decisions were made concerning whether and how to combine data. Precisions of results were indicated. Limitations and inconsistencies were also stated.	No Response
All study designs were considered in the synthesis and reasonable decisions were made as to combining the data. Precision was reported and limitations described. Limitations and inconsistencies were stated along with limitations of the review process. The meta-regression was used in an attempt to compare treatment across trials.	No Response
At one or more points the authors used the term "cathartic". I am not certain I know what they mean by that . Do you mean laxative?	Yes.
I believe that there are at least three major reviews related to this topic to varying degrees that merit mention. I believe the authors have mentioned at least two of these three. The three of these are: the CANTOX Report; Frank Greenway's recent review, and a review by Allison and colleagues which appeared in critical reviews in Food Science and Nutrition in 2001. Each of these reports addressed the use of ephedrine products for weight loss in part or in whole. I do not think any of them need to be discussed at great length, but it should be mentioned an the authors of the current report should briefly mention whether their conclusions largely agree or do not agree with those prior reports. The authors are probably also aware that, subsequent to their producing this draft document, there was a Senate hearings on the use of dietary supplements for weight loss, at which a number of experts provided testimony. The written testimony from several of these experts are available on the Senate's website. The authors may wish to briefly mention this in their report and cite any key relevant information that appeared in that testimony that was not available to them this report was written.	In this revision we do not review previous reviews, therefore we did not act on this comment.
I think that as it intimated above, any studies excluded must be carefully accounted for. The exclusion of studies opens up the reports to potential allegations of bias. Therefore, I would advocate that the authors include a very detailed table of all excluded studies giving the reason for their	We considered doing so, but felt the report had so many tables already that this table was of marginal extra benefit.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
exclusion and a full reference to the study. If the authors have already done that and I missed it (the tables were quite extensive and I confess that I did not go through them with a fine toothed comb), I apologize.	
I was confused on one point. I thought that the authors only included studies that were randomized, double-blind, placebo-controlled trials. If I understand the scoring of the Jadad system correctly, a study that is randomized would get at least one point, and a study that was placebo-controlled (and therefore presumably double-blind) would get a second point. Therefore, all studies should receive a score of at least 2 on the Jadad scale. However, I thought that I saw some point that some of the studies received scores less than 2. Can this be clarified?	Studies were included if they were controlled clinical trials. Such studies can score zero on the Jadad scale.
I was somewhat disappointed by the authors discussion and use of effect sizes. First, the discussion is slightly simplistic at points. For example, it seems to imply that the particular effect size metric they use is "the" effect size rather than "a" specific metric of effect size. Moreover, it is generally well-recognized that when the outcome measure in a field of study is something that had intrinsic or accepted meaning it is perfectly reasonable to use this outcome measure rather than the particular effect size metric the authors used, which scales things relative to within group standard deviations. This is the case with body weight, where most investigators and people in general understand pounds and kilograms. There is no reason to standardize by the standard deviation, which makes the data less interpretable. In fact, several meta analyses have appeared in the literature on obesity and simply use pounds or kilograms. I agree that several meta analyses have as appeared in the obesity field that have used the standardized effect size that the authors use, but I personally see it as unnecessary. It is not only unnecessary but it can create situations in which there is less clarity. For example, two studies can achieve the same absolute weight loss and yet one study because it is much more tightly controlled may have a smaller standard deviation. This latter study would achieve a larger effect size and yet I do not think that most people would see it as more efficacious if the same number of pound or kilograms were lost. The authors themselves seem not to accept this metric of effect size because they later back transform if two pounds as a way to help the reader interpret the results. Finally, although a minor point, the authors should use metric units of weight (kilograms) rather than pounds. The metric system is the accepted system in current scientific practice and the authors report will be perceived as less professional otherwise.	The effect sizes were transformed back to changes in weight loss in pounds. We also assessed whether conducting the analysis entirely in pounds changed the results. It did not. We also continue to report the results in pounds because US audiences are more familiar with this unit of measurement.
Being not a statistician, I do not understand certain parts in	The methods text has been revised to

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
the Methodology. For example, effect size is not defined in the section “Weight Loss Effect Size” (page 25), the paragraph in the Safety Assessment, Controlled Trial Adverse Events, Meta-Analysis section on exact conditional inference methods versus asymptotic methods (page 29) was not very clear.	improve clarity.
On p. 2 you describe the "effect size" determination. Although it becomes clear later that you are comparing them with placebo, this one sounds like it is only for single treatments. I am confused.	An effect size is calculated for any comparison of two groups.
In general I think that this a reasonable objective, and have little doubt that this will make a useful contribution to the field. That being said, I think there are a number of points that, if carefully addressed, could improve the document. I detailed specific comments below. In the main summary (I.e. pages V and VI) the authors made no mention of dose. I think that this is a marked oversight. It could inappropriately be taken as an indication that they statements they make apply to all doses. Clearly this is not the case as the statements they make can only apply at best to the doses for which they observed the data. I believe the authors should consider substantially softening several of their statements. The first one to catch my eye was the statement on page VI that "the effects on weight loss of synthetic ephedrine plus caffeine and Ephedra-containing dietary supplements with herbs containing caffeine are equivalent..." As I am sure the authors are well aware, lack of evidence for an effect is not the same as evidence for lack of an effect. We can never marshal sufficient evidence to unequivocally prove the null hypothesis. We can only fail to reject a null hypothesis. If the authors had access to multiple, very well controlled studies comparing herbal and non-herbal ephedrine, this conclusion might be warranted. However, based on the data they have observed, a far softer statement such as "We observed no statistically significant difference between the effects of herbal and non-herbal sources of ephedrine and caffeine" would be much more appropriate. The authors may perceive me to be a stickler on this point. I am suggesting that the authors try to particularly cautious throughout this report in framing their conclusions because of the highly contentious nature of the topic they are studying. Even if these authors never enter a courtroom, it is highly probable that they will "speak" in one or more courtrooms through this document. That is, lawyers and expert witnesses representing multiple diverse interests are likely to cite this document in court cases. For this reason, it is crucial that the authors say exactly what they mean and state exactly what can be supported by data and be vary catious about making statements that could be	A dose analysis is included with this revision. We also tried to make sure our statements were accurate.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
misinterpreted or overextended	
On page 11of the report, the authors state some numbers regarding how many billions of dollars obesity costs. Although I do not think these numbers are especially relevant to the report and could easily be eliminated without any loss, if the authors are going to cite them they should cite the most accurate information available. My colleagues and I published a report in the American Journal of Public Health in 1999 in which we showed that prior estimates of the costs of obesity were almost certainly inflated by a fact of approximately 25%. If the authors are going to cite cost figures they should probably cite our paper and lower costs showed therein.	We stated that the reported value is only one estimate. The point, we think, is that obesity has an enormous cost in terms of health. We also included this reference stating that another estimate was 25% less.
On Page 14 under Pharmacokinetics, I thought the authors may wish to consider softening their statements about the lack of difference between herbal and non-herbal sources of ephedrine in terms of Pharmacokinetics. It seemed to me that the studies they reviewed did not fully support what appeared to be their conclusions, namely that there were no important differences in pharmacokinetics between herbal and non-herbal ephedrine.	We made this modification.
The authors state that two physicians working independently extracted data in duplicate and resolved disagreements by consensus. It would be interesting to know how often such disagreements occurred. That is, can the authors present any indication of the reliability of their coding scheme.	We did not assess in this project (or any similar project) a measure of disagreements, such as Kappa, and therefore cannot report this.
With respect to the search strategy, the authors seem to have been quite thorough. However, there are two sources they did not mention using that, in my experience can be extremely useful for this type of work. The first is the United States Patent and Trade Office which now has all patents on line. The online data base is searchable. One can often obtain quite a bit of additional information on this topic by finding companies' patents. Second, although, in my experience a less important source, Dissertation Abstracts International, Which also had on line searchable databases can occasionally help uncover additional studies. I can certainly understand the last thing the authors probably wish to hear at this point is a suggestion they go back and search for more literature. Whether they ultimately choose to do so is obviously up to them. However, at minimum they might want to do a type of "sensitivity search" to see if it seems likely that they would have missed a great deal of information by not searching these databases.	We did not go back and search these databases. No reviewers identified any missed trials, so while we can never be sure, we judge it unlikely that there are significantly large and well done RCTs that were not included in our analyses.
Regarding herbal ephedra for weight loss: There are apparently no studies addressing whether weight loss is maintained after ephedra use is discontinued. This is a very important gap in our knowledge and should be explicitly	An addition was made to the limitation section.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
pointed out. With all other weight loss medications, weight is regained after their use is discontinued, suggesting that life-long use (whether continuous or intermittent) is likely needed to maintain weight loss. If this is also true for ephedra, adverse events must be considered from the perspective of chronic ephedra use rather than episodic use. This has important implications for clinical use, public health, and study design to detect adverse effects.	
The HHS requested this analysis to evaluate the safety and efficacy of ephedra/caffeine products when used for weight loss or exercise enhancement used in the absence of medical supervision. However, the reports that have been analyzed have all been performed on subjects that were screened for pre-existing medical conditions and were followed during the trials with medical supervision. For example, in the most recent study by Boozer et al., the investigators excluded one of every ten subjects they screened for medical history or for conditions that made ephedra/caffeine, in their estimation, to be unsafe. The only trials that could have adequately addressed the question posed by HHS would be any that enrolled an unscreened population and followed them with little, if any, medical supervision. Such studies are not feasible or ethical because of the general knowledge that ephedrine-containing products are dangerous. An Institutional Review Board would not accept this study design. The report should note that the clinical trials reviewed (at least the ones with which I am familiar) had strict criteria for medical exclusion and require careful monitoring for safety during the study. This is the result of the general understanding of the medical community that these products are dangerous and therefore requires medical supervision during their use	The issue of studying select populations was added to the limitations.
The other major flaw in the analysis is that it failed to adequately consider the pharmacology and clinical pharmacology of sympathomimetic amines. The consistency of the evidence across a range of chemically related substances must be considered. The relative safety and efficacy of other drugs that have similar pharmacologic actions is absolutely relevant. Every drug with sympathomimetic actions that have been studied adequately has been associated with serious cardiovascular and neurological adverse events. Likewise, the actions of drugs that antagonize the effects of ephedrine should be considered. For example, adrenergic antagonists reduce the incidence of strokes and heart attacks.	These are topic areas that may be worthy of review but were outside our scope of work.
The questions are clearly formulated, but some of the answers are difficult to find. For example, the answer to question 7 on page 19 was buried in paragraphs on page 14. I might suggest adding to the summary chapter, brief answers to the questions you posed that are based on your	We reported the questions as we received them. We tried to reword our conclusions to better match the questions.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
analysis. Although the summary does address many of the questions, it would be nice to see the answers lined up with the questions. For example, the answer to questions 7 may be something short like this: " Ephedrine releases norephedrine from nerve terminals stimulating alpha and beta adrenergic receptors. Caffeine magnifies this effect by slowing the breakdown of cyclic-AMP inside the cell through the inhibition of phosphodiesterase." This is not an attempt to suggest text, but just to give an idea of how it might be possible to address the questions you posed in a two or three sentence answer. The questions are understandable but not well formulated. The questions should be specific for the way the ephedrine/caffeine products are being used. To ask whether they are safe without specifying how they are used ignores the potential selective bias.	
The selection of 24 hours as a window for exposure to ephedra is conservative. It is quite possible that ephedra could cause coronary or cerebral vasospasm that could persist much longer. This certainly has been described for other sympathomimetic drugs such as cocaine.	The 24 hour criterion was set by the TEP and not something we can change.
In evaluating the adverse events, why was documented use of ephedra with 24 hours made a criterion? This timing interval is far shorter than that used in the PPA epidemiology study [use ~72 hours]. Furthermore, this criterion tends to exclude those adverse events that are not necessarily time or dose dependent or whose effects are not ascertained until some critical threshold is exceeded [e.g., immunological reactions; hemorrhagic stroke with symptoms of an antecedent headache not considered “typical”].	This criterion was set by the TEP
Requiring documentation of ephedra exposure within 24 hours of the acute event may be biased against the most serious cases when a patient cannot provide a history of recent use because of death, coma, aphasia, or other sever impairment. In the absence of toxological results a reliance proxy history by a household or family member should be adequate.	We did count as satisfying this criterion a report of the subject consuming ephedra or ephedrine within 24 hours.
Most likely the only detail analysis of all the trials of ephedrine in the literature. The detail explanation of the method with tables and graphs are helpful to the reader.	No response
Is the FDA data on the products that contain ephedra based on label claims that it is ephedra or there are lab analysis confirmation? This should be stated as many of the products tested claim to contain “ephedra only” contain ephedrine and pseudoephedrine and no other ephedra alkaloids. This is likely due to a non-naturally occurring source.	It could be based on the label or on direct analysis.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
"In September 2001, the FDA's Office of Nutritional Products, Labeling, and Dietary Supplements produced an excel spreadsheet...for the dates specified" What is the inclusive date of the requested report?	From inception to September 2001.
The literature on herbal supplements/medicines is replete with reports based on undefined or poorly defined research materials. This occurs at least one of three levels: (1) research, (2) reporting research findings in journals, and (3) abstracting/indexing journal articles for database entry. Unless serious efforts are made immediately to set criteria for researchers at all three levels to follow, further research in the herbal supplements/medicines field will only continue to generate data that will continue to lead to ambiguous conclusions and hence, controversy. Ephedra and ephedrine are no exception.	No response
Ephedra herb (defined as the green herbaceous stem) sometimes contains up to 30% root material, which has different types of chemical constituents than those of ephedra herb. The root has completely different traditional uses than the stem as well (e.g., antiperspirant vs. diaphoretic). And the root contains macrocyclic spermine alkaloids (ephedradines) that are hypotensive, as opposed to the hypertensive effect of ephedrine in ephedra herb. Also, ephedra herb from different sources (Ephedra sinica, E. intermedia, E. equisetina, etc.) contains widely different levels of ephedrine among the ephedrine alkaloids (30%-90%) present in the herb.4 We can't assume the results from ephedra herb containing 'ephedrine' are equivalent to those based on the single-chemical drug ephedrine unless both the following two conditions are met: (1) the efficacy and safety evaluation is only based on ephedrine and (2) the concentration of ephedrine in ephedra has been specifically defined by definitive chemical analyses. Otherwise this 'ephedrine' could only be 30% ephedrine, with the rest (70%) being made up of other phenethylamines (e.g., pseudoephedrine, norephedrine, etc.) as well as ephedradines (from root material present as adulterant in the raw material used for extraction); the latter have different pharmacological activities and toxicities than ephedrine.	We added to the limitations the lack of standardized products for ephedra.
Line 3 should read this: "Less than 20 percent attrition is a commonly accepted threshold above which concerns about bias increase due to loss to follow up.	Change made.
The cases of seizures (n=70) and fainting/loss of consciousness (n=63) may represent serious cardiovascular events such as syncope due to cardiac arythmia.	The seizure cases are now included in this report.
You may want to add to table 1, Bitter orange extract (Citrus aurantium) and Garcinia Combogia.	This portion of this Table was deleted in this revision.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
In contrast to the Rand draft report, the relation of the potential of consumption of ephedra to the dosage involved was the central point of the ephedra risk assessment contracted by the Council for Responsible Nutrition and performed by Cantox Health Sciences International, Mississauga, Ontario (http://www.crnusa.org/CRNCantoxreportindex.html). The Cantox report reflects a true risk assessment that includes (1) evaluation of the evidence for a hazardous effect, (2) dose response relationship evaluation, (3) uncertainty assessment, and (4) identification of a dose that does not carry significant risk under specified conditions of use. The Rand report includes one important topic not addressed by Cantox—the benefits of ephedra.	No response
Given the animal toxicology data that includes well-documented toxicity at high doses, as well as many anecdotal cases from the drug and dietary supplement literature that point toward ephedrine or ephedrine alkaloid toxicity, examination of the ephedra adverse event report (AER) dataset for possibly, likely, or even "definite" causality seems to be a moot point unless the dosage that produced that causal case is identified and put into context with the recommended dosages. There are abundant examples among the essential nutrients of the absolute necessity of applying this principle. For a comparative example, a conclusion that vitamin A can cause liver damage may be true but is misleading, and actually harmful, as a generality. Clearly, scientists should recognize the critical importance of dose in any evaluation of causality, but not all policymakers or legislators, much less the general public, can be expected to do so. Thus, it is critically important to recognize and evaluate the dosage involved in any possibly or likely causal cases of adverse effects by ephedra. The Rand evaluation of risk stops a major and critical step short of the Cantox risk assessment in that little attention was paid to the dosage involved in adverse effects that might be casually related to ephedra ingestion. The absence of any significant dose-response consideration in the evaluation and conclusions is very clear in the Structured Abstract sections Main Results and Conclusions (page vi). This omission inexplicably occurs even though in the Methodology section (page 19), the Safety Assessment list of considerations asks the appropriate dosage question. This virtual absence of dose-response assessment in the entire report is reflected in the section on Attribution of Adverse Events (pages 20-21). In that section, the "dose question" is asked mainly in relation to the temporal relationship, not a dose-response quantitative relationship. Likewise, in the section on Causality Analysis of Case Reports, the three key points of the causality algorithm do not include any evaluation of the dose that produced the	A dose analysis is now included in this revision and this revision no longer assigns causality.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
adverse effect. A complete evaluation requires an answer to the following question: If the answer is affirmative on all three key points, what dose was involved? Paracelsus got it right some 500 years ago—"the dose makes the poison." Without consideration of dose, we can justifiably conclude that anything, indeed everything, is a poison. The necessity of adequate information to answer the dose-response question is exemplified by AER 13408, released by the Food and Drug Administration (FDA). In contrast to the labeled dosage of up to six capsules per day, the wife of the 26 year-old male in this case acknowledged to the FDA investigator that he "took a handful at a time, several times a day." This case is mentioned only to illustrate actual dosage may bear no resemblance whatever to labeled or expected dosage. Regardless of oral reports of specific dosage, the actual dosage should be assumed to be completely unknown, without confirming pharmacokinetic information or other objective information.	
Pharmacokinetics. There are two published studies of the pharmacokinetics of ephedra, both from the same laboratory (Gurley, references 74, 75). Unfortunately, the results are not consistent, but rather conflicting. It is strongly recommended that future research include a carefully designed comparison of the pharmacokinetics of ephedrine and two ephedra formulations, one comprised of powdered whole herb and the other powdered extract of the whole herb, in human volunteers. This should resolve the issue of a potential difference between purified ephedrine and the herbal products.	We agree this is an important line of research, but think this falls somewhat lower in priority than our first three listed recommendations.
The report notes the similarity between ephedrine and phenylpropanolamine (PPA) but fails to consider the relevance of the data with PPA to ephedrine. The suggestion that a trial similar to the one with PPA should be performed ignores the fact that most would consider the study to be unethical. The only ethical way to do the study would be to exclude patients at risk for cardiovascular events but that would make it impossible to accurately define the safety in an unscreened population of patients. Again, it would not be ethical to conduct a case control trial to quantify the magnitude of harm from a drug known to have the ability to cause strokes and heart attacks. The only reasonable recommendation from this analysis is that the drug (ephedrine/caffeine) has modest short-term efficacy and probable safety when used under medical supervision. If it is to remain available to the public it should only be used under medical supervision, i.e. dispensed only by prescription.	The case control study suggested is an observational study design that does not compel subjects to take anything, and in most situations starts after exposure has already occurred. We do not think it any more unethical to conduct this study than the PPA study.
The review of the safety of a drug with potentially rare adverse events must include a complete consideration of the polymorphisms of adrenergic receptors that have been	We judged this beyond the scope of our report.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
identified that could explain variable response and idiosyncratic reactions (Am. J. Human Genetics 2002: 70; 935-42). The polymorphisms that result in failure to develop tolerance are especially important to be considered. The section on metabolism should include some mention of the metabolic polymorphisms that result in deficient metabolism and accumulation of excessive drug levels.	
The evidence report questions were easily understandable.	No response
There was a paper published several years ago in the American Statistician, Unfortunately, I do not recall the authors' names. However they presented a particular method as a way of analyzing MedWatch Report data from the FDA. In brief, the method entailed creating a contingency table between types of events on the one hand, and drugs or substances ingested on the other hand. By looking for cells with larger than expected frequencies, one can potentially identify drugs with particular hazards. The authors might consider adapting this method to their data, or at least mentioning it.	We could not find this paper so we could not include this.
On page 54, the authors state that "a sensitivity analysis on only those studies scoring 3 or greater on the Jadad scale yielded a pooled estimate of effect size substantially lower than the main analysis…this difference…did not quite reach the conventional levels of statistical significance (p=.053)." In my opinion this is an extremely important finding. The literature on supplements for weight loss is riddled with a large number of trials of a very dubious quality. It is often difficult to know how to interpret such trials. It is easy to point out the flaws in these trials, but the obvious question is do these flaws matter? No study is perfect, and defenders of the claims companies make based upon these flawed trials are quick to point this out. The finding from the current authors suggest that such flared trials may be giving misleading answers. I believe that the authors should much more carefully describe this result and its implications and portray it much more prominently in the report.	We do note this prominently in the text but also note this effect was only observed for studies of ephedrine without caffeine.
On page 58 the authors state that there are data from the pharmaceutical literature that support the contention that patients taking pharmaceuticals outside of clinical trials may have a greater risk of certain adverse events than do patients selected to participate in clinical trials. The authors should supply one or more references supporting this statement.	Reference added.
On page 59 the authors state "Thus bias may exist, as the events we included were different in terms of type vs. those we had to exclude." It is unclear to me exactly what they meant by this. I suggest that they describe exactly what bias they are referring to.	We revised the text to try and clarify this point.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
Health Canada discourages its citizens from using ephedra for weight loss. They say they have at least 60 reports of adverse events. It's not enough to say that you haven't received them. You must get them and include them in your analysis. The US military discourages its people from using ephedra. A Col Mike Health, identified as an Army pharmacy consultant, states on the armymedicine.army.mil website: "There were 25 documented active-duty deaths of soldiers, sailors, airmen or Marines who had died and were coincidentally taking ephedra-containing products." You must get these and include them in your analysis. The American Association of Poison Control Centers collects information on human poison exposure cases, including cases attributed to dietary supplements. In 2000, 2.2 million cases of poisoning were reported to 63 centers. The Los Angeles Times reported on September 2, 2002, that the nation's Poison Control Centers collected 9,000 cases of ephedra poisoning since 1993. Where are these? You must include them in your analysis. E'Ola, a manufacturer of ephedra products, admitted in a lawsuit deposition in 1999 that it had received 3,500 complaints about ephedra from its customers that it had not forwarded to FDA. Where are these? They should be included in your analysis. There are at least 25,585 reports of adverse events associated with ephedra that you have not included	The EPC did request adverse event reports from most of these sources. We did not receive any. The EPC does not have the power to compel organizations to provide any data. Furthermore, the adverse events that were assessed leave us unable to conclude anything about causation. Therefore, our expectation is that the inclusion of additional case reports is unlikely to increase our certainty about a causal relationship between ephedra use & serious cardiovascular or neurologic events.
This report must deal better with the issue of dosages. Some people dismiss reports of ephedra-induced reactions as the consequences of over-dosing. Which events among the likely or possibly associated with ephedra use involved subjects taking only the recommended dosages?	It is not possible to tell which patients were taking the recommended doses.
Why were the criteria for high blood pressure set at systolic BP > 180 or diastolic > 105 mm Hg? More reasonable measures for serious or clinically significant hypertension would be to capture all cases of hypertension where pharmacologically management is indicated [class 2 and 3 hypertension] . This would be consistent with the definitions of serious adverse event as defined by MedWatch and in CIOMS [required intervention to prevent serious outcome….] Ascertainment of the rate and risk of clinically significant hypertension would be particularly critical in any safety assessment of ephedrine alkaloid containing products for use the general population where a “learned intermediary” is not required.	The criteria were set at a level sufficiently high that treatment would be warranted that day.
p. 50 Figure 4 brief data collection form for case report: what criteria were used to establish the categories under psychiatric [e.g., severe depression, psychosis]	The implicit review of experienced clinicians.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
On page 59 it is stated that "we did not examine the remaining 251 adverse events because the descriptors in the master excel spreadsheet were of conditions less severe….." The descriptors mentioned in the listing were of the 'adverse event as reported' [usually by a consumer] rather a diagnosis or precise description of signs and symptoms. Consequently, this description may be an unreliable or inadequate characterization of the adverse event, its severity or seriousness. It may be better to state that the remaining 251 AERs appeared to fall outside the focus of serious adverse events [deaths, cardiovascular, CNS, etc.].	We changed the text to reflect this.
Were other measures of variation included, e.g. confidence intervals or limits? Could these be used instead of having to impute standard deviations?	If possible we back calculated the standard deviation from other information include din the report. Otherwise, we reported the standard deviation.
The questions were clearly formed.	No response
The search methods were appropriate and resources were clearly documented.	No response
Inclusion of Non-Scientific Adverse Event Reports Invalidate the Integrity of the Study There is a potentially fatal weakness in the report in that there is no inclusion of a discussion on the peer-reviewed animal and laboratory research but extensive discussion of the non-peer-reviewed, non-scientific FDA Adverse Event Reports (AERs). The inclusion and heavy dependence on data that the General Accounting Office has already concluded was flawed is likely to nullify the scientific integrity of the report. The GAO report stated:	Animal and laboratory data were outside our scope.
While FDA's conclusions regarding the desirability of the proposed action may be valid, we believe these conclusions are open to question because of limitations and uncertainties associated with the agency's scientific and economic analyses. The GAO found that the AERs were poorly documented; that the FDA did not perform a causal analysis to determine if, in fact, the adverse events reported in the 13 AERs it used to set dosing levels were caused by supplements containing ephedrine alkaloids; and that the FDA indicated in its proposed rule that 10 to 73 percent of reported adverse events might not be related to consumption of dietary supplements containing ephedrine alkaloids.	We do not see how this critique of FDA is applicable to our report.
Have AERs ever been included in an AHRQ or RAND Evidence-based Center review before? An important hallmark of the evidence-based review or meta-analysis is the establishment of strict criteria prior to the review and an adherence to the established criteria once the review begins. Any deviation from criteria once the study begins may result in a flawed analysis and a loss of credibility	Case reports have certainly been included as a course of evidence in other AHRQ evidence reports, for example our own report on a "Best Case Series for CAM Treatments of Cancer".

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
The inclusion of AERs as part of this review appears to be a serious deviation both from what AHRQ requested and from the standard criteria used in conducting a meta-analysis. On page 4 of the draft, the authors conclude, "the majority of FDA case reports are insufficiently documented to make an informed judgment about the relationship between the use of ephedra-containing supplements and the adverse event in question." Devoting approximately 50 pages to AER reports in the report seems incongruent with the space devoted to descriptions of the peer-reviewed scientific data.	
There is also no explanation in the AER evaluation of products that were found to be illegally marketed as dietary supplements, which in fact were misbranded. Some of the early and most serious adverse events were from products that were adulterated with high doses of synthetic ephedra.	In fact, as our analysis shows, there were more deaths as a percentage of total AERs reported in the more recent data compared to the older data.
Several of the preliminary questions provided to RAND have not been addressed in the report. We expected a review of the literature to be included in the report: Questions about Dosage: What dosage of ephedra produce risk of CVD or other life threatening events? This may be because there is little or no data available. If so, it should be made clear in the report. The CANTOX report drew conclusions about a safe upper limit. While these were based on the results of a single study, they were somewhat corroborated by others. This is not to say that the CANTOX report is definitive.	A dosage analysis is included in this revision.
Also not addressed: Do ephedra-containing dietary supplement products alter physiologic markers of cardiovascular function?	This was addressed to the extent that RCT data in humans was identified. Blood pressure and ventricular tachycardia were two physiologic measures of cardiac function included in the analysis.
Adding AER analyses of ephedra AERs in the published literature, ephedrine AERs from the FDA's Adverse Event Database, and those for seizure and would make the report more complete and well balanced.	These have been included in this revision.
While it's useful to analyze the controlled trials for evidence of adverse effects, we're not likely to find significant effects in them because if adverse events were that common the studies wouldn't have been permitted in the first place. We must rely instead on case reports and adverse event reports for evidence. Therefore, every effort should be made to assemble all the credible case reports and adverse event reports associated with ephedra use. That was not done.	We disagree strongly with the contention that we did not expend every effort to obtain case reports. We have extensive documentation of our efforts to identify and obtain case reports for this analysis. Within the resources available to this project every possible effort was made.
The report should also make a better attempt at comparing the commonly reported adverse symptoms with those symptoms observed upon exposure to ephedra/ephedrine in controlled experiments. If the symptoms are consistent, or inconsistent, that's important to know.	This has been done in this revision.
The case reports cited are difficult to evaluate as they	We were limited by what was available in

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
contain clinical terms incorrectly used, incomplete descriptions and use an algorithm for causality that is impractical and unrealistic when using FDA reports. A vigorous documentation and search for better records at the time the case-reports were received would have improved the utility of the case reports. We regard the handling of adverse consequences as incomplete and unrealistic. The review by the Clinical Research and Review staff of the Center for Food Safety and Applied Nutrition of the Food and Drug Administration represents a more comprehensive and scientifically valid approach to reviewing adverse events associated with ephedra and ephedrine.	the files sent to us. It is not within the EPC scope to "search for better records at the time the case reports were received." We disagree that clinical terms are incorrectly used; in most circumstances we are reporting the clinical terms used in the source documents. Finally, if there was agreement about the best "scientifically valid approach to reviewing adverse events" then there would exist standardized methods for so doing and we would not have received the same level of peer review comments that we did.
Overall Evaluation. (i)The means used to evaluate the AERs is not clear. It is difficult to determine what role, if any, the TEP actually played in the review process. From the description given in the text, it would appear that most members of the TEP never even saw the AERs. (ii) It is not clear why an eight-week exclusion criteria was chosen the review of earlier safety studies. The exclusion of double-blind placebo control studies of less than 8 weeks duration resulted in the loss of valuable information about acute toxicity (and excluded most of the existing data not demonstrating toxicity). (iii) Important epidemiologic and scientific data has been omitted. This omission severely limits the value of this study.	The trials of less than 8 weeks duration were not excluded from the safety analysis and epidemiologic studies were outside our scope of work.
Question Formulation. Questions are well formulated and easily understood. All of the defects in the study, and there are many, stem from the methods used to answer the questions.	No response
Study Identification. Appropriate search criteria were not used. Not all episodes of ephedra/ephedrine toxicity are a consequence of chronic exposure. The exclusion of all studies of less than eight weeks duration may strengthen conclusions about effectiveness, but it weakens conclusions about safety. There are, for example, dozens of double blind placebo control studies where clinically relevant doses of ephedrine were found to have no effect on blood pressure or cause arrhythmias, even in asthmatics with heart disease. There is no reason to exclude such highly relevant data. Studies where ephedrine was compared to placebo should not be excluded just because they were not about weight loss or athletic performance. The scientific credibility of the report was weakened by the search strategy that was chosen. Clearly, the authors of the report assume that (1) all episodes of ephedra/ephedrine toxicity are a consequence of chronic exposure, and that (2) clinical trials of ephedrine have been limited to studies assessing the effect of ephedrine on weight loss. All these assumptions are easily shown to be incorrect.	There is no assumption that chronic exposure is necessary and we did not assume that trials of ephedrine have been limited to studies of weight loss. We do not agree that studies of safety in healthy adults are necessarily relevant to studies in obese individuals who are at greater risk for comorbid conditions.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
Data Synthesis. The analysis of the weight loss achieved by ephedrine versus placebo, and ephedrine plus caffeine versus placebo etc., is very problematic because one has assumed that the weight loss rate is high initially and subsequently lowers, so that the weight loss from months 3 to 6 is typically very small. It is therefore invalid to simply calculate the mean rate of weight loss as pounds weight loss per month when trials of very different duration are included. Those who are familiar with placebo controlled weight loss and weight maintenance trials know that most of the difference between the active and placebo arms is achieved during the first 3 to 4 months, and that the difference is subsequently maintained even up to 2 years. The way the data are handled in this report has therefore produced projections that severely underestimate the real efficacy of ephedrine and ephedrine plus caffeine. This has been carried over into the conclusions, where it is stated that ephedrine/caffeine is not as effective as other anti-obesity medications currently on the market.	We disagree. We tested whether weight loss was linear over this time period and we could not prove that it was not.
Data Synthesis. This must refer to Orlistat (the pancreatic lipase inhibitor from Roche) and Sibutramine (the centrally acting compound from Abbott). If one looks at the long-term of Orlistat ones sees that the mean weight loss difference between Orlistat and placebo after 6 months to 2 years are of the order of between 2-5 kg in all the large trials. Ephedrine plus caffeine produces at least an equivalent effect. For example: If the weight loss on an active compound after 3 months is 10 pounds more than on placebo, and this result is maintained also after 6 months, it is clear that rate of weight loss would be calculated as 10 pounds divided by 3 (=3.33) if the trial is stopped at 3 months. Whereas the result from a 6 month trial would give 10 pounds divided by 6 months (=1.67), which is exactly half of the weight loss. This issue should be addressed and the efficacy section should be revised accordingly. The way the panel has calculated the weight loss rate actually assumes that the weight loss rate is linear and that it continues at the same rate with prolonged use. Obviously, this is not the case.	We were careful to state in the text that our results could not be extrapolated beyond 4-6 months. We added data on other weight loss products for comparison.
Data Synthesis. I note that Astrup et al. International Journal of Obesity 1992;16:269-77, listed in the bibliography (accepted articles) as number 1, is not included in the analysis! The Danish double publication of this is the Quaade et al., listed as number 48 in the same bibliography. It is hard to see why the panel quotes the Quaade et al. publication in Danish, which a condensed version of the Astrup et al. paper, which I assume must be the paper the panel had taken the study information from in English. The panel has used pounds in the analysis of weight loss, but it would be more appropriate to use weight loss in percent of initial body weight, because the weight loss in pounds in not	We identified these two trials as reporting identical data, and the inclusion of either (but not both) should make no difference in the results. Our practice is to include the most informative article. For these reasons we note, a percent weight loss analysis has several limitations.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
independent of initial body weight. This may introduce a bias if the initial body weight and body mass index in the 2 arms were not comparable.	
Page vi. , paragraph 3; page vi, paragraph 4; page;4, paragraph 4; and page 30 last paragraph: Some may not classify anxiety, change in mood as psychiatric symptoms. Emotional/ mood adverse effects might be more appropriate.	Psychiatrists may disagree with the statement that anxiety and change in mood are not psychiatric symptoms. No change made.
On page 55, the authors state "the effects of ephedrine and caffeine appear to be additive. I do not understand the basis for the authors statement. Unless there is a 2 x2 design in which to have the opportunity to observe an interaction between ephedrine and caffeine and observe that no such interaction occurs, how can the make a statement of additivity? I believe what they mean to state is that there is an effect of the combination of ephedrine and caffeine combined that is greater of the effect of either alone. This is not that same thing as stating that the effect is additive.	The reviewer is correct. We clarified the language so that we do not imply the effects are “additive” in the arithmetic sense.
On page VI, the authors state that there are no data from studies of herbal ephedra-containing dietary supplement products without caffeine. This is not correct. There is at least one study. My colleagues and presented an abstract at the 2002 Experimental Biology meeting from suck a trial. Unfortunately, we did not present efficacy data. Moreover, I had thought the community sponsoring the study had provided the safety data to the NIH for this review. Although, it is not within my authority to release the data themselves, I can certainly provide the authors a copy of the poster presented if they do not have access to it.	Without efficacy data we cannot include this in the analysis. We did not receive this study in response to our requests to industry for unpublished studies.
There is no published study of the efficacy of ephedra without caffeine, but a large, industry-sponsored study was done by Coffey et al. at the 2002 Experimental Biology Meeting the authors reported that there had been no adverse events, but did not report on efficacy.	
I think you did an excellent job in collating the important available data.	No Response
The Danish study reports a 100% increase on post-exercise O2 consumption by 100%. Was that immediately after exercise and for how long, this seems like a very big increase are you sure this is correct? If you are unclear have Mary Hardy send me this paper and I will have a look at it.	On further review, we determined the Danish study was not relevant to the report since it did not report differences in performance between groups.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
As far as athletic performance goes, no studies were available on herbal ephedra and only a modest affect on "very short-term immediate performance" was observed with ephedrine, only when caffeine was co-administered. The report states that there was one study that assessed the effect of "sustained use of ephedrine on performance over time", an " reported that the addition of caffeine to ephedrine necessary to produce an effect on athletic performance." But in the structured abstract it is not made clear what the extent of the effect was. The report text states on pg. 57, however, that "a study116 published in Denmark concluded that aerobic training enhanced the effect of ephedrine on energy expenditure. After, 8 weeks of aerobic conditioning, ephedrine increased post-exercise energy expenditure by 100%. N. B. The reference numbers in the section are incorrect e.g. ref 116 cited above should be 43, 115:52	
Regarding the 1986 Denmark study – do you mean that ephedra increased energy expenditure during exercise, or that it increased energy expenditure after the completion of exercise? Please clarify.	On re-examination we determined this study should have been excluded, as it did not measure the effect of ephedrine on physical activity but rather on basal metabolic rate.
Were there a disproportionate number of case reports of adverse events that occurred during or after the performance of exercise training or physical activity?	Not assessed, and probably not possible to assess.
The conclusions regarding herbal ephedra for weight loss are reasonable and defensible as far as they go, but are too conservative. Saying that there is no evidence for sustained weight loss with use of these preparations for more than 3-4 months is true, but does not translate into conclusions that are useful for the clinician or regulatory agencies. The data presented and summarized support the use of herbal ephedra or ephedra + caffeine for weight loss.	Our charge was to present the evidence. Translating the evidence into clinical recommendations or regulatory decisions is specifically beyond the scope of the EPC.
Important parameters were properly identified and addressed, such as study population and design.	No Response
In general, the appropriate study parameters were examined. However, I would have liked more data on dose-response, especially with the efficacy trials	A dose response analysis is included in this revision
Most of the important parameters were systematically addressed.	No Response
Should a descriptive statement be made on possible non-statistical publication bias. For example, funding source for the published clinical trials may also bias the quality of the results (see BMJ 2002;325 (August 3):249).	This was added to the limitations.
There should be an expanded general statement regarding the medical exclusion criteria that are used in all the published clinical trials. This needs to be emphasized in the analysis and should specifically point out that many patients	This was added to the limitations.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
with underlying diseases (hypertension, heart failure...etc) were excluded from trials. The emphasis is needed due to common argument that safety data from clinical trials do not support the potential serious adverse reactions collected by FDA.	
None of the information provided addressed analytical methods with the ephedrine or herbal ephedra products used (e.g., certificate of analysis verification that the product used in the study met label claim) and other issues of quality.	We included this information where it was available.
I think you did a good job in defining the methods you would use in appraising the studies. Open label treatment following a controlled clinical trial is often thought to be a way to screen for safety, but I understand how incorporating that into your assessment might inject bias due to lack of a control group. Although caffeine and ephedrine has been evaluated in a controlled clinical trial for 6 months followed by an additional 6 months of open label treatment, I understand the statement that trials do not last more than 6 months refers to the double blind period.	No Response
The studies that were obtained for review were evaluated carefully. Objective criteria were established prior for inclusion into the planned meta-analyses. The evaluation of the case studies provided by FDA was also performed in an objective manner. The limited number of studies on exercise and athletic performance are presented objectively. The limitations of these studies is accurately noted.	No Response
It appears that a thorough search for relevant data was undertaken. I can find no evidence of bias or intentional inclusion/omission of data or search strategies.	No Response
Criteria for clinical study inclusion and exclusion were well defined and adhered to. Little bias seemed to be introduced by the selection process, at least for the clinical studies. All important relevant studies were evaluated.	No Response
The inclusion and exclusion criteria for the selection of articles is adequate. I am unaware that any crucial data is lacking; however, it would be helpful if Dr. Phil Waddington's data from Canada could be included in the final report.	We contacted Dr. Waddington but did not receive any data from him.
Should the search term "adverse reaction" be part of the strategy? If not, why not? Please see my comments on funding source of the clinical trial. In an ideal situation, a trial funded by a neutral party will probably yield the most un-bias information.	Generally, these terms act as "limiters" and would exclude studies if they were not tagged in this fashion. We prefer not to limit the search in this way. Since tagging articles is not always accurate and we did not want to prematurely exclude potentially relevant studies. We did include "adverse events" as text search terms.
There were appropriate inclusion and exclusion criteria for	This revision includes the results of a few

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
the studies selected for the meta-analysis and these were clearly stated and explained. There may have been some bias based on the studies selected, even though the inclusion criteria were clearly spelled out. This was mentioned in the Limitations section of the report. Efforts were made to identity unpublished studies and several studies were missed in the evaluation as they had not been received. The researchers indicated that they will be considered for future assessment. When this will occur is unclear.	additional studies that are relevant. We do not judge that any of the handful of requested but unretrieved articles are RCTs.
The discussion of the results of the weight loss trials makes no mention of doses. Doses should be mentioned either in the text or in the table. It is stated that "all of these studies had an attrition rate of greater than 20%…" Was the attrition rate higher in the active vs. placebo treatment group? This should be clarified. If the rates were higher in the active group, which I suspect, then a specific analysis should be done as to the cause of attrition.	The mean attrition rate was not higher in the active treatment group. This has been added to the results
Athletic Performance, second paragraph. Please clarify the duration of the exercise test. This is important because some drugs (like creatine) may produce benefit with short duration exercise (a few seconds) but not longer duration exercise.	The exercise tests varied in duration from short (weight lifting) to an hour or more (endurance).
The originally proposed key questions included inquiries regarding dosage levels if ephedra with respect to weight loss, athletic performance, and safety. The Report does not address dosage levels with respect to weight loss, athletic performance, and safety assessment (except for the mention of possible future research study). Clearly this is an extremely important concept, and issue, regarding to these materials, especially with respect to the review of case reported obtained from the passive AE reporting system. It is our view that the general omission of dosage considerations should be mentioned with regard to the AER case report reviews and that it receive some attention in weight loss and athletic performance assessments.	A dose analysis was added to the RCT portion of the report. We did not judge analysis to be possible in the case report portion of the report.
Are all of these studies single dose studies? What was the interval between taking the dietary supplement and the performance of the exercise test?	This has been clarified in the results.
The meaning of this sentence is obscure. What is meant by "enhanced mechanisms of heat loss?"	This sentence has been reworded.
See page 26: none of the weight loss studies were beyond 4 months. Therefore, I would rewrite the first paragraph of the Main Results to say the "longest published weight loss intervention was 4 months". If there was a study with a post-intervention follow-up then this should be stated as well. (See comments for page 5)	Three studies had 6 months of treatment, which was too few to perform meta-analysis on this time period specifically. For studies that reported only 6 month data, we included these in the "4 month" time point, and specified in the methods section in the report that this could include data out to 6 months. Hence the "4 - 6

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
	months" statements in the text.
On page 3 there is mention that 19 of the 48 controlled trials were excluded from pooled analysis because they had follow up of less than eight weeks. There should be some description of the findings of these studies in this report.	These studies were excluded as evidence on the advice of our TEP. To discuss them as evidence of efficacy would be inappropriate in our view. We did include them in our safety analysis.
It may not be standard format for the EPC's to cite patient numbers at this early point in a report, but to make sure readers understand what a small number of people have actually been studied in controlled trials, it would be helpful to include this near the beginning of the report.	This was added to the report in the safety assessment, where the possibility of a type II error is increased due to low numbers of studied patients.
Figure 2, question 18: Chemical analysis of ephedrine alkaloids was part of the quality review form, but data for individual studies are not provided. Given the variability of herbal ephedra, if chemical analysis was not performed in a particular study, does that call into question the results of the study?	We do not think so since the results for the ephedra studies were remarkably consistent.
"In order to improve health outcomes, long-term weight loss is necessary." Do you really mean that long-term follow-up would be necessary to determine health outcomes?	No, we meant maintenance of weight loss, since the relationship to health outcomes is known. We have clarified this.
Page 53: Results Section. Weight loss. It might be helpful to provide a table of 5 types of comparison studies indicating sample size in each trial and the power calculations for each. It is important to highlight when sample sizes are small and individual power calculations are insufficient.	Considering that we pooled data, we do not think the addition of our assessment of power of individual studies is very useful. We did include this in specific circumstances where it seemed warranted.
"Use of ephedrine, ephedrine + coffee, or dietary supplements containing Ephedra and herbs with caffeine is associated with a statistically significant increase in weight loss (compared to placebo) over relatively short periods of time (no more than a few months)." Please Clarify "(no more than a few months)". We assume you are not saying that the data show loss of effect after a few months, just that the studies don't extend beyond a few months.	The data cannot be extrapolated beyond a few months. We have clarified this. We earlier explained the reason for the "4 – 6" month designation.
It would be helpful to define what is meant by "sufficient evidence."	This is defined as statistically significant.
Several of the preliminary questions provided to RAND have not been addressed in the report. We expected a review of the literature to be included in the report: Questions about Dosage: I. What dosage levels of ephedra are necessary to achieve weight loss?	We now include a dose analysis in this revision.
When describing the efficacy studies, it might be helpful to include a table delineating the key elements of the weight loss studies: Dietary prescriptions, Description of subject characteristics, Mean weight or BMI at initiation of study, and inclusion and exclusion criteria for studies.	We considered this change or addition but decided there were already a great deal of tables and therefore we did not add this table.
Although ephedrine is the chemical drug that has been	We are sympathetic to this comment but

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
found to be effective in short-term weight loss and athletic performance, this result cannot be extrapolated to the herb ephedra that contains multi chemical components. In order to study the efficacy of herb ephedra, the amount of ephedrine present in the herb ephedra being tested must be precisely defined. This 'ephedrine' must be pure ephedrine and not, say, 50% ephedrine with 25% pseudoephedrine and 25% norephedrine or other related or unrelated alkaloids, as are normally present in herb ephedra. Ephedra is not ephedrine and vice versa, even though ephedrine is one of ephedra's active components. I personally don't see how one can generate meaningful results from a study using a material, such as ephedra, which is not clearly defined. There are just too many variables. Good science requires a well-defined test material. For example, we would never accept a single-chemical drug like cortisone with even 25% impurities when performing a clinical trial on cortisone. Why should we accept the chemical drug, ephedrine, present in herb ephedra, whose concentration can vary by 300% (from 30% to 90%, with the balance composed of other alkaloids)!? Until this problem (which is not insurmountable, as product definition criteria have been and can be set)5 is resolved, any studies on herb ephedra for weight loss or athletic performance (both based on ephedrine) will not yield meaningful results. It is possible that some of the papers the RAND report selected do clearly define the ephedrine content in the herb ephedra (though I seriously doubt it), the fact still remains that there are related and unrelated alkaloids also present in addition to ephedrine. For example, if a product containing an ephedra extract has been analyzed to contain specifically 20mg ephedrine per tablet/capsule to conform to the required amount of ephedrine for efficacy, what happens to the other alkaloids also present, which could easily be twice the ephedrine amount, or 40mg, making the total alkaloids content 60mg? This is a natural scenario unless made 'unnatural' by manufacturers or suppliers who take spent ephedra herb (from which all alkaloids have been extracted) or a token amount of ephedra herb and add the prescribed amount of ephedrine, thus rendering the product basically a single-component drug (ephedrine), formulated with inert spent ephedra or token ephedra herb as carrier/excipient. In this 'unnatural' case, the 'ephedra' product is basically an ephedrine drug dosage form and has nothing to do with the herb ephedra. I doubt there is any published scientific information on herb ephedra based on sound scientific definition of the test products containing ephedra. Furthermore, whatever reports available most likely have not clearly identified and characterized their test materials	believe that the consistency of our findings supports the decision to pool studies of weight loss and compare ephedra to ephedrine.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
hence rendering their findings of little value to us. Based on the current state of published information in this field, I don't believe we will be able to obtain meaningful conclusions relating to herb ephedra's toxicity through adverse events analysis alone or based on modern published experimental data. In case of the former, since no precise standards are required for commercial ephedra products, few if any of the reported adverse events can be reasonably traced to ephedra herb. In case of the latter, there are simply too few useful published reports whose findings are based on work that used well characterized and well-define ephedra. Modern medicine and traditional Chinese medicine are two parallel and distinctly different healthcare systems, each has its own merits and defects. While modern medicine is based on scientific experimentation, TCM is based on empirical practice and trial and error in humans over time. The latter has accumulated a vast amount of recorded information, including cautions and contraindications. This has been an ongoing process and it continues to accumulate data as TCM practice continues to generate them. It would be our loss if this valuable resource was not somehow utilized. Since herb ephedra has a long use history in traditional Chinese medicine with an extensively documented record (safety, cautions, contraindications, etc.) over a 2000-year period, this should be taken into consideration. Also, common TCM traditional practice should be heeded. For example, some of the adverse events reportedly due to herb ephedra alone may not be so at all, but rather due to the concurrent and inappropriate use of other common herbs such as Asian ginseng which is traditionally cautioned against use in healthy persons with a vigorous (yang) constitution and which has been known to cause serious toxicity, including death when used improperly. 6 If one combines the indiscriminate use of even such common herbal tonics as Asian ginseng with a relatively potent herbal drug like ephedra as dietary supplements, to be used daily with no prominent warnings or precautions, serious adverse effects are bound to occur. In order to meaningfully study or evaluate the safety and efficacy of traditional medicines such as ephedra (not ephedrine, the chemical), apart from ensuring that the ephedra has been well characterized and defined, we should also consider taking its historical record and its traditional use context into consideration as well as keeping an eye open to the simultaneous but inappropriate (outside of tradition) use of tonics such as Asian ginseng. Furthermore, we should keep an open mind to the possibility that the efficacy and safety of herbal medicines simply cannot be determined by Western 'hard' science alone. Common sense and well-documented historical use and safety data should constitute	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
part of the evaluation protocol.	
Regarding safety or adverse effects of ephedrine and herb ephedra, the two drugs need to be evaluated separately. With ephedrine, there should not be much of a problem because there must be copious amounts of data on the drug ephedrine, which can be accessed in various databases. However, with ephedra herb, it is quite different. Since ephedra has not entered the market through the usual drug-development-and-approval route, which would have generated toxicity data during that process, evaluating its safety as if it were a standard pharmaceutical (the single-chemical drug ephedrine) is not appropriate. There are few modern scientific or clinical reports published in the field.	For the RCT data, the numbers of patients studied with ephedra have been too small to assess adverse events without a high probability of a type II error. For the case report analysis, we did separate ephedrine from ephedra.
You are to be commended for the comprehensive search you conducted. Your methods seemed well-defined and unbiased.	No response
I think you did an excellent job of selecting articles using specific criteria and limiting bias. I know that one of the major issues prompting this review was concern regarding the safety profile of caffeine and ephedrine. In the United States this combination is sold in an unregulated fashion, so the only estimate of the denominator for adverse events is the number of doses manufactured. In Denmark, caffeine and ephedrine is a prescription preparation for the treatment of obesity. Orlistat and, before 1997, dexfenfluramine were approved prescription drugs in Denmark competing with caffeine and ephedrine. I assume that sibutramine is also approved in that country, but I do not know that for sure. It may be too late to at this point to include in this report, but information must exist for the incidence of reported adverse events to obesity drugs in Denmark. Although this is not a perfect way to assess safety, it might be useful to determine the relative incidence of serious adverse events reported with various prescription obesity drugs in Denmark. Based on conversations with individuals familiar with the Danish experience, I suspect that the safety of caffeine and ephedrine would compare favorably with sibutramine. One advantage of such an analysis is that one would be comparing alternative drugs for treatment of obesity in the same population. The second advantage would be a better estimate of the denominator based on prescriptions written rater than manufactured pills.	It was beyond our resources to obtain safety data (other than published data) from Denmark. We added to the future research that this would be a good study to undertake.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
Denmark Experience. Ephedrine/caffeine combinations are used extensively in Denmark for weight control purposes. There is a long history and experience that should be considered by RAND. Dr. Astrup has indicated that there are very few adverse event reports associated with such products in Denmark. RAND should contact the Denmark health authority, and/or Dr. Astrup, in order to obtain more information regarding these reports - and should include this information and Denmark experience in the final report.	
With regard to the adverse event reports, significant amounts of information were often missing; however that was not the fault of the authors, but rather a shortcoming of the MEDWATCH program. If anything the study highlights the inadequacy of voluntary reporting systems for adverse health effects and the confusion that results when a "systematic" analysis is attempted on such data. It is well known among the legal community and the FDA that thousands of adverse events have been reported to ephedra supplement manufacturers. Access to these reports might have affected the outcome of the present study. If anything these additional reports would have magnified the gravity of the public health threat attributable to ephedra-containing supplements. Moreover, Poison Control Centers throughout the country also log calls on ephedra supplements. Were attempts were made to access these additional resources?	We did not contact Poison Control Centers. We did include in this revision an assessment of the reports made to one manufacturer.
There was a thorough search of relevant articles using 9 electronic databases. Both national and international journals were included and the searches appeared to capture most of the relevant studies. The majority of the accepted articles for the meta-analysis were from the U.S. It seemed that 3 were from Germany. There were no studies from Asian journals.	No response
Table 19, I do not understand the point of this table or the conclusions being drawn. Please clarify for the simple minded.	We have added text to explain this table. The point is the later cases, that we did not have access to, contained proportionately more deaths.
Timing of last ephedrine (ephedra) dose? If it was > 24 hours, because of the relatively short half life, one would not expect to detect much or any in the blood at autopsy. Is timing of last dose with a tox screen negative test taken into consideration when determining causality?	Such cases were not reviewed, so a negative toxicology screen would not even have been assessed.
Insert percentage next to the # of adverse events for easier direct comparison of the placebo and intervention groups	We do not feel this comparison is justified due to small sample sizes, that is why we did not perform meta-analysis.
Instead of the 5x4 (nxn) test, it may be better to perform the chi-square test on event type vs. data type, ie. death (vs. other) x data type (2x3), stroke x type, etc.	We are not sure what this comment applies to.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
There were clear criteria used to select studies for inclusion in the report.	No response
The appropriate criteria were used to assess the studies on efficacy of weight loss with ephedra or ephedra/caffeine.	No response
The initial draft failed to include a review of the published case reports of adverse events. Given that these case reports were published in peer-reviewed medical journals and are prepared by medical professionals, they are likely to contain data that is more complete, accurate, and of scientific merit. It would seem more appropriate for these reports to have been evaluated and included in this report, especially since no other entity has conducted a review of them, than to once again include an evaluation of the FDA's evaluation.	These published case reports are now included in this revision.
Currently one must read to page 54 to get an answer to the key question, "We interpret this data as indicating the use of ephedra is associated with a statistically significant 1.3 pounds of weight loss per month more than is associated with placebo for up to four months of use" and "We interpret these data as indicating that the use of ephedrine and caffeine is associated with a statistically significant 2.2 pound weight loss per month more than is associated with placebo up to four months duration." This should be in the very beginning of the report.	This information is in the appropriate place for an EPC evidence report.
The 19 efficacy studies not pooled in the analysis because they had a duration of less than 8 weeks, and the 9 studies eliminated for a variety of reasons should be accounted for in the document, and any serious adverse event reports described should be included in the report.	These studies were included in the safety analysis.
“Even in aggregate the clinical trials only enrolled sufficient number (how many?) of patients to detect a serious adverse event rate of one per one thousand.” And again on page 58: “For studies of ephedra, there was only sufficient statistical power in aggregate to detect a rate of serious adverse events if three in one thousand.” For the reader, it would be helpful to know how these event rates, 1/1000 and 3/1000 compare with those reported in the literature for drugs in the same usage category as ephedra (three billion servings in 1999), i.e. to HRT (approx, 3.8/1000 women has an MI or developed breast cancer), or to event rates for aspirin and GI bleed.	We added that these events would be classified as “rare.”
References 104 and 108: The Nasser study (ref.108) is the same as the first Boozer study (ref.104).	This duplicate study has now been removed from the pooled analysis.
Under Bibliography Accepted Articles, pp 135 –137, references 12 and 41 are the same study [12=published study, 41 = published abstract], leaving 4 studies that assessed the effect of ephedra + herbal caffeine. Data reported in Chapter 3, Results, including Table 14, will need	This duplicate study has now been removed from the pooled analysis.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
to be corrected and re-analyzed.	
Monthly Weight Loss. In your results section you report the data as "Monthly" weight loss. The rationale for this escapes me. Weight loss with all medications slows with time and a plateau is reached between 4 and 6 months. Thus, the most rapid weight loss occurs in the first month. In trials that last 6 months, the only weight loss will be slower than one that lasts 2 months. How do we compare them with this criterion?	We tested, given the data available in these trials, whether weight loss differed across the different months. We could find no evidence that it did. Therefore, within the limited time frames of these trials (4 months), we included all relevant data points, as it increased our statistical power.
Miscellaneous Comments. On page 11, the draft report acknowledge that the "estimate of use of ephedra containing products may be low." Despite this, the report fails to emphasize denominator-related concerns associated with adverse event report reviews. This major scientific weakness needs to be acknowledged as part of the adverse event report analysis in a manner similar to that used in prior AHRQ studies (such as the Garlic Report). On page 29, the draft report indicates that although certain studies did not record any data for certain even category or indeed any adverse events at all, such studies were not included in the adverse event meta-analysis as RAND did not assume zero observed events if a study did not mention a particular type of event. Is this approach consistent with most scientific reviews?	We did assume zero events for serious events like death or stroke even if they were not recorded in the RCT. We did not do so for other events because we could ever know whether those events were sought by the investigators if they were not recorded. In other words, we did not assume zero for the entire universe of adverse events, only for those specifically mentioned and sought and recorded as zero. This is consistent with most high quality scientific reviews.
The Garlic Report. The draft report, as noted above, is in many ways inconsistent with prior AHRQ reports that address adverse event case reports. The AHRQ Garlic Report (Garlic: Effects on Cardiovascular Risks and Disease, Protective Effects Against Cancer, and Clinical Adverse Effects), and statements contained in the Garlic Report, should be reviewed by RAND as a potential model for the ephedra report - particularly in the manner case reports are assessed and described. For example, the Garlic Report provides the following with regard to adverse event reports and confounding factors: Adverse effects of oral ingestion of garlic are "smelly" breath and body odor. Other possible, but not proven, adverse effects include flatulence, esophageal and abdominal pain, small intestinal obstruction, contact dermatitis, rhinitis, asthmas, bleeding, and myocardial infarction...The frequency of adverse effects with oral ingestion of garlic and whether they vary by particular preparations are not established...Furthermore, the causality of the adverse effects was not clear, except for the breath and body odor, and the expected frequency of adverse effects was not determined... In addition to the RCTs, 73 studies were found that addressed diverse effects. Most (97 percent) were case reports or small case series (Evidence Table 9). The literature reviewed gives a limited picture of adverse effects	The inclusion criteria and reporting of studies in the Garlic Report were shaped by their TEP and their Partners. The inclusion criteria and reporting of studies in our ephedra report were shaped by our TEP and our Partners. There is no requirement that these reports be the same in inclusion and reporting.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
attributable to garlic for many reasons. First, searching for studies that report adverse effects is difficult. Many studies may mention adverse effects in passing, but do not use adverse effects as a key index word or in their abstracts. If these studies do not otherwise meet selection criteria in a review, they will be missed. Second, in most case reports and case series, adverse effects cannot be directly attributed to garlic because chance, coincidence, or confounding factors could have been responsible for the adverse effect. For example, alternative causes of reported adverse effects were possible in 22 percent of the reviewed studies and could not be excluded definitively in 69 percent. Third, case reports and case series may miss delayed adverse reactions because such associations are more difficult to make than those that occur immediately after garlic is administered. Fourth, although case reports and case series can provide qualitative information about the nature of an adverse effect, incidence cannot be estimated from such evidence... The frequency and severity of adverse effects that are related to garlic should be quantified. Whether adverse effects are specific to particular preparations, constituents, and doses of garlic should be elucidated. Whether certain adverse effects are unique to particular types of garlic exposure (e.g. inhaled, oral, or topical) should be clarified. The most serious potential adverse effects of garlic that have been cited are complications related to bleeding. Whether particular preparations and constituents of garlic affect physiological parameters related to bleeding such as platelet adhesiveness, prothrombin time, and partial thromboplastin time, as well as whether particular preparations lead to clinically significant bleeding, warrants more study. (emphasis added).	
The limitations of the review process are not stated. You have adverse reactions and inconclusive studies on the effectiveness of ephedra/caffeine. That leaves us with insufficient information to make an assessment of either safety or efficacy. In the data synthesis, the impression is given that there is more precision than can be justified based on the nature of the data.	We disagree with regard to the reviewers comments on precision. Precision is determined mostly by sample size and number of studies. We believe our pooled results adequately reflect the degree of precision the data allow.
Possible bias of the report due to members of the TEP and literature captured. Limitations and quality of the studies and short-term studies used in the systematic review. Combining the systematic review by meta-analysis with the analysis of AERs. No conclusions for herbal ephedra and weight loss. The AERs evaluated were limited to those provided by FDA. AERs from the studies were not reviewed. Report did not emphasize any potential benefits of ephedra/ephedrine and weight loss.	The AERs from other sources are now included in this revision. We disagree that the report did not emphasize the potential benefits of ephedra/ ephedrine use and weight loss. The other limitations are noted.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
Reference 29 can provide a national estimate of 2.5 million individuals using ephedra wt loss products (during 1996-1998), which is probably an underestimate since 33% of one-states respondents did not know that their nonprescription weight loss product contained ephedra.	The imprecision noted by the reviewer even in this one estimate is, we believe, good reason to avoid its use in trying to calculate a rate using case reports.
Some of the questions guiding the evidence report were not answered but are available in the literature search and data collection. For example, the question regarding the dosage level of ephedra necessary to achieve weight loss was not answer. There is no summary statement on the dose of ephedra or ephedrine other than Evidence Table 1 and 2.	A dose analysis is included in this revision.
Conclusions regarding the efficacy of ephedra-containing supplements in promoting weight loss and the enhancement of exercise performance were supported by the available data.	No Response
The conclusions are clearly and concisely laid out and are consistent with the evidence presented.	No Response
The conclusions of the efficacy of ephedrine and related compounds are valid for the short term studies evaluated and conclude that longer term studies need to be conducted. The remaining conclusions are valid and appropriate.	No Response
For the efficacy and minor adverse effect evaluations, the evidence does support the conclusions.	No Response
Finally, there was no demonstrable effect of sustained ephedrine supplementation on strength training.	No Response
According to the report, caffeine appears to enhance the effect of ephedrine yet there is no mention of assessing caffeine in the diet. This would have to be done in a case-control study. It may be appropriate to add this separate bullet under conclusions.	The need to control for caffeine intake was added to a bullet in the conclusion.
Bullet 1: Conclusion of "sufficient evidence" should be tempered with reference doses used and duration of treatment in the studies.-Bullet 2: This is just a statement not a conclusion. It would become a conclusion by adding that out of 1848 cases known, 1344 were selected for review and of these, 158 showed serious adverse events. Within this subset of 158 reports, 11 were identified in which ephedra was possibly causal. Conclusions section should be able to stand alone. This would require adding more detail including doses and duration.	Changes made to bullets.
It is generally acknowledged in the field of obesity research that any study with less than a 2-year follow-up is misleading and not relevant to the evaluation of the anti-obesity modality under investigation. The reason for this is that most individuals regain lost weight on any diet, drug or weight loss program after 24 months. There were no long-	We believe the text is clear that these results cannot be extrapolated beyond four months. The data support a linear relationship for weight loss over 4 months in these studies.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
term weight loss studies of ephedra or ephedrine reviewed in this report, yet a mathematical formula for weight loss associated with ephedra was presented. This seems highly speculative and hard to defend. The conclusion that short-term weight loss can be achieved with herbal ephedra and caffeine rest on studies of herbal concoctions that are not adequately characterized, chemically and pharmacologically, to permit pooling of studies; The attempt to quantitate the weight loss per month attributed to the use of ephedra-containing herbs is invalid. Further, the model assumes constant weight loss over time, an unlikely outcome.	
The need to add caffeine to ephedrine to produce any measurable degree of enhanced athletic performance suggests caffeine alone may suffice to achieve this effect, which, in any event, is evident for only short time periods.	The reviewer is incorrect, since one study reported in the athletic performance section compared ephedrine, caffeine, and their combination, and reported only the ephedrine/caffeine combination produced an effect. This result refutes the reviewer's hypothesis that caffeine alone may suffice.
It seemed that the decision to review FDA adverse event reports produced very little useful information and was almost a duplicate effort given Haller and Benowitz study. It may be more useful to use the same strategy on the case studies reported in the literature.	The literature cases are now included in this revision.
The bullet point "Scientific studies (not additional case reports) are necessary..." should be deleted as this is not the conclusion of the Adverse Consequences but rather Future Research, which has been stated already.	We think it important to also include this as a conclusion.
Dr. Leung made the point that there were thousands of Chinese literature on the Ma Huang, and if it wasn't safe the literature would say so. The literature comments that some people should not take it, but there is a lot of information to show that it is safe. Dr. Leung also made a point on the credibility of this information by stating that for other herbs the literature shows they are not safe.	We do not disagree that there may be extensive Chinese literature on MaHuang but we did not find controlled trials in our literature search, nor was this literature offered to us by any of the many reviewers of this report. Furthermore, we believe that there is ample evidence to support that the most valid conclusions come from properly designed hypothesis testing studies, not a collection of anecdotal literature, either supporting or refuting safety.
I disagree with the safety conclusion of this report for two reasons: 1. The clinical trials excluded patients at risk, thereby reducing the study's ability to detect harmful effects of the drugs. 2. The totality of prior pharmacologic information was ignored in the analyses of the FDA cases. The consistency between the type of events reported and the known actions of ephedrine to increase heart rate and blood pressure must contribute to the assessment of causality. The similarity of	We added to the limitations the issue of select patient populations. It was outside our scope to assess other chemicals.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
the reports seen with similar chemicals, e.g. phenylpropanolamine, must be given consideration. The cases that the authors classify as "possibly caused" by ephedrine/caffeine I would classify as "probably caused by" ephedrine/caffeine.	
I believe the second bullet under "Adverse Consequences" (page 112) is unclear and unqualified. This statement reads, "There have been a great number of adverse event reports filed with FDA regarding herbal ephedra-containing dietary supplements." I find this statement unqualified and unhelpful. A great number of reports compared to what? – Total dietary supplements sales? The number of AER's reported for other dietary supplements? The statement should be better qualified, in my judgment.	We believe 2000 is "a great number" by most people's definition and have not made any changes.
The first bullet on page 113 is a very key issue and I believe deserves further comments. I agree with the conclusion that, given the rarity of serious adverse events associated with ephedra, properly designed case controlled studies would be appropriate. However, I believe it will be difficult to develop such a properly designed case controlled study, as the underlying factors (that appear to be idiosyncratic) are not well understood. How, then, would a case controlled study be designed to take such unquantifiable factors into account? This is precisely the continuing problem in deciding how best to approach both the regulation of and further scientific research into ephedra	It is beyond the scope of an EPC evidence report to go into such details of study design. We note, however, that others have made detailed proposals to governmental agencies for just such a study.
The conclusions seemed fair and stated appropriately. I would suggest adding a section answering the questions you posed at the start of the report in a summary fashion.	We organized our bullets to follow the order of the questions.
Most of the studies reviewed were on synthetic ephedrine and weight loss, therefore, a relationship between herbal ephedra and weight loss cannot be made and this appears problematic. The clinical data that were examined only included ephedra in combination with another herbal stimulant. While it is true that many weight loss products contain a combination of ingredients, not all do (NNFA's database of ma hang or ephedra reveals that almost half of the products do not contain another stimulant). A concern is that the conclusions drawn in the report may be applied to all ephedra products, regardless of use and regardless of whether other ingredients are present.	We now include one RCT of ephedra without caffeine. We have limited the conclusions to only those concoctions studied.
It seems important to mention that apparently healthy individuals have died from the use of these products, possibly related to exacerbation of previously undetected disease. And, because these products are available without prescription, and not even regulated as OTC drugs, the risk associated with unsupervised use are potentially greater than with drug formulations of ephedrine.	We have added to the conclusions and limitations both of these points.
The conclusions of adverse consequences are internally	We disagree, since the finding of serious

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
inconsistent. Once the limitations are listed, then the number of serious adverse events (death, stroke, and myocardial infarction) should not be enumerated. This is the same thing as saying that results of a scientific study are not statistically significant and then still enumerating them in the conclusion. The most important conclusion is listed second to last on Page 113 and should be moved up to the front of the conclusion section of the report. That is, scientific studies (not additional case reports) are necessary. I disagree that a case-control study would be the next step. Rather I would recommend strongly a prospectively randomized controlled study design with appropriate data safety and monitoring in place. A population-based study will have the same drawbacks as the phenylpropanolamine study and will have the same risk of spurious associations rather than cause and effect relationships. An intervention study and not an epidemiologic study is needed to clarify the situation. As already stated, it is clear that the evidence does not support the conclusions. The adverse event reports are that they are. The attempts to connect them to the use Ephedra/Caffeine remain unconvincing both in this report and in the New England Journal of Medicine article.	adverse events in otherwise healthy young people is a cause for concern. We also disagree that the phenylpropanolamine study found a “spurious” association. We note that a case control study is the accepted study design to quickly assess a possible relationship between an exposure and rare adverse events.
Specific issues related to ephedra are not addressed adequately, and notions for which there is no proof are presented as if they were accepted scientific fact. For example, in the Pharmacology section, the report states "ephedrine increases peripheral resistance and can lead to a sustained raise in blood pressure...Elevations in blood pressure appear to be dose dependent in humans. However, does under 50 mg do not always result in increased blood pressure." The report fails to state that the sustained raises seen in hypotensive patients occur after the intravenous, not oral administration of ephedrine. The only citation for the dose dependency of an ephedrine-related rise in pressure is a review article, and that article does NOT say that ephedrine causes hypertension! It says, ephedrine and caffeine cause a greater increase in systolic pressure than ephedrine alone, that there is no effect on diastolic pressure, and that hemodynamic effects are transient. The statement is quoted out of context and is therefore misleading. The way the sentence is written, readers would be likely to assume that, even though "doses under 50 mg do not always result in increased blood pressure, a series of double-blind, placebo control trials have shown that at most the effects of oral ephedrine on blood pressure are negligible (as opposed to intravenous dosing used b anesthesiologists). A partial listing of some of these studies is cited here [10-23]. The lack of effect on blood pressure is even supported by the list of TEP "accepted articles" cited in	We clarified that the use of ephedrine to raise blood pressure intraoperatively is with parenteral use.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
the bibliography. All of the cited articles were from controlled clinical trials, and none reported clinically significant blood pressure elevations. These studies should be included as part of the RAND review, and the statements regarding increases in blood pressure should be substantially revised. Indeed, if the authors of the report cannot cite published, double-blind placebo control studies showing that taking oral ephedra/ephedrine significantly increase blood pressure, that that claim should not be included in the report.	
Suggested directions for research were provided, but they are narrow in scope and may suggest "gaps" that do not really exist. To date, more than 2000 ephedra/ephedrine users have been enrolled in clinical trials. Given the consistently benign results of all the previous trials, are still more trials needed before the issue is put to rest? On the other hand, cutting edge issues in obesity-related research are completely ignored. Does ephedrine interact with uncoupling protein? Does use of ephedra supplements have any effect on the production of inflammatory cytokines by adipose tissue? Or upon lepton homeostasis? On Lipotoxicity? If supplement manufacturers are to be believed, they have thousands of testimonials from satisfied users reporting weight losses of 50 pounds or more. Why not study these individuals and compare them with other product users who were unable to achieve weight loss? Having identified a population of proven ephedra responders, and non-responders, comparing the two groups medically, chemically, or genetically, may provide some truly useful insights.	We would ask the reviewer whether 2000 successful airplane flights mean that airplanes never crash. The point is that 2000 studied patients is insufficient to detect a rate of 1/1000 events, and even rare events, when multiplied by the millions of people who may be consuming ephedra, add up to numerous serious adverse events, if such an association exists.
The report draws conclusions about efficacy and safety that are not sufficiently supported by the data. As I have pointed out below the efficacy of ephedrine/caffeine is underestimated due to the incorrect method of analysis. In addition, in my view a number of shortcomings in the safety assessment tend to exaggerate the adverse events. My overall conclusion is that in several aspects the report needs some important revision. This includes the identification of studies, selection of studies for efficacy and safety. The data handling is also inadequate in some aspects. Consequently the report's overall conclusions are not supported in the current version and I believe that the revision suggested below will produce a substantially changed conclusion.	No response to this general comment. Specific response made to specific comments. We disagree that the identification of studies, selection of studies, data handling, etc. are inadequate. We also disagree that our results underestimate the efficacy of ephedrine/ caffeine, or that our analysis is incorrect.
Weight Loss. In the first bullet it is stated that compounds produce weight loss over relatively short periods of time (no more than a few months). This is misleading as there are trials for a duration of 6 months. The same applies for the 3rd bullet where the expression "short-term weight loss" is used. Bullet 6 is outrageous, here it is concluded that	Six months is still “a few months” when one year data are considered necessary by FDA to assess pharmaceuticals. The data about phentermine are taken directly from the graph in the cited reference. We have added data about weight loss using

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
ephedrine and ephedrine plus caffeine produce a weight loss somewhat less than the effect reported for FDA approved pharmaceuticals for weight loss. The panel has used phentermine as an example and state that the effect is "reported at about 20 pounds of weight loss at 6 months". This is certainly not the weight loss produced by phentermine above placebo, the weight loss produced by phentermine from baseline including a diet. For comparison one can take the Astrup et al. study from 1992 where the weight loss in the ephedrine plus caffeine arm was about 16 kg. But of course, the weight loss in the placebo arm must be subtracted, giving an additional weight loss produced by the compound of 3.6 kg.	other pharmaceuticals.
Adverse consequences. Again, this reviewer suggests that the open trials should also be included. In the first bullet it is stated that it is not possible to separate out how caffeine contributes to the side-effects. This is actually possible. In the Astrup et al. in International Journal of Obesity in 1992 there was a separate caffeine arm in the 6 months trial. Side-effects are shown in one of the tables in this paper, and here it is clear which side-effects can be attributed to caffeine.	Our statement refers to the data included in our review, which was restricted to RCTs and CCTs.
A long-term study of comparing ephedra + caffeine with ephedrine + caffeine at promoting weight loss and adverse reactions. Expand the pharmacokinetic study of ephedrine (pharmaceutical preparation) and ephedra (botanical preparation) absorption (as part of the dose response studies).	We think this is already subsumed under the first bullet point.
The most basic and important aspect of any research in natural products and in the reporting of findings is the characterization and clear definition of the products or materials being studied. Without this, research findings cannot be reproduced and thus are meaningless. In our case with ephedra evaluation, we not only need to set criteria for selecting articles for study, but also be sure to clearly understand what it is that we want to study – ephedra herb or ephedrine. The whole field of 'ephedra' in weight loss and athletic performance is twisted backwards. The herb ephedra has never been traditionally used for either function, nor has it been first clinically reported (before ephedrine) to have these effects. Only the drug ephedrine has. Yet ephedra is being used for these effects and is touted as natural ephedrine and thus safer. So far, there has been no credible clinical evidence that ephedra itself (and not synthetic ephedrine in an inert 'ephedra' carrier) has these actions, despite the conclusion reached in this report. It is worthwhile to reevaluate references 104-108 to determine whether the 'ephedra' used in those studies was actually natural ephedra containing the total complements of	The included ephedra studies said they assessed herbal ephedra. We agree with the reviewer that a future study of ephedra should adhere to these recommendations.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
ephedrine alkaloids in their natural proportions. If not, was it composed mainly of synthetic ephedrine formulated with carriers (e.g., token ephedra or exhausted ephedra marc) into an 'ephedra' dietary supplement that contains little or none of the usual complements of other ephedrine alkaloids? If it is the latter, then this 'ephedra' herb has no place as a dietary supplement in weight loss or athletic performance. Such 'ephedra'-containing products should then be more appropriately placed under the OTC-drug category which would eliminate much of the problems currently associated with its abuse and also would save us taxpayers much money trying to resolve these problems. In order to show ephedra herb (not synthetic ephedrine) to also have efficacy in weight loss and/or athletic performance, it is necessary to first characterize and standardize ephedra products to specific amounts of ephedra's alkaloids in their natural proportions, before subjecting them to clinical trials. This would eliminate the drug ephedrine being formulated into a dietary supplement to bypass the OTC-drug regulations. Unless ephedrine-containing products (whether natural or synthetic) for weight loss and athletic performance are all considered OTC drugs, adulterated, poorly characterized, and undefined dietary supplements containing ephedra herb will continue to be sold and abused. We need to set standards for manufacturers to meet and follow in order to be able to label and market their ephedra-containing products as dietary supplements. As I have repeatedly stressed, the most important aspect of any research in herbal medicines/supplements is characterization and precise definition of the test materials, without which no meaningful and reproducible results can be achieved, no matter how well designed and how well executed the rest of the research. In order to reduce the continued accumulation and dissemination of ambiguous, meaningless, and useless research data in the natural products field, we urgently need to set criteria for the characterization and precise definition of test materials at three levels: (1) research; (2) publication; and (3) abstracting, indexing, and data input into databases. Such criteria have been published and are available.	
The conclusion that "a properly designed case control study would be the appropriate next step" would require a study so large, lengthy and expensive it is unlikely to ever be funded or completed. (If one assumes a prevalence of use of ephedra of 1%, an alpha of 0.05, a power of 80% and if one seeks to detect a doubling of risk, then 2,400 stroke cases and 2,400 unaffected controls would be required. See Schlesselman, Case-Control Studies, Oxford University Press, 1982). The presumed benefits of ephedrine, should they exist, (improved athletic performance, enhanced	We agree that a properly designed case control study would need to be large (perhaps not as large as this reviewer conjectures). However, the PPA case control study was also large and was successfully completed, and we do not favor substituting opinion for science when the scientific study is feasible.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
energy, short-term weight loss) are likely due to the sympathomimetic effects of the drug and the adverse consequences are predictable as they were with isoproterenol, amphetamines and fenfluramine. Since sympathomimetic drugs have never been shown to result in safe and sustained weight loss, it is highly unlikely that this will be the result of long-term controlled trials of ephedrine and weight loss. But the known adverse effects of the drug and its congeners are now well characterized.	
First, you need to organize your listing of sources of adverse event information better so that readers can see which sources you have included and which you have not. Right now, it's difficult to follow what you have gathered. A table would be ideal, with each row specifying the source, the number of complaints, the number of deaths, serious injuries, non-serious injuries, and the numbers of each of these you concluded are likely or possibly related to ephedra use.	All of the serious adverse events came from FDA data, so in the draft report such a table would have no meaning. In the revised report, such a table is included.
You do not provide clear evidence of the pharmacological and pharmacokinetic equivalences in the use of the herb or of ephedrine alone. The complexity of the Phytochemistry of Ephedra (p 13) reinforces the point that the whole herb contains other alkaloids that are likely to be active or to qualify the effect of ephedrine. More should be made of this deficit at various points in the text.	We agree that there is likely heterogeneity in the herbal concoctions, but note the striking consistency of our findings relative to amount of ephedrine alkaloid and weight loss.
There is insufficient evidence that dietary supplements made up of the herb Ephedra spp. have any of the effects or risks identified for the alkaloid ephedrine.	We disagree and believe the data speak for themselves.
The future research directions proposed are reasonable. One addition may be to recommend examining the interaction of ephedrine and exercise training on weight loss and adverse events. Is there some interaction between physical activity and ephedrine?	This was added to the future research.
If the majority of ephedra users are seeking long-term weight loss, it would be very helpful to better understand the age, gender, race, temporal use patterns, concomitant drug use and other risk factors associated with ephedra usage. These points are underdeveloped and are, I believe, central to understanding safe and appropriate use of ephedra in the general population.	We agree in principle with this comment, but think it might wait until there is better evidence of sustained weight loss in any population.
I agree with the suggestion to analyze and compare the adverse events reported for ephedrine to ephedra. I would also add PPA.	This report now includes an assessment of ephedrine so this bullet has been eliminated.
The suggestion to consider a dose response study to determine a minimum effective dose of ephedra would be difficult, at best. Effectiveness criteria should be identified in these comments.	While it may be difficult, it is certainly feasible and consistent with the way FDA evaluates pharmaceuticals for weight loss.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
There remain open questions whether there is a difference between synthetic ephedrine and naturally extracted ephedrine alkaloids. It would be a very useful research activity to analyze the branded products which are identified in AER's using both AOAC and USP methods to try and determine whether synthetic or naturally occurring ephedrine alkaloids are present.	Agreed.
Future Research. The "numerous gaps" in the literature regarding the efficacy and safety of ephedra is a central point.	No response
FDA has recently taken action against six companies selling synthetic ephedrine as dietary supplements. This is not permitted under current law but, unfortunately, synthetic ephedrine dietary supplements are being sold to the general public.	No response
As a final thought, the inadequacy of FDA's adverse event reporting system is clear as it relates to ephedrine. I believe it is appropriate for RAND to recommend that, with respect to ephedra products, FDA/CFSAN's process and systems to evaluate and capture ephedra-related AER's be thoroughly reviewed, as it is likely that continued reliance will be placed on this system, despite its weaknesses.	This is not a proper role for an EPC evidence report and we decline to make such recommendations.
A chapter was devoted to future research. The researchers addressed the gaps in a variety of areas and suggested meaningful recommendations for further research. Most significant is the need for long-term studies of ephedra/ephedrine and weight loss and athletic performance including both anaerobic and aerobic exercise. This was emphasized in this chapter.	No response
It might be beneficial to explain the pathophysiology of how ephedra/ ephedrine can contribute to an acute cardiovascular event in the setting of mild-moderate underlying disease. For instance, in individuals with non-critical coronary artery disease, ephedrine alkaloids can produce platelet aggregation with resultant thrombus, increased myocardial oxygen demand, and cause vasospasm, all of which can result in decreased perfusion and ischemia. The same contributory actions could be expected in individuals with congenital cerebral aneurysms, and other underlying abnormalities in the cardiovascular system.	While we agree that biologic rationale is an important criterion when assessing causality, we think that direct evidence of an association is most important, and therefore recommend a hypothesis-testing study.
The implications for future research are fairly stated. I would suggest adding the analysis of safety or adverse events reports in Denmark comparing available prescription obesity drugs, since Denmark uses caffeine and ephedrine as one of it's approved prescription drugs for obesity.	This was added to the Future Research Section.
Are implications for research discussed? Not adequately. The major implication of the research is whether the	The role of the EPC is to report the evidence, which we believe we have done

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
analysis will be adequate to advise the FDA and HHS on whether they should take action to protect the public. This aspect of the analysis is ignored.	The judgment about the adequacy of the evidence to make a judgment is not a role for the EPC.
What directions for future research would you recommend based on this report that we have not covered? As I discuss in the following general statement, the missing ingredient in this (in addition to the issue of how the products are being used by the public) is the need for an analysis of the complete pharmacology of ephedrine/caffeine products. This must include consideration of the modern science of pharacogenomics and genetic polymorphisms of receptors for these products. This type of analysis adds relevance and credibility to adverse events that occur in low frequencies. It explains how some patients can have little or no change in blood pressure or heart rate and how some can be placed at risk of stroke, seizures or heart attacks.	A good suggestion, but one that we feel is probably some years off, as opposed to the three studies listed first. A genetic analysis could conceivably be added to a case control study and used as an effect modifier in the analysis.
Implications for future research were discussed. Physicians and most pharmacologists seem to want to lay the blame for problems associated with ephedra supplements at the feet of ephedrine/ caffeine. This narrow view excludes the pharmacological activity or potential interactions with other phytochemicals present in these products. In the opinion of this reviewer, the problem is more complex than simply ephedrine and caffeine.	This is a good point. An assessment of ephedra use may be able to take advantage of the heterogeneity in concoctions to perform subgroup analyses looking for ingredients other than ephedrine and caffeine.
Regarding future research aimed at stroke aspects, it would seem valuable to pursue case-control studies along the lines of that by Kernan and colleagues cited above but considering both hemorrhagic and idiopathic ischemic stroke in relatively young adults.	Agreed. The proposed case control study should assess all of the serious outcomes we assessed.
Page 5, paragraphs 3 and 4 would it be appropriate to mention ethical considerations for case-controlled studies in the summary?	We do not think so. The exposure has already occurred.
Page 5, paragraph 4: “Pre-clinical studies should also be considered to determine the use of ephedrine or ephedrine containing the alkaloids increases the risk of development of heat related conditions such as heat exhaustion, heat stroke, and rhabdmyolysis, if an appropriate animal model can be found.” What specifically would be learned from this? Could it be extrapolated to humans?	It might help establish a biologic rationale, but in this discussion we have deleted the “animal model” and “pre-clinical aspect” to this and suggest it be included in a study of adverse outcomes in humans.
Rewrite to redirect emphasis of sentence by placing, “If an appropriate animal model can be found, pre-clinical studies should be….and rhabdomyolysis.”	We actually eliminated the “preclinical studies” part of this and suggest a study assessing this as a potential adverse event.
I would suggest come additions to the section of future research including an interaction study to investigate the effects of ephedra with not just caffeine-containing herbs, but also combined with botanicals such as citrus aurantium, garcinia cambogia, and the herbal diuretics and cathartics	These suggestions have been added.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
as listed in Table 1. Also, I would specifically suggest that studies on athletic performance be conducted in women and adolescents, since these populations are known users of these products. Finally, I would recommend that the association between ephedra and seizures be formally explored.	
Future Research Section. You favor a case-control study. The case-control trial with phenylpropanolamine was sufficient to remove the drug from the market, but it was a pretty poor study. The controls and cases had very different lifestyle habits. There was a no-dose-response to PPA. The effect was only detected in women. Because of the large number of things used in the many products on the market, and the relatively high rate of deaths and disability from heart disease and stroke, it is not clear that a useful answer would emerge.	We disagree. We think such a study would tell us something useful about ephedra products. We do not think the heterogeneity in the components of the products will be any greater impediment to the analysis of safety than it was to efficacy. Our data are consistent with the hypothesis that the only active components with respect to efficacy and safety are ephedrine and caffeine.
Another point that might be useful to make is that from the available data it is not possible to determine which populations are at greatest risk for serious adverse events, and that this could only be determined by additional research.	This point has been made in our suggestion for a hypothesis testing study.

CAUSALITY COMMENTS

The remaining reviewer comments from the first review concern an attempt in our draft document to assess causality for some adverse events. We did so using our own modification of published methods. These comments varied widely, ranging from critiques of our method for being too conservative (meaning, in the opinion of some reviewers, we had excluded or assigned too low a level of causality to certain cases) to critiques for being too liberal (meaning, in the opinion of some reviewers, we assigned too high a level of causality to certain cases). Often, these conflicting comments concerned the same cases. We believe these peer review comments demonstrate that case report reviews involve considerably more subjective interpretation than do reviews of randomized trials. Because our goal in this evidence report is to report the evidence as objectively as possible, we ceased to assign assessments of causality to the case reports. Rather, we tried to identify those cases that would be classified medically as "idiopathic" in etiology, meaning the cause is not known. For such cases, given the known pharmacology of ephedrine, if use of ephedra or ephedrine was documented, a potential role for ephedra or ephedrine in causing the event must be considered. We classified such cases as "sentinel events." Other than correct typographical errors and respond to questions of fact, we do not provide a response to the numerous criticisms of the causality algorithm or suggestions to change our interpretation of these case reports based on the reviewer's opinion or "additional information" they posses that we did not have in the documents available to us to review.

Reviewer Comment	Rand Response
Although the present study was compared to that report by Haller and Benowitz (New England J Med), why was a comparison not made to the Samenuk et al. study (Mayo Clin. Proc.)? Were individual case numbers not available from Dr. Samenuk? On page 51, in Level 2 of the Causal Flow Model, what is meant by "in more than minimal dose"? What constitutes a minimal dose? Are you talking about ephedrine, ephedrine/caffeine, or ephedra supplements? It must be emphasized that ephedrine and caffeine in conjunction potentially hundreds of other pharmalogically active phytochemicals constitute an ephedra supplement? Accordingly, the pharmacodynamic effects for ephedra supplements are not directly comparable to synthetic ephedrine or ephedrine/caffeine combinations. Furthermore, given the heterogeneity of ephedra supplement formulations, the pharmacodynamic effects of individual ephedra supplements are expected to vary.	
It should also be acknowledged in the final report that there is very little consistency in the results of any expert attempts at assessing causality with this same set of AERs. The draft touched on this issue in discussing comparison with other reports and in presenting information in Table 22. The language of the draft is not, however, consistent with the data in the table. The draft states that the current judgments "are more conservative than those of Drs. Haller and Benowitz" but that there was agreement that some cases cannot otherwise be explained. It is difficult to understand how such statements, with their implication that any differences are either minimal it immaterial, can possibly be associated with the actual information in the table. The table identifies 24 cases, 20 of which were evaluated by both this	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
group and by Haller and Benowitz. In only 5 of these is there full agreement. Twelve of the 20 cases are cases where Haller and Benowitz report possible or probable causality while the current reviewers reported insufficient information! It is curious that the Table does not note that, of the 11 cases reported by the draft as probably and the 26 cases identified as possibly related to the use of ephedra, Haller and Benowitz only classified 2 of the first and 7 of the second as either probably or possibly related. Information is not provided to assist in understanding whether Haller and Benowitz had classed these as not associated with ephedra of whether they had not evaluated these cases. If the first case, it should be disclosed if the second, there should be some mention that the current causality assessment is preliminary and subject to review by other qualified experts.	
It is also curious why reviews with different conclusions by other parties were not acknowledged. For example, in at least one case where both this group and Haller and Benowitz agreed that there was a possible causal relation between use of ephedra and death, the local coroner ascribed the unfortunate incident to congenital problem. In addition testimony was given by Theodore Farber, Ph.D. on August 8, 2000 at the HHS Office of Women's Health to discuss the issue of inconsistency at length and in detail. As Dr, Farber noted " There was a sufficient lack of concordance between the FDA's causality analysis and the causality analysis performed by it outside experts." Other presenters at this meeting provided analyses of these AERs that found quite different conclusion that have been drawn in the Draft. It must be assumed that the record of this meeting and possibly more specific information, was accessible as the draft was being prepared but it does not appear that any attention was paid to any other commentary on causality reviews to date. This must be corrected.	
In discussing the case reports the Draft states that "events related to synthetic ephedrine" were removed. Notwithstanding out earlier attempt to clarify that synthetic ephedrine probably means ephedrine in isolation (or its salts, e.g. ephedrine hydrochloride), at least 8 of the cases reported on were associated with a product that was labeled to contain ephedrine hydrochloride (the E'OLA product) or was subsequently found, or at least has come out to be assumed to have been manufactured with undisclosed ephedrine salts (Formula One). At lease two cases do not identify the brand so it is not known how this determination was made from these cases.	
Almost every 'Probable" case and "Possible" case had either a preexisting condition that could have contributed to the adverse event or exhibited unhealthy behaviors	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
(excessive drinking, smoking, "intense effort to lose weight") which should be noted.	
On p 69 and Table 22 there is a comparison of the results of Rand evaluation of FDA AER with those of Dr. Benowitz – the specific criteria to meet definite, probably or possible causality are explicitly stated for the Benowitz evaluation [as was the case with other expert evaluations of these data in the FDA docket]. It would be useful to specifically list the criteria used for the Rand Evaluation for their classification. It is stated that the Rand evaluation is more conservative than the analysis by Benowitz, but what about comparisons with other expert reviews of these data [2 FDA reviews, Woosley, Benowitz, Ricaurte and Stoll]?	
Table 22 Summary of comparision with other reports of ephedra adverse events there is an error in the table for FDA case number 12720 and 12722. According to information elsewhere in the report, the following appear to be the correct data entries: Case# Adverse event Benowitz EPC 12720 Death Possible Insufficient information 12722 Death Possible Possible	
What was the classification of the CanTox study commissioned by CRN?	
Do you have information as to how soon an autopsy was performed? Could this have any impact on the toxicological screen results?	
I would assume that the prevalence of pre-existing coronary artery disease is very high. Therefore, when interpretations are made as to causality and risk for most Americans, it may be important to have some reference numbers as to how many Americans have pre-existing CAD.According to the Am Heart Association (using NHANES III data), 1 in 5 males and females has come form of cardiovascular disease, see this website for details about CAD.See American Heart Association. 2002 Heart and Stroke Statistical Update. Dallas, Texas: American Heart Association, 2001.http://www.americanheart.org/downloadable/heart/1014832809466101319099012 3HS_State_02.pdf	
My prior bias about ephedra and stroke was based on influential case-control study of the relationship between the use of phenylpropanolamine, a compound with related physiological effects, and hemorrhagic stroke (Kernan WN et al. NEJM 2000;343:1826-32). I believe that it is likely, re-enforced by the data in this draft report, that ephedra use occasionally leads to stroke. However, for the purpose of this review, I have elected to play the devils advocate in considering the specific question: "how strong is the existing	A neurologist was included in the review process in this revision. "Grand mal" seizure was the description of the event in the original source material.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
evidence that use of ephedra can cause a stroke?" Because determining the cause of strokes among young people is not that often straightforward, it would have been optimal to have the stroke cases reviewed and classified by a stroke expert with experience in evaluation of young stroke. The case reports (p.63) suggest a lack of neurological sophistication (i.e. grand mal seizure in case 11062 is not technically correct; generalized convulsive seizure is probably what was intended). OK,so this is an irrelevant elitist comment, but in the absence of hard evidence, credibility is a subjective issue. Case 10874 is categorized as "probably causal" : along time intravenous drug abuser with phenylpropanolamine on toxicology screen. Case 9335 is classified as "possibly causal": 56 year old woman with hypertension, tabacco use, elevated cholesterol and triglicerides, an MRI with microvascular changes and whose event was lacunar infarct. Case 12713 was "possibly causal": a 63 year-old woman with artificial fibrillation with acute loss of conscious and embolic stroke. It would be easy to take issue with classification of likelihood of causality in each case. In short, I agree with the appropriately cautious conclusion that "there is sufficient evidence to suggst a possible causal role of ephedra-containing dietary supplements in rare, but serious adverse events, particularily cerebral hemorrhage." (p.vi) Support for this statement would be better served by have a stroke expert review the case reports and perhaps tossing our the marginal cases (such as noted above). Further, since this authoratative report may eventually be used for medical-legal purposes, it would seem responsible to include a caveat that it is not sensible to consider all strokes of idiopathic cause in people taking ephedra as caused by the agent. These comments are not meant to disparage the overall quality of this impressive report. As noted at the outset, I have elected to play the devil's advocate concerning this specific aspect.	
The criteria for determining causality were arbitrary and did not address the true causality. In fact, the term "causality" is misleading in this connection. Rather, the term "association" should be used as in "guilt by association". Lay people will read this report and "probable causality" will be interpreted as a cause-effect relationship which is not warranted by the data available from medical record reviews.	
The limitations of the data collection are not emphasized enough in this report. Adverse reactions reports by definition have no denominator and are subject to reporting bias. In the famous Phen-Fen debate, initial reports on which the FDA took action suggested at 35% incidence of valvular abnormalities. Subsequently, this was found to be less than 8% and reversible following discontinuation of the	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
medications.	
It is very important to the integrity of this report that the basic questions asked in the contract are answered, that the report is well ordered, and that only scientifically valid information is included. If the AER information, which is not scientifically valid is included, it should be included as an appendix, not in the body of the report, as this takes away from the science.	
Use of the terms probably and possibly causally related may make the causality assessment sound more objective than it is. Would be the subjective natures of the assessment be more effectively conveyed by changing those AERs currently designated as probably causal to possibly causal, such as events if uncertain relationship? Instead of specific designations it might be adequate to describe the results in narrative form. The narrative could explain that although in some cases cofounders make it difficult to attribute causality, there is a subset of cases in which cofounders make it difficult to attribute causality, there is a subset of cases on which cofounding factors are minimal or absent as far as can be determined, and it is these cases that raise concern over safety. Whatever terms or phrases are used, defining them early in the document will help even those unfamiliar with adverse event causality analyses understand their meaning.	
Should make it clear that it is not possible to determine the actual level of risk for people taking ephedra or ephedrine because the number of people who actually take it is not known.	
It would be helpful to provide possible reasons for the differences between the RAND causality assessment and the one done by Haller and Benowitz, this could be done by adding text to point out that: I. Each group used different criteria. II. The same group of experts would come to different conclusions of they were using different sets of criteria for evaluating the same set of AERs, and III. The RAND report use more stringent/restrictive criteria for assigning causality than were used in the Haller and Benowitz review, resulting in more conservative assessment.	This table has been dropped from this revision since causality is no longer assessed.
Requirement of angiography for assigning causality for M.I.s to Ephedra (similar comments from two reviewers): Page V Paragraph 5:" for cases of myocardial infarction, we required coronary angiography to have been performed and the results available." This seems like a very restrictive set of myocardial infarction cases. What would be the effect on the results if angiography had not yet been done? Why was this restriction used? Results section explains this better, Assume all such cases would have been classified as	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
possibly causal , so the data is still listed. Pages 3 and 32: While we understand the importance if documenting the occurrence of myocardial infarction by restricting the documentation of the event so that the cardiac characterization is required to assign causality to ephedra, are the number of MI events being underestimated? Why not also use enzyme changes in laboratory specimens and Q-wave changes on the EKG to assign causality?	
"'Probably not causal' was used for events that had clear other causes discovered on detailed investigation." This assumes all events had a single cause. But can't someone with known atherosclerosis die suddenly because of superimposed effects of a substance.	
This sentence is confusing " In the 935 reports, there were data in 968 subjects of which 925 reported taking ephedra." Not clear how there can be more subjects than reports.	A single FDA MedWatch report can contain information on more than one person.
A case presented (#12843, 15-year-old female) without any reference to ephedra exposure. Absent that information, it would be hard to make this even a possibly causal classification.	
A couple of reviewers were confused by the mention of AERs that took place after September 30, 2001. Perhaps a footnote on the table would be informative to remind readers if the timeframe for the AERs analyzed.	We added text to try and help explain this.
Table 20 provides a lot of useful information, but it might be easier for readers to interpret the data if another table were added. This table would have 5 columns across the top, labeled Product, #Probably Causal, #Possible Causal, # Insufficient Information, and Total (terminology may change based on other comments).	We considered but did not make any changes to this table. A different kind of Summary Table is included with this revision.
Data from the case report of the death of a 28-year-old female indicates that the MiniThin was one of the products that she was taking. MiniThins were shown to contain synthetic ephedrine, not ephedra, and the FDA required the company stop marketing it for weight loss and change the name and marketing focus (product name was subsequently changed to MiniTwoWays and was marketed for use on people with bronchial asthma). Should this AER be included there?	This case was included in the ephedra FDA MedWatch file. The patient was also taking Yellow Jacket.
Specific cases. Page 60, Deaths, Probably Causal: A 21-year-old male collapsed…": This patient has been taking hydroxycut, which I assume is hydroxy citrate. Hydroxy citrate is probably quite toxic, though it has not been systematically assessed in clinical trials. Biochemically it may be assumed to have a substantially liver toxic effect. I think it is therefore very difficult to attribute the case to ephedrine. I think that there are too many examples of patients with many other risk factors, such as those included	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
under the "probably causal" myocardial infarctions, e.g. a 54-year-old, who has smoked for 30 years and been an alcoholic.	
Specific cases. Page 62, Deaths, Probably Causal (cont'd): Page 62: Another example is the "Stroke, Probably Causal": "She was a long-time intravenous drug abuser and alcohol abuser. She also smoked cigarettes for 10 years." She tested positive for benzodiazapines and phenylpropanolamine, whereas there was no positive test for ephedrine. I strongly disagree with the conclusion that this case can be classified as probably causal with respect to ephedra use. It is more likely, with the given history and the positive test of the patient, that the stroke was caused by other vaso-active drugs taken by the patient. These weaknesses apply to several of the other stroke cases, and I think this is particularly interesting in light of the meta-analysis of adverse events reported from control trials (Table 17, page 80) where it is found that there is no statistically significant increased risk of hypertension. This also quite clear from the control study by Ingerslev et al. on hypertensive patients treated with ephedrine/caffeine. One should therefore be cautious about drawing conclusions on the causality with respect to stroke.	
HHS and GAO Statements Regarding the Same AERs that RAND Reviewed. The draft report attempts to ascribe degrees of causality to the ephedra AERs, thereby ignoring recent statements by the Department of Health and Human Services and the General Accounting Office ("GAO"). HHS and FDA recently reviewed the same AERs that RAND reviewed, and issued a response to Public Citizen on June 14 that provides the following: The primary purpose of a voluntary adverse event reporting system is to generate 'signals' of potentially related events, rather than assessing product safety. While a 'signal' has been generated by these reports, FDA has determined that questions remain on the likelihood and strength of association between ephedrine alkaloids and the adverse events reported to the FDA... There are situations when background rates of the observed event are so rare or unusual that, in combination with physiologic responses and biologic plausibility, a significant relationship between the events is self-evident from the reports in a voluntary reporting system. However, the FDA has advised me that the types of observed outcomes reported in relationship to the ingestion of ephedrine alkaloids are not uncommon in the general population and therefore the reports alone do not provide a scientific basis for assessing the safety of ephedrine alkaloids or establish a link between the reported adverse events and the ingestion of ephedrine alkaloids. (emphasis added). The draft report should mention and cite the above-	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
mentioned letter and FDA's and HAS' position with regard to the AERs and attempts to ascribe causality. The draft report also does not acknowledge the GAO report on ephedra AERs, which reviewed the same AERs and determined that they were "poorly documented," further weakening RAND's reliance upon reports for causation analysis. GAO also noted that the AERs have "inherent weaknesses" and lacked or had inconsistent information...such as the amount of the product used, how often it was used, or for how long it was used. These limitations were not prominently identified in the RAND report - nor were the general limitations associated with attempts to ascribe causality based upon review of information obtained from a passive surveillance system. GAO also noted that, based upon its review of specific AERs, it was not possible to draw conclusions regarding the "causal relationship between ingestion of the implicated product and the adverse event observed." Potential Product Variation. As noted above, the draft report does not even acknowledge the possibility that certain ephedra supplements may not be standardized and/or manufactured according to GMPs. Accordingly, even assuming causation (which is a major assumption), it is conceivable that certain adverse events may have been caused by problems with a specific product (such as having more ephedrine alkaloids or caffeine than stated on the label). Although it is my understanding that most manufacturers of ephedra employ stringent quality controls, this is still nevertheless a significant possibility that should be reflected in the report. Attachment B contains an article prepared by Dr. Gurley entitled "Content versus label claims in ephedra-containing dietary supplements." Although this article reviewed only a small subset of ephedra products, and only reviewed a small sample-size of bottles, it nevertheless supports the conclusion that it may be inappropriate to assume that all ephedra products are identical with regard to product quality. Accordingly, even assuming causation, it should be noted that there is no assurance that any potential adverse health events were not caused as a result of consumers ingesting non-standardized products that contain too much ephedrine or caffeine. It would be inappropriate for RAND to assume that consumption of ephedrine and caffeine within labeled amounts is a potential problem if the possibility exists that certain incidents may have been caused by consuming non-standardized products. Accordingly, my strong opinion is that as part of its review, RAND should call for FDA to finalize dietary supplement GMPs and impose stringent quality control requirements on ephedra manufacturers to ensure that such products contain what they are claimed to contain. Haller/Benowitz Review of AERs. The draft report places	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
great importance of the review of ephedra adverse events reports conducted by Dr. Haller and Benowitz. In fact, the draft report compares RAND's assessment of AERs with the Haller/Benowitz assessment. The draft report, however, fails to mention that Drs. Haller and Benowitz subsequently wrote a letter to the editor of the New England Journal of Medicine (Attachment C) indicating that their review did NOT prove causation or provide quantitative information with regard to risk. Specifically, their letter provides the following: Finally, our report describes a series of cases in which the use of ephedrine-containing dietary supplements was associated with a diverse cardiovascular events. Our report does not prove causation, nor does it provide quantitative information with regard to risk. A large-scale case-control study similar to the Hemorrhagic Stroke Project for phenylpropanolamine is needed to determine the risks associated with these dietary supplements. Based upon this letter and clarification, it is unclear why the Haller/Benowitz review of the AERs is used as a baseline for purposes of comparison. Moreover, it is unclear why their statement regarding causation and the recommendation of a case-control study is not highlighted - as this would appear to be information that RAND should consider as a recommendation for further research. Importance of Background Risk - Kimmel Study. This draft report fails to cite favorable analyses of FDA AERs - including a detailed study conducted by Dr. Steven Kimmel that was presented before the Office of Women's Health (Attachment D). Dr. Kimmel reviewed the AERs and determined that the number of events was consistent with background rates in the general population. His report highlights the importance of background risk - an issue that should be highlighted in the RAND report. He concludes that the AERs - even assuming significant under-reporting - are not suggestive of causation. In this regard, he quoted FDA" "it is possible that the reported serious adverse events are reflective of coincidental background spontaneous occurrences in the population and are not necessarily causally related..." OTC Drugs Containing Ephedrine and Caffeine. The draft report does not prominently refer to the wide usage in the United States of ephedrine in OTC drug products. The report should include use-data for OTC drugs, and should explain that FDA has already determined - under the OTC Drug Review - that ephedrine is safe and effective (as a bronchodilator) in does well over 100 mg per day (the maximum dose level for the vast majority of ephedra supplements on the market). In addition, the FDA does not require such products to contraindicate caffeine ingestion (i.e. consumers routinely ingest such products along with coffee, tea, and other beverages that contain caffeine). This	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
information must be factored into the final report, as they have a direct bearing on any safety assessment of ephedrine and caffeine. OTC drug use-data indicates that the combination of ephedrine and caffeine is safe. Scientific Data and the Landmark Six-Month Harvard-Columbia Trial. The draft report acknowledges that "there were no serious adverse events reported in these clinical trials." Despite this, the report barely addresses this issue. Rather, the vast majority of the report reviews and subjectively interprets AERs that GAO, HHS, and FDA have already reviewed. Even though the number of subjects in the clinical studies is limited, and therefore it is conceivable that small subsets of the population may have some susceptibility, the clinical data is far more reliable than the anecdotal adverse event reports and should receive greater prominence than the AERs. In addition, the report does not place enough significance upon scientific data such as the Cantox Report and the landmark six-month Harvard-Columbia trial (published in the International Journal of Obesity). The Harvard-Columbia trial addresses the review of adverse event reports, and makes suggestions regarding future research. The RAND report should contain a more detailed discussion of this landmark trial - including the researchers assessments regarding product safety and efficacy. For ease of review, sections of the report addressing adverse event reports, product safety and efficacy, and future research are included below: In a FDA-sponsored analysis, Haller and Benowitz categorized 140 adverse-event reports based on how likely they believed the reported events to have resulted from the use of ephedra supplements. The difficulty in making such judgments is illustrated by the controversy regarding their conclusions. With millions of American consuming ephedra-containing products it is obvious that some number of adverse events is expected each year regardless of consumption of these products. The real question is not whether adverse events occur in a population undergoing treatment, but whether these occur at a rate that is higher than that of a matched, untreated group. This is impossible to determine from adverse event reports alone. The randomized, placebo-controlled trial allows evaluation of cause and effect relationships vs. coincidental events. Most clinical trials purposely exclude individuals with pre-existing medical conditions to avoid confounding of results. It is therefore not justified to extrapolate results from such trials to individuals with such exclusionary medical conditions or to extrapolate results beyond amounts or time periods that have been studied. The possibility of unfavorable interactions between herbal combinations and other medications, either prescription or	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
illicit, should be recognized and warning labels present on herbal products should be adhered to. Some have expressed the theory that adverse event reports may reflect an unusually high degree of sensitivity in a small fraction of individuals. Because of the low suspected incidence, this type of sensitivity might not be revealed in a clinical trial, but requires a case-control study of a very large number of individuals. Such a study would be difficult to conduct, but may be the only way to address the question of rare hypersensitivity. In total, these [ephedra studies] suggest that herbal ephedra/caffeine herbal supplements, when used as directed by healthy overweight men and women in combination with healthy diet and exercise habits, may be beneficial for weight reduction without significantly increased risk of adverse events. In total, these [ephedra studies] suggest that herbal ephedra/caffeine herbal supplements, when used as directed by healthy overweight men and women in combination with healthy diet and exercise habits, may be beneficial for weight reduction without significantly increased risk of adverse events. The current widespread usage of herbal products and the increasing incidence of obesity warrant additional clinical trials to confirm and extend these results. (emphasis added). Finally, it should be noted that the Harvard-Columbia Trial researchers, and Drs. Haller and Benowitz appear to agree that in order to evaluate the safety of ephedra, a long-term control study would be beneficial.	
The draft report in the end recognizes the futility of trying to reach scientific conclusions from the AER's, recommending that a case control study be done to assess risk and recommending against further AER analysis. Nevertheless, a detailed causality assessment was performed and included in the draft report, and conclusions of this assessment are presented without context. Further, the draft report describes the involvement of the TEP in a way makes it appear this assessment was done on the recommendation of the TEP., when I and others thought that there was a general agreement within the TEP and RAND further assessing causality based on the AER's was not recommended and would not be part of the ephedra review. The draft report on page 21 states that the "Highest level if causality that could be ascribed. Was "probably" causal". This is not a position taken by the TEP at the November meeting; in fact it is contradictory to the TEP's position as quoted above. There was a discussion of the characteristics of the case reports, but not in the context as stated on Page 21..."that would be necessary in order to assign a classification of "probably causal." The criteria quoted were of causality. The report not only implies (also	The causality analysis was dropped from this revision. Regarding the involvement of the TEP, there was a lengthy discussion of the criteria by which we would assess case reports, so we are surprised this TEP member concluded "the TEP" agreed that such a review was not warranted. We did not receive any such comment from any other TEP member, all of whom reviewed this report.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
on page 31) but also states that the TEP agreed with this causality classification categorization, which was not the case. The classification scheme was developed after the November meeting, and any category such as "possible causal" or "probably causal" that suggests a causal relationship is in contrast to the position of the TEP at the November meeting. Terminology of "possibly" or "probably" causal is too strong and more than suggestive of causality, and to suggest that these terms are less than "definitely" causal is too fine a point for readers of this report and also an incorrect representation of the TEP's position. If the causality assessment remains in the final report, I suggest that all reference to this classification scheme be changed accordingly, to omit the word "causal". The conclusions could be characterized as weak evidence or possibly suggestive evidence, but the words causality and causal are too strong. The draft report notes correctly about the AERs that "The most important limitation is that the study design, that is an assessment of case reports, is insufficient to warrant definite conclusions regarding causality." Yet when it came to assessment of individual case reports, there were definite conclusions that the AERs prove ephedra to be unsafe. The result is a misleading presentation of the available information. For the reasons explained above, I feel the report requires major revision and subsequent further review by the TEP. Because I have been focused on the fundamentals of the reports as described above, I haven't even considered the comments on details that are included in the draft report. The weakness of the FDA AER database must be better addressed, and the causality analysis should either be removed from the report or substantial revised to, among other things, provide the necessary context and to change the classification of the AERs to avoid using the terms "causal" and "causality". The fact that the safety section is dominated buy the AER analysis reduces the credibility of the section and indeed the whole part of the report.	
However, in the opinion of the reviewer, those conclusions regarding the case reports are limited by a combination of the conservative causality assessment criteria and the limited medical records and toxicology data available for most case reports. For example, hypertension was defined as a systolic pressure in excess of 180 and a diastolic pressure in excess of 105. Also, no consideration appeared to have been given to the contribution multi-component ephedra-containing dietary supplements might have had in those individuals with underlying cardiovascular or cerebrovascular disease. I think it is generally accepted among the medical and scientific community the presence	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
of sympathomimetic agents could potentially exacerbate the likelihood of adverse events in such populations? I would think most clinicians would factor such information into their differential diagnosis rather than dismiss them altogether.	
Adverse Events Reports: One limitation with your approach taken in evaluation of the adverse events reports made to the FDA is that your causality algorithm does not include an assessment of whether ephedra played a contributory role on the adverse event. Because ephedra is available as a dietary supplement, it is likely that many persons taking these products are not using them under doctor's supervision, and may have medical contraindications to their use. Therefore, the role of underlying disease becomes a crucial factor in causation assessment, particularly when a potential risk factor often goes undetected (i.e. essential hypertension, structural heart defect), or when a condition is omitted from the ephedra product label warning (i.e. family history of premature CAD, sickle cell trait).	
Two AERs that you assess as no higher than possibly causal illustrate this point. AER 12485 did indeed have a moderate degree of coronary artery disease detected at autopsy. However, he was reportedly in good health without history of angina, and had been jogging regularly without adverse effects. Because he collapsed suddenly after returning from jogging, we felt this was a primary arrhythmic event due to ephedra. Similarly AER 12843 was a healthy, adolescent who had participated in competitive sports for many years. She had appeared to have been well-compensated for a serious underlying coronary artery abnormality that was clinically undetected since birth. Only with use of Ripped Fuel, did she suffer a catastrophic cardiac event resulting in death. We felt that the cardiac stimulant effects of ephedra resulted in myocardial ischemia in this case.	
It would be helpful to specify what degree of pre-existing coronary artery disease would constitute a significant risk factor to result in myocardial infarction or sudden death in the absence of stimulant use, thereby ruling out ephedra in the causation assessment. (page 32 of chapter 2 methodology). In the case of AER 14530 (page 63), I would disagree that 20-30% stenosis would be significant enough to result in acute M.I. in a 43-year-old female smoker without a significant contributory effect from the ephedra alkaloids in Metabolife.	This case was reviewed by a cardiologist who made this judgment.
On page 59 I suggest the authors be slightly cautious with their use of language such as "probably causal" and "no other possible explanation." The latter phrase is particularly troubling. What they mean is no other explanation that they could identify. Similarly, on page 69 they state that there are a certain number of cases of serious adverse events that	We have revised this language.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
"cannot be explained by causes other than ephedra use." While I do not deny that it is extremely likely that many cases of adverse events happen due to ephedra use, simply because we do not have in our hand an explanation of why an event occurred other than a particular explanation under consideration, does not mean that the particular explanation under consideration is the correct one.	
Because of the paucity of large randomized trials, evidence concerning stroke and ephedra by necessity consists of analysis of case reports. "…An assessment of case reports is insufficient to warrant definite conclusions regarding causality." (p.110) Nevertheless, arbitrary criteria are used to define "probably cause": documentation that a stroke occurred, that ephedra was used, and that there was exclusion of other potential causes. The definition may be too liberal. Of ischemic strokes in relatively young adults (i.e. those <50 years old), perhaps 20-35% are "idiopathic" despite thorough evaluation. The definition implies that all idiopathic strokes would be classified as "probably causal" if ephedra was used in any dose in proximity to the event. Given the frequency of idiopathic stroke, many (perhaps most) neurologists would consider "possibly causal" to be a better designation in this situation. Are there specific clinical circumstances in which the relationship of ephedra use and idiopathic stroke could be certain? Perhaps if acute, striking elevation of blood pressure were known to precede the stroke onset of of angiographic features characteristic of vasospasm were present in the abscence of migrane? Arbitrary to be sure, and not very helpful.	
The evidence for effectiveness supports the conclusion. Except for AERs, however, little evidence of toxicity is actually provided, and evidence of safety has been largely ignored. No evidence is provided to even suggest "a possible causal role of ephedrine-containing dietary supplement in rare, but serious events," let alone extremely common events such as heart attack and stroke. Even critics of ephedra have concluded that the clinical effects of pharmaceutical ephedrine, and the ephedrine contained in herbal preparations, are indistinguishable. Gurley states that the increased incidence of ma huang toxicity does not stem from differences in the absorption of botanical ephedrine compared with synthetic ephedrine." Haller and Benowitz, in their most recent publication, conclude, "Botanical stimulants have disposition characteristics similar to their pharmaceutical counterparts..." The Cantox Report, and the Report of the Expert Panel of the EEC reached similar conclusions. Since there are no real differences, studies demonstrating the safety of pharmaceutical ephedrine and pharmaceutical ephedrine in combination with pharmaceutical caffeine, should not be	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
excluded when considering the safety of herbal equivalents. The explanation most frequently offered for alleged cases of ephedrine-related stroke is drug-induced blood pressure elevation, this in spite of the fact that no clinical trial, of any duration, has ever demonstrated that a clinically significant effect on blood pressure exists. Indeed, the studies that have addressed this question. including the most recent paper by Drs Haller and Benowitz, have shown diminishing cardiovascular effects over time. In other words, if dangerous blood pressure elevations do not occur with the first dose ephedrine, they are even less likely to occur with prolonged dosing. These studies should be included in the RAND review of ephedra and should be used to address the question of potential increases in blood pressure and other safety issues. Ephedrine has been studied in more than 50 double blind, placebo-controlled clinical trials, some of long duration. A far from exhaustive literature search produced the attached list of peer-reviewed, published, clinical trials. Most have compared ephedrine to placebo, and to other sympathomimetic drugs used to treat asthma. However, others have involved smoking cessation, sexual function and athletic performance. Nearly a dozen of these trials involved caffeine/ephedrine combinations using does exceeding those found in herbal supplements. In total, more than 2000 individuals have been enrolled in these trials. In several studies there was even continuous cardiac monitoring in middle-aged patients with known heart disease; no effect was observed. No clinically significant episodes of toxicity were reported. Including these and other studies on ephedrine that have been excluded from the RAND review will increase the power of the safety calculations that can be derived from clinical data. One of the major limitations of the report was the composition of the TEP and the reviewers who made subjective assessments of the AERs. Given the importance placed on assessment of AERs, it is unfortunate that no pathologist was included in the view or on the panel. The lack of expertise is obvious from the comments made about the individual AERs. The failure to provide information about any potential conflicts of reviewers also detracts from the study. Why were the findings of Expert Panel of the Ephedra Education Counsel not considered? The analysis of this panel was in some ways unique, as it is the only consensus opinion on ephedra safety. In addition, this panel conducted the most comprehensive review of the ephedra AERs to date, and yet the causality assessment, which conflicted the findings of the draft report, are not even mentioned. If RAND believes that the EEC review and analysis was, in some way scientifically flawed, then the reasons for that belief should be stated.	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
The danger of drawing conclusions from AERs without a control group can be illustrated by examining data from randomized trials in which participants are blinded to whether they are receiving the study treatment or inactive placebo. In the placebo group of a recent randomized trial,2 there was an increase in ventricular couplets (extra heart beats) at the 4th week of the study 9from 3% at baseline to around 14% at week 4). This is, of course, not due to the placebo, which is inactive, but rather just spontaneous ventricular couplets that occurred by chance. However, if these participants had been given ephedrine alkaloids in an uncontrolled study (without placebo), this change could have been attributed, incorrectly, to the ephedra. That is, these could be AERs that were attributed to ephedra. In fact, in the controlled trial, a similar increase in ventricular couplets was not seen in the ephedra/caffeine arm. Another example is the 15-year-old female (case 12843) with Bland-White-Garland syndrome who died while playing soccer. This disorder has been associated with sudden death after physical exertion. In the absence of a unique pathologic process, it is almost never possible to establish a causal association on the basis of adverse event reports. There is nothing pathologically or diagnostically unique about the adverse events noted in the ephedra database (e.g., myocardial infarction, stroke) that allow one to distinguish a spontaneous event from one caused by use of Ephedra products. In fact, a review of all autopsy data from ephedra AERs by Dr. Grover Hutchins, a Professor of Pathology at the John Hopkins University School of Medicine and member of the Expert Panel of the EEC, concluded that "The pathology data available do not show any pattern consistent with ephedrine alkaloid-containing dietary supplements as a cause of death."5 Similarly, 10 participants in the ephedra/caffeine group (12% of these participants) withdrew because of cardiovascular symptoms (palpitations, elevated blood pressure, arrhythmias). If there were no control group, these also might have been attributed to the ephedra/caffeine combination. However, the same proportion of participants in the placebo group (13%) withdrew for the same reasons. The withdrawal rate in the ephedra/caffeine group thus was consistent with the background rate of these events in a placebo group unrelated to ephedra use. A second limitation of adverse event reports is that other potential causes of the event are often present, making it extremely difficult to determine if an event truly is related to the exposure. As an example, it is well established that physical exertion can trigger myocardial infarctions and cardiac arrest (up to a 74-fold increase in the risk of sudden death, according to a recent report in the New England Journal of Medicine6)	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
Therefore, a case such as the 38-year-old man with three vessel coronary disease and a dilated heart who died after jogging could very well have been related only to the physical exertion and not the ephedra. According to his wife, his heart problems had been known for at least five years. In addition, this man had been taking an ephedra product for one year without apparent ill effect. The third limitation of AERs is that their interpretation remains subjective. Even experts in the field will disagree about the possibility that an event may or may not be due to an exposure. The RAND report (Table 22) displays a comparison of the causality assessment between their review panel and a published report by reviewers for the FDA.3 In only two out of 20 cases that both groups reviewed was their agreement about the highest level of possible causality. For example, with respect to the 38-year-old discussed above, the Haller and Benowitz review recorded this event as "Definitely or probably related" to ephedra while the RAND report classified it as only "Possibly causal." Another review of AERs performed by Dr. Theodore Farber and Dr. Norbert Page, members of the Expert Panel of the EEC, reveals similar disagreements. The two "probable" cases reviewed by EPC were rated as "Low Possible" (case 12980) and "Improbable" (case 13418) by Drs. Farber and Page. The fact that the etiology of events can be debated simply illustrates the substantial limitation of case reports that lack a comparison control group. This echoes reviews done by FDA and its consultants in which agreement about causality was poor. For example, two consultants from FDA, Drs. Ricaurte and Stoll, reviewed 28 AERs related to neurological events. Dr. Ricaurte classified eleven cases as "attributable" while Dr. Stoll classified only five as "attributable."5 Only two of the consultant's findings overlapped - that is, there were only two cases that both Dr. Ricaurte and Dr. Stoll agreed should be categorized as "attributable."5 This disagreement is not a flaw of the reviewers, but rather a flaw of AERs. A fourth limitation of AERs is that ingestion of the substance in question cannot always be substantiated. For example, case 10276 is a 32-year-old with an enlarged heart who was found dead in his truck. Although a product that contains herbal ephedra was found in his truck, so were several bottles of cold medications, including Nyquil. Toxicology revealed no ephedrine, but did identify pseudoephedrine and doxylamine, both components on Nyquil. Thus, there is no evidence that this person even ingested the herbal ephedra. A similar case (13096) revealed no ephedrine in a toxicology screen again suggesting that the man had not ingested ephedrine around the time of death. Equally importantly, this man died of a disease that appeared to run	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
in his family (aortic dissection), including an 18-year-old niece. In summary, AERs cannot be used to assess causality. As stated by several authors with experience in interpreting adverse event reports for the FDA: "It is probably impossible to do comparative analyses employing ADE [adverse drug event] reports for drugs that have received extensive publicity in the mass media for an adverse event..."7 In fact, the FDA's Center for Drug Evaluation and Research pointed out, in a February 10, 2000 memorandum concerning ephedra products, that "it is possible that the reported serious adverse events are reflective of coincidental background spontaneous occurrences in the population and are not necessarily causally related to [the use of dietary supplements containing ephedrine-type alkaloids]." The RAND review notes this as well, stating that "The most important limitation [of their assessment of adverse events] is that the study design, that is an assessment of case reports, is insufficient to warrant definite conclusions regarding causality." They list this limitation as one of the "most important...gaps" in the current knowledge-base. The RAND review also states that "Disentangling the relative importance of the pre-existing condition and the ephedra use is not possible." They state further that "Continued analysis of case reports cannot substitute for a properly designed study to assess causality. A case control study would probably be the study design of choice." Their Technical Expert Panel also "judged that case reports alone would be insufficient to establish definite causality between ephedra use and serious adverse events." Because of these limitations, terms such as "probably causal" and "possibly causal" in AER reviews are potentially misleading (see, in particular, the "Conclusions" section of the "Structured Abstract," page vi, and the "Conclusions" section, page 112). They represent only reviewers' assessment of causality based on uncontrolled data and subjective assessments. Although these terms are often used in scientific publications, their use in the RAND report may suggest a level of evidence that does not exist from the current data. These statements, therefore, should not be taken out of context. RAND's stated limitation that "Definite causality cannot be determined from case reports" must be kept in mind when interpreting this report. It is also critically important to remember that case reports can produce false signals of cause and effect.8,9 Most importantly, because of the limitation of AERs, it is unclear why the RAND report states, in the Structured Abstract, that "These [sic] is sufficient evidence to suggest a possible causal role of ephedra-containing dietary supplements in rare but serious adverse	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
events, particularly cerebral hemorrhage." With respect to a possible causal role of ephedra in adverse events, RAND acknowledges that AERs are not sufficient to draw conclusions about causality, consistent with the known limitations of AERs discussed above. In addition, this comment is puzzling given that their "Conclusions" section (ages 112-113) does not state this at all. With respect to cerebral hemorrhage, the only comment in the body of the report on cerebral hemorrhage refers to a case-control study of phenylpropanolamine (PPA) and cerebral hemorrhage.10 However, the RAND report refers to this study only as an example of case-control methodology that could be applied in the future to ephedra. The report does not discuss the PPA study further. In fact, this study has been heavily criticized. Despite this, there was not a formal review of this study by RAND (and there were no members of the technical expert panel listed with expertise in epidemiology to perform such a review). In addition, PPA and ephedra have different chemical structures and different pharmacological activities. Finally, the RAND report does not mention that the PPA report also presented data on non-PPA, ephedrine-alkaloid containing products. These agents included medications that contained pseudoephedrine hydrochloride, phenylephrine, ephedrine, and epinephrine. In the report, there was similar prevalence of use of these products among those with and without hemorrhagic strokes. Although this is not a definitive analysis, it suggests that there was no association between these ephedra-containing products and hemorrhagic stroke. Therefore, the statement in the Conclusions section of the Structured Abstract of the RAND report is inconsistent with currently available scientific data. In summary, the RAND review supports the use of herbal ephedra and caffeine for weight loss, an effect that may have beneficial health consequences. The report also suggest, from controlled studies, that adverse events following ephedra use are, at most, rare. (The should not imply that the events are even causally related to ephedra us.) Most importantly, because of the reliance on AERs, the report cannot establish a causal effect of ephedra on serious adverse events.	
Review of Anecdotal Adverse Event Reports - Limitation of Review. On pages 109-110, the report identifies potential limitations associated with the review of anecdotal adverse event reports. These limitations are buried at the end of the report, rather than being incorporated into the appropriate sections of the report (as per prior AHRQ reports, such as the Garlic Report - see Section IX, herein). Moreover, a number of limitations are not prominently identified, including but not limited to: a) The poor quality of the data and information contained in the anecdotal adverse event	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
reports; b) Inherent problems associated with voluntary reporting systems; c) The number of individuals in the general population who experience the adverse events identified (i.e. background risk); d) The possibility that certain products may not have been standardized and/or manufactured according to Good Manufacturing Practices ("GMPs") - resulting in potential adverse events that have nothing to do with ephedra or caffeine when consumed in recommended amounts. Specifically, in the absence of standardization and quality control, it is conceivable that certain products may contain far more ephedra or caffeine (or other constituents) than indicated on the product label. This possibility must be considered when evaluating anecdotal adverse event reports. In the absence of identical product identity, any general conclusions regarding ephedra and caffeine are inappropriate and highly suspect based upon adverse event reports (see the AHRQ Garlic Report for an appropriate way to address this issue). In my opinion, RAND should strongly support immediate issuance by the FDA of dietary supplement GMPs and should endorse stringent quality control measures to ensure that all ephedra supplements contain what they are claimed to contain; e) The possibility that the consumer abused or misused a product by ingesting more than the recommended amount - or that the consumer ignored detailed product warnings and contraindications. RAND should emphasize the detailed warning label contained on the vast majority of ephedra supplements - and should acknowledge that there is little way to know from anecdotal data whether a consumer abused a product (either intentionally or more likely inadvertently); f) The possibility that the anecdotal adverse event reflects chance, coincidence , or confounding factors - including but not limited to the possibility that ingestion of a different product or substance led to the stated event. I I. Review of Anecdotal Adverse Event Reports - B. Ascribing Causality to Adverse Event Reports - 1. Reliance on Unpublished Article. In order to establish a framework for analyzing the adverse event reports, the draft report relies upon an unpublished article written by Cynthia Mulrow, M.D. Reliance upon an unpublished, non peer-reviewed article to establish the framework for a critical portion of the report is entirely unacceptable. AHRQ studies have not, to our knowledge, ever relied upon an unpulished article to establish the framework for this type of analysis. In addition, reviewers such as myself have no way of analyzing the article - thereby defeating one of the primary reasons for review of the draft report. II. Review of Anecdotal Adverse Event Reports - B. Ascribing Causality to Adverse Event Reports - 2. Failure to Review Factors Critical to the Interpretation of Anecdotal Adverse Event Reports. Table 4 of the draft report	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
summarizes the various methods researchers use to assess causality from adverse event reports. The following nine factors are identified: a. Temporal relationship. b. Dechallenge response. c. Rechallenge response. d. Could there be an alternative explanation? For example, dehydration or consumption of other toxic substances. e) Prior reaction to same substance. f) Dose response. g) Objective evidence of adverse event. h) Previous conclusive reports. i) Definition of substance. Despite the report's reference to these nine factors, it is my understanding that the report concludes that events are "probably causal" based upon a review of only two factors - a and g. The draft report does not explain why RAND believes only two factors out of nine can be used to ascribe degrees of causation to anecdotal adverse event reports. II. Review of Anecdotal Adverse Event Reports - B. Ascribing Causality to Adverse Event Reports - 2. Failure to Review Factors Critical to the Interpretation of Anecdotal Adverse Event Reports (cont'd). The draft report also indicates that three factors are used to determine if an event is "probably causal": a) Reasonable certainty that the adverse event occurred. b) Reasonable certainty that the patient took ephedra in a dose and timing compatible with the known pharmacology of ephedrine. c) An adequate evaluation must have been done to rule out other potential causes for the adverse event. The third factor (factor c, above) is exceptionally problematic from a scientific perspective. The report acknowledges that the third factor is subjective. Specifically, in an effort to rule out other potential causes, the report indicates that such a determination was made by determining if the subject had a pre-existing condition that was identified in the adverse event report.	
Adverse Event Reports. While CRN acknowledges that the judgments made about the AERs were, overall, much more conservative than those made by other reviewers, there is concern that, in some cases, a much more likely explanation was evident, but still possible causality was assigned. Some examples follow, although they are not all-inclusive: 1. Case 10276. A deceased truck driver was found with cold tablets, Nyquil and Vick's Formula 44, in addition to ephedra-containing supplements. The toxicology screen was negative for ephedrine, but positive for pseudoephedrine. The much more obvious and likely culprit here would be one of the pseudoephedrine-containing cold formulas, since ephedrine is the dominant alkaloid by far in ephedra products, and the clearance rates for the alkaloids are roughly the same. 2. Case 12843. A 15 year old died of a congenital abnormality of the left coronary artery. No ephedrine was	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
reported in her system. How could this possibly be causal? 3. Case 10874. A woman with considerable substance abuse problems and some use of ephedra supplements for weight loss tested positive for phenylpropanolamine (PPA), but not for ephedrine or pseudoephedrine. The former is but a minor constituent of ephedra and products derived there from, but were a common primary component of a number of OTC weight loss products until recently. Such products are a much more likely source of the PPA than the "possibly causal" ephedra supplements.	
We have several concerns about the way the information is presented in the sections related to safety. It is AHPA's position that the report's safety assessment section reviews case reports from a passive event reporting system without fully and redundantly disclosing what has already been determined about the nature of the FDA's current AER system. Appropriate disclosures include, at a minimum, a reference to the GAO report on the subject and recognition that the Special Working Group of the office of Special Nutritional (FDA: Food Products Containing Ephedrine Alkaloids, Washington D.C., October 11th-12th, 1995) explicitly stated that such a system cannot, by it's nature, show causality.	We acknowledge that the case reports cannot show causality. We do not need to discuss the findings of other with respect to the adequacy of the FDA AER system. We assessed the information we did receive using explicit criteria, and our findings are reported.
AHPA recognizes that limitations in the clinical trial data lead one necessarily to consider case reports for an assessment of serious adverse events. The fact that there are no serious adverse events reported in any of the clinical study should however be stressed, even as it is identified as of insufficient statistical power to detect a rate of serious adverse events. This fact should be repeated at the Structured Abstract and in the Conclusions, for example, and the total number of patients in these studies (is that 2319 in the intervention groups?) should be identified. In addition, Table 17 should be expanded to include each of the serious adverse events that are subject to safety review in the draft (e.g. death, myocardial infarction, and stroke) and the number "zero" should be entered in both the placebo and intervention columns, if that was in fact the published observation.	We do not favor, as a general rule, adding rows or columns to a Table when all the entries in each cell will be the same. Such information can more expeditiously be conveyed in the text.
The Draft contains an extensive review of the specific AER case reports. This inherent emphasis presents an unbalanced appearance with respect to the intent of the original key questions. In comparison, assessment of efficacy is presented in a much more summarized fashion.	We cannot change the amount of space needed to describe what we did.
In arriving at criteria for judging the causal relationship in case reports of adverse effects from ephedra, the concept of biological plausibility is conspicuously absent. All else bring equal, adverse events that are biologically plausible (consistent with the mechanism of action of the drug in question) are more likely to be causally related to drug use	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
than those, which are not. Sudden death, myocardial infarction, and stroke are biologically plausible toxic effects of ephedra. This should be taken into account when interpreting the data.	
Clearly there is no right or wrong answer with regard to how much data is needed to judge the causality of an adverse event case report as "probable" or "possible" but I believe that the criteria used in this review have resulted in conclusions that that are understated. There is nothing wrong with the criteria per se, but using the term "probable" for a death that has fulfilled every review criteria (except re-challenge which is by definition impossible) understates the quality of the data and implications of the case. Similarly, the term "possible" for cases that have satisfied several but not all criteria makes it sound like these importance of these cases should be minimized, which I do not think is the intent of the report. For example, requiring negative angiographies to support causality for myocardial infarction has a rationale, but will necessarily exclude many or most cases because not all patients have this procedure.	
The importance of this wording is illustrated by the comparison of the Benowitz ephedrine data and your group's reanalysis of it, which would have downgraded so many case reports as to make the report unpublishable. Instead, It was published and shows a remarkable similarity in adverse event profile with the current report. This congruence of findings is in fact some of the strongest literature support for the conclusions of the current report regarding toxicities from ephedrine and ephedra, and these two reports suggest just that.	
The alternative to changing the terminology of the report is to provide additional commentary on the interpretation of the findings; that, in a view of 1) biological plausibility, 2) the considerable number of case reports emanating form a spontaneous reporting system, 3) similar toxic effects of pharmaceutical ephedrine, and 4) similar toxic effects of phenylpropanolamine, the findings of the current review are highly suggestive of a relationship between herbal ephedra and serious adverse events such as sudden death, myocardial infarction, and stroke.	
Discussion at the NIH requested that the title "possibly causal" was misleading and should be retitled to indicate more accurately the Rand staff interpretation that there is no proof of causality and while causality is possible, it is not probable.	
Study Selection. Study selection was not appropriate. Partly for the reasons stated above, but also because of the reliance placed upon passively collected anecdotal data. The mere fact that existing clinical trials contained "too few	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
(subjects) to allow adequate statistical power to access the rate of serious adverse events," does not make AER analysis any more reliable or probative. In fact, case reports cannot be "relied upon to assess serious adverse events," except, perhaps for the occurrence of rare disorders such as coronary artery dissection. The decision to include an analysis of AERs is particularly puzzling, given that the TEP chose to reject the Haller and Benowitz analyzing the same AERs (see "rejected articles" #195, record #116)! Heart attack and stroke are common disorders in our society, and thousands of ephedra product users would be expected to experience vascular events, even if ephedra did not exist. Analysis of AERs for common disorders, which are even more frequent among the overweight, is virtually guaranteed to show a connection with ephedra use, even if no such connection exists (for example, see the August 1 article, "Obesity and the Risk of Heart Failure" in the New England Journal of Medicine).	
Appraisal of Studies. Important parameters that could alter study results have not been systematically addressed. The brief discussion of obesity is confined to generalities. Obesity is a prothrombotic disease. [1]. Overweight people, presumably the majority of ephedra supplements users, are at greater risk for sudden cardiac death (SCD), and heart disease [2]. The report fails to provide any sort of epidemiologic prospective, leaving the false impression that the occurrence of these disorders among ephedra product users is somehow surprising or unanticipated. In fact, when the GAO wrote its highly critical analysis of the FDA's proposed rule on ephedra products, one of the issues raised was FDA failure to account for the reality that "there is almost always an underlying background rate for any clinical event in a population, regardless of whether there was exposure to a particular product..." The RAND report states that 3 billion servings of ephedra were sold in 1999. Assuming that 3 servings are used per day for 12 weeks (as Haller and Benowitz do in the NEJM paper), then there were 12 million users. The accepted rate for sudden death, heart attack, and stroke in the U.S. is 0.1, .5, and .2 percent per year respectively [3]; which means that even if ephedra/ephedrine has absolutely no relationship to any of these disorders, 12,000 cases of SCD, 60,000 cases of myocardial infarction, and 24,000 cases of stroke would still be expected among ephedra users each year. Not providing this information to general readers paints a completely misleading picture and leads to a misinformed, if not false, impression of relative risk. It also repeats the same FDA error already criticized by the Government Accounting Office.	
Data Collection. No effort is made to reduce bias in the	The reviewer is incorrect. The autopsy was

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
data collection, or even to assure that the data is valid. This is immediately apparent in the discussion of the first AER. The history given for AER #13914 is simply incorrect. The autopsy was, in fact, included in the docket (copy attached). The heart was actually examined by a consultant cardiac pathologist, and that report states "the cause of death may be attributed to myocarditis in the absence of other demonstrable cause." The summary in the report misstates the data available and misrepresents the conclusion in the report.	not included in the docket we received.
Data synthesis. Limitations of the review process are not adequately stated. Many, if not most, of the interpretative problem seem to be the result of the medical experts on whom RAND relied upon the review the AERs (it is not clear from the draft report who reviewed which AERs). Medical-legal death investigation is customarily performed by pathologists with specific training and expertise in sudden death investigation, yet it appears that not a single pathologist was included among the reviewers. As a consequence, many of the AERs were almost certainly misinterpreted. I do not have access to many of the AER files that RAND reviewed, so I cannot comment specifically on RAND's subjective assessments of causality in most of the cases. However, the errors and misinformation in the AERs that I can check show clearly this experienced death investigators were not involved in the project. For example, Case #14390, classified by the panel as "probably causal", was said to have had a "shunt" in place. It follows that the brain could not, as the report states, possibly have been "normal". For one thing, there would have been a shunt in place, which is decidedly not normal. There would have been tissue reaction around the shunt, both the heart and brain. If the shunt had been placed for traumatic injury, trauma residuals would have present. If the shunt had been placed for traumatic injury, trauma residuals would have been present. If the shunt was for congenital hydrocephalus, it is quite likely that the abnormalities associated with the Arnold-Chiari malformation would have been present. Sudden cardiac death secondary to acute obstructive hydrocephalus is well recognized by forensic pathologists [4, 5]. Even if a pathologist was on the AER review panel, the methods section is so vague that it is not clear whether the pathologist would have been asked to review this AER. One can only conclude that the panel of reviewers do not know the accepted causes of sudden death in young people with ventriculatrial shunts. Evidence of potential bias in the AER reviews is provided by the AERs chosen as "probable" and "possible" as well as by the summaries of the AERs contained in the draft report. For example, Case #12485 was classified as "possible causal," and the presence of	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
"mild cardiomegaly," noted. The autopsy report in the docket shows a heart weight of 490 grams, two standard deviations above predicted [6]. No one with training in vascular pathology and sudden death investigation would call that degree of enlargement "mild." Cardiac enlargement is a recognized risk factor for SCD [7], over and above the severe coronary artery disease that was also present [8]. Referring to such an important pathologic finding as "mild" can only be explained in two ways: either the reviewers were unaware of the significance of this abnormality, or they were trying to minimize the finding because it did not fit a bias towards finding evidence of ephedrine toxicity. Preconceived bias is strongly suggested by the inclusions of AER #12722, a child who died of a type of congenital heart disease where the left coronary artery arises from the left pulmonary artery (not vein, as stated in report). No ephedrine was detected in tissues analyzed by the FDA, and there was extensive scarring of the myocardium, reflecting early episodes of healed myocardial infarction. By their very nature, the morphologic changes detected, which almost certainly were the cause of death, antedate by weeks or months the alleged history of ephedrine ingestion. Classification of this case as "possibly causal", violates the report's own stated criteria, which specifically reserve the "possible" category for those cases where "another condition by itself could have caused the adverse event, but ephedra use may have helped precipitate the event." No ephedrine was detected at autopsy, and anatomic changes were present that had to have occurred weeks or months before the first use of any ephedra product is even alleged. Bias is also suggested by the discussion of phenylpropanolamine. The section on Phytochemistry correctly states that the phenylpropanolamine (PPA) content of ephedra is low. But then in the discussion of AER #10874, the report also states "there is a described association between PPA and cerebral hemorrhage, PPA is also a component of some herbal ephedra." Had a balanced presentation been intended, the report would have provided the additional information that the most PPA ever detected in a serving of an ephedra supplement was half a milligram, and that in studies with volunteers, 50 mg of PPP is needed just to modestly raise blood pressure [9]. The emphasis of the potential linkage between PPA and ephedra in the report is distressing for three reasons. Firstly, it reiterates the same unproven argument propounded by Drs. Haller and Benowitz in their most recent publication (Haller, et al., Clin Pharmacol Ther 2002; 71:421 432). The unquestioning repetition of an unproven hypothesis tends to lend legitimacy to that hypothesis, even in the absence of evidence, thereby detracting from the value of the report. Secondly, and quite improperly, the	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
report nowhere discloses that this unproven mechanism of toxicity is, in fact, an unproven theory propounded by one of the TEP members. Thirdly, the report makes no mention of the dispute between professional epidemiologists over the validity of the study that lead to the withdrawal of PPA in the first place. Indeed, the entire discussion of PPA in this report, the failure to mention the PPA content of the average herbal supplement, the failure to mention the fact that ephedrine is minimally metabolized to (-) norephedrine, and the failure to mention existing disagreements among epidemiologists about PPA, is simply not scientifically supportable. A consumer would have to simultaneously ingest 100 servings of an ephedra supplement in order to receive enough PPA to minimally raise blood pressure, and probably twice that amount to cause a clinically significant increase. In AER #14019, death was attributed to "dissection of a left anterior coronary artery" in a 26-year-old woman. This is not a supportable conclusion. Coronary artery dissection has never been reported in an ephedrine user, or even in amphetamine abusers (a drug to which ephedrine has frequently been compared). Almost all reported cases of spontaneous coronary artery dissection involved young women following childbirth, or in cocaine users. Had the decedent just delivered? Was she a cocaine user, or both? Was toxicology testing performed? This case and others classified as probable and possible are examples where "crucial information is missing" and they should have been classified as "insufficient".	
Safety assessment. As mentioned above, the panel should also include the non-placebo controlled and non-randomized trials in the safety assessment. The information obtained from such trials is superior to that from case reports. Page 58, 4th section: Here it is stated that patients taking pharmaceuticals outside of clinical trials may have a greater risk of certain adverse events than patients selected to participate in clinical trials. I strongly disagree. In all the clinical trials we have conducted, which have been conducted in Denmark, it is quite normal that the patients are referred by general practitioners or hospital departments because they have a high degree of overweight (are typically obese, with a body mass index of 29-40) and suffer from complications to the obese state. This may not necessarily be ischemic heart disease, heart failure or type 3 diabetes, because these subjects will typically be excluded, but patients with pre-diabetes, dyspnoea, osteoarthrosis in knee or hip, etc. In addition, one of the large trials was conducted on the hypertensive obese patients (Ingers, et al.). In contrast, individuals in the community taking preparations containing ephedrine will typically be less overweight and be	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
generally healthier. They will be less likely to experience serious adverse events than subjects in clinical trials I think the conclusion reached by the panel should therefore be reversed.	
The Danish Experience. It is quite natural that the panel has received the list of case reports from FDA's office of nutritional products, labeling and dietary supplements. However, why did the panel not ask the Danish FDA for their full report of collected adverse events during the 12 years from 1990 to 2002 where an ephedrine/caffeine prescription compound has been on the market in Denmark? This is a substantial body of experience that could give more valid conclusions than those received the American FDA alone. During the last 8 years, the defined day doses have ranged between 3.6 and 4.6 per 1,000 inhabitant/day in Denmark. It also means that the Danish Drug Administration has, in its surveillance program, obtained anecdotal data regarding reported side effects from General Practitioners and other Doctors in Denmark. The post market surveillance program is very effective in Denmark and there are 134 reports of side effects, but they are all very mild side effects and the all the well-known side effects we know from the pharmacological action of ephedrine/caffeine. They include tremor, insomnia, palpitations; side effects from the use of ephedrine/caffeine even though Denmark has had a substantial amount of sales and ten years experience.	
First there is no control group. Because there is often an underlying baseline risk of disease unrelated to exposure to a product, there will be events reported in people exposed to that product that are in no way associated with use of the product. Given the large number of users of ephedra to a product, there will be events reported in people exposed to that product that are in no way associated with use of the product. Given the large number of users of ephedra-containing products,4 there will be events among users that are coincidental with the use of ephedra (i.e., not causally related) even in the absence of other explainable causes of these events. In an analysis performed by the Expert Panel of the Ephedra Education Council (EEC), an estimate of the rate of serious events in ephedra users was compared with the rate of events expected in the general population. 5	
Although this analysis was not designed to rule in or out a possible cause-and-effect relationship between ephedra and the outcomes evaluated, it did suggest that the adverse events reported among ephedra users may very well represent simply the background rate of events expected among such a large number of users of ephedra-containing products, unrelated to ephedra itself. This was true even under assumptions that inflated the risk from ephedra and	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
even assuming the events reported to FDA represented a small percent of the adverse events occurring among ephedra product consumers.	
I. Research Parameters - Overview. Based upon my review of the draft questions provided to RAND, it is my understanding that RAND was asked to provide a science-based review in order to identify gaps in the data relating to ephedra safety and efficacy. Identification of these data gaps was deemed essential to prepare an agenda for future research. Despite this charter, RAND instead acknowledged that it engaged in a subjective review of anecdotal adverse event reports in order to ascribe potential causality to such reports. As explained below (see Section II, herein), the inherently subjective nature of such a review by definition reduces the objectivity of the report and may lead one to question the proposed research agenda. It is my strong belief that RAND should issue a report that is entirely objective and science-based. A subjective review of anecdotal adverse event reports that attempts to ascribe potential causality to such reports in my opinion provides no meaningful benefit with regard to identifying data gaps and developing a research agenda.	
I. Research Parameters - Overview (cont'd). If the goal of RAND's review is to identify data gaps, RAND should address this issue by reviewing the scientific studies and the types of anecdotal adverse events that have been reported - which should be used as a signal to identify additional research projects. Ascribing degrees of causality to poor data is not helpful in this regard. Moreover, by including a subjective component to the review, and indeed emphasizing this component, one does not control for potential inadvertent reviewer bias. One of the reasons placebo controlled double-blind trials are considered to be the gold-standard for scientific research is that researcher bias must be accounted for and eliminated to the greatest degree possible. It is understood in the scientific community that despite the best of intentions, inadvertent bias can impact researcher conclusions.	
I. Research Parameters - Overview (cont'd). RAND acknowledged that its review of FDA's AERs was subjective, and even focused on potential publication bias by certain researchers and organizations, yet failed to even acknowledge the possibility that inadvertent bias could potentially impact its own report based upon the subjective nature of the review of adverse event reports. I am not in any way alleging that the draft report is actually biased - either intentionally or inadvertently. Rather, my point is that a subjective review is subject to inherent inadvertent bias and therefore the report should focus on objective information. If the goal of RAND's report is to identify data	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
gaps and suggest areas for additional research, ascribing degrees of potential causality to anecdotal adverse event reports appears to be entirely counter-productive. Finally, the General Accounting Office and Department of Health and Human Services (along with the FDA) have already acknowledged that it is not possible to determine causation based upon a review of these anecdotal adverse event reports.	
I. Research Parameters - Overview (cont'd). In light of this determination, it is difficult to comprehend the benefit that can result from RAND attempting to ascribe degrees of causation to such reports. From an objective perspective, the adverse event reports should not be used to ascribe degrees of causation - but also should not be ignored. Rather, as noted, the reported adverse events should be documented and identified and then used to help target endpoints for future research.	
The draft report's causality assessment is not consistent with how other Agency for Healthcare Quality and Research (AHRQ) reports have addressed dietary supplement adverse events. From our review of other reports, the draft reports causality assessment is unprecedented. The central question is why RAND conducted a causality assessment of the AERs and prominently reported it admittedly subjective attributions of causality. RANDs causality assessment obscures the objective findings of the report to such a degree that the research agenda RAND recommends will not be achievable. Further. Rends subjective attributions do little or nothing to answer the questions that RAND has been asked to address. In a broader context, the final report, if published with the causality assessment, as it now exists, will threaten future support for similar reviews of other dietary supplements. Despite serious concerns, the draft report contains a very worthwhile and comprehensive review of the objective data on ephedra and ephedrine, and the recommendations for future research are commendable and attainable, and will serve the valuable function of answering important questions concerning ephedra products. The final report can become a monument for how to address controversial safety issues such as those that exist for ephedra, provided the final report focuses on objective science rather than subjective assessments of the AERs. The purpose of the RAND review is to perform an objective review of the science pertaining to ephedra in order to answer specific questions from the current data, and to identify research gaps and recommended additional research to answer the questions where the current data are insufficient or do not exist. All parties to the ephedra discussion agree that the AERs raise serious questions that	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
deserve serious attention. Industry has enthusiastically supported the concept of the RAND review for that reason. However, as the draft report recognizes, the AERs do not represent the objective data. RAND's review of ephedra is necessary in large part to resolve the controversy that has been created by widespread reporting of the subjective assessments of the AERs on ephedra and the conflicting opinions as to what the AERs mean. Informed critics and supporters of the ephedra agree that the AERs cannot be used to assess safety, to establish whether there is in fact a risk of serious adverse events, or to quantify that risk if the risk exists. These limitations on the use of the AERs for ephedra were the focus of the HHS and the Food and Drug Administration's (FDA's) recent statements on ephedra on June 14, 2002 1) and have also been noted by critics such as Drs. Haller and Benowitz, as well as industry-supported panels such as the Expert Panel of the Ephedra Education Panel. 1)"The primary purpose of a voluntary adverse event reporting system is to generate 'signals' of potentially related events, rather than assessing product safety...There are situations when background rated of the observed event are so rare or unusual that, in combination with physiologic responses and biologic plausibility, A significant relationship between the events is self-evident from the reports in a voluntary reporting system. However, the FDA has advised me that the types of observed outcomes reported in relationship to the ingestion of the ephedrine alkaloids are not uncommon in the general population and there for the reports alone do not provide a scientific basis for assessing the safety of ephedrine alkaloids or establish a link between the reported adverse events and the ingestion of ephedrine alkaloids #2 "[O]ur report describes a series of cases which the use of ephedrine-containing dietary supplements was associated with adverse cardiovascular events. Our report does not prove causation, not does it prove quantitative information with regard to risk. A large-scale case-control study similar to the Hemorrhagic Stroke Project for phenylpropanolamine is needed to determine the risks associated with these dietary supplements" Christine A, Haller & Neal L. Benowitz, Dietary Supplements Containing Ephedra Alkaloids: Letter to Editor, 344 New Eng. J. Med 1095,1096-1097 (2001). #3 the consensus conclusions of the EEC Expert Panel, as well as extensive reviews of the published literature and the most comprehensive review of the AERs that has been conducted to date, were submitted to the FDA's published docket on ephedra in October 2000 and were made available to RAND. Since the Expert Panel Report is nowhere mentioned is referenced in RAND's draft report, a copy is included with these comments. The Expert Panel	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
concluded that the available data and information including the AERs, "do [] not demonstrate an association between the use of dietary supplements containing ephedrine alkaloids and serious adverse events when used according to the American Herbal Products Association (AHPA) trade recommendation for ephedra products." Expert Panel, Ephedra Educ. Council, Comments of the Ephedra Education Council on the Safety of Dietary Supplements Containing Ephedrine Alkaloids and on the AERs and Health Assessment Released by the FDA on April3, 2000 on 6 (2000) (on file in FDA docket 00N-1200 as C30) RAND has also concluded that, because of the subjectivity of assessing causality from AERs, further analysis is additional case reports will not lead to any objective scientific conclusions and would not be useful to establish whether there is any causal connection between ephedra and the type of serious adverse events. Continued analysis of case reports cannot substitute for a properly designed study to assess causality." Draft Report at 5. Nonetheless, RAND has conducted a causality assessment of some of the AERs deemed to report serious adverse events and has described the results of this review in terms of "probable" and "possible" causal relationships between ephedra, death and other serious events. RAND's findings are presented in a way that will expand the controversy surrounding the AERs on ephedra rather than resolve it through objective analysis and recommendations for additional research. Further, the draft report is open to the inappropriate interpretation than RAND has concluded, based on the AERs, that ephedra causes serious adverse events, and that is exactly how the ephedra critics will interpret the draft report. This interpretation is a result of the wording of RAND's findings as well as the presentation of the causality findings without necessary context. The headlines that will result concerning RAND's findings on causality will make is difficult if not impossible to justify the research agendas that RAND recommends at the very end of the draft report. Again, given the lack of any scientific value to RAND's analysis, the ability to discuss the "signal" that the reports raise without reporting subjective attributions of causality, and the importance of the report as a means to resolve the controversy that the AERs have created, the best solution is to remove the causality from the assessment report. The problems created by the RAND causality assessment are aggravated by a number of factors: 1) The vast majority of the draft report's text and tables relating to safety address AERs (10 pages of text, 20 pages of tables) rather than clinical data (one page of text, three pages of tables), even though RAND acknowledges that clinical data are if much greater value, and even though in the end of the report	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
RAND concludes that causality assessments do not provide objective, science-based answers to the questions that it was asked to address. Although the draft report discusses a review of selected AERs that was conducted under the contract with FDA by Drs. Haller and Benowitz, and the draft report compares the results of this review to the RAND review, the draft report excludes any mention of an extensive review by Cantox Health Sciences International, and the most comprehensive review of the AERs to date by the Expert Panel of the EEC. A copy of the Expert Panels Report is enclosed with a hard copy of these comments. While the issues of publication bias and other coursed of bias were carefully analyzes in the draft report for published studies, other than a brief mention of the subjectivity of RAND's causality assessment, the draft report does not adequately address the potential for reviewer bias in RAND's causality findings. Because the AER files that RAND reviewed not the same as those that FDA had provided to the public, and RAND has not made its AER files available to the technical expert panel or the peer reviewers, none of these reviewers will have the ability to do an in-depth analysis of RAND's causality assessments. Some reviewers may have their own files on some of the AERs, but this will not permit the type of analysis that would be needed for a thorough peer review. There is a simple solution to the problems that the RAND causality assessment, and the manner in which the assessment is presented in the draft report, have created. The causality assessment should be removed from the report. The final report will then be focused on the objective assessment of the data and answering the questions that RAND was asked with addressing. In addition, dropping the causality assessment will permit the recommended research agenda to proceed as intended. To do otherwise places the whole RAND ephedra project in jeopardy, as well as future support for similar projects. If RAND does not remove the causality assessment from the report, RAND should drop the attributions assigned to specific reported as either probably causally" or "possibly causally" related to ephedra consumption. Instead, RAND should state the certain reports were reviewed for causality and that those reports raise a "signal" that additional research is needed to address that signal. This would be a neutral and objective statement that would be accepted by parties, and would be consistent with the recent statements of HHS and FDA on the inability to make safety determinations or regulatory decisions based on the AERs for ephedra.	
However, in the opinion of the reviewer, those conclusions	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
regarding the case reports are limited by a combination of the conservative causality assessment criteria and the limited medical records and toxicology data available for most case reports. For example, hypertension was defined as a systolic pressure in excess of 180 and a diastolic pressure in excess of 105. Also, no consideration appeared to have been given to the contribution multi-component ephedra-containing dietary supplements might have had in those individuals with underlying cardiovascular or cerebrovascular disease. I think it is generally accepted among the medical and scientific community the presence of sympathomimetic agents could potentially exacerbate the likelihood of adverse events in such populations? I would think most clinicians would factor such information into their differential diagnosis rather than dismiss them altogether.	
Adverse Events Reports: One limitation with your approach taken in evaluation of the adverse events reports made to the FDA is that your causality algorithm does not include an assessment of whether ephedra played a contributory role on the adverse event. Because ephedra is available as a dietary supplement, it is likely that many persons taking these products are not using them under doctor's supervision, and may have medical contraindications to their use. Therefore, the role of underlying disease becomes a crucial factor in causation assessment, particularly when a potential risk factor often goes undetected (i.e. essential hypertension, structural heart defect), or when a condition is omitted from the ephedra product label warning (i.e. family history of premature CAD, sickle cell trait).	
Two AERs that you assess as no higher than possibly causal illustrate this point. AER 12485 did indeed have a moderate degree of coronary artery disease detected at autopsy. However, he was reportedly in good health without history of angina, and had been jogging regularly without adverse effects. Because he collapsed suddenly after returning from jogging, we felt this was a primary arrhythmic event due to ephedra. Similarly AER 12843 was a healthy, adolescent who had participated in competitive sports for many years. She had appeared to have been well-compensated for a serious underlying coronary artery abnormality that was clinically undetected since birth. Only with use of Ripped Fuel, did she suffer a catastrophic cardiac event resulting in death. We felt that the cardiac stimulant effects of ephedra resulted in myocardial ischemia in this case.	
It would be helpful to specify what degree of pre-existing coronary artery disease would constitute a significant risk factor to result in myocardial infarction or sudden death in the absence of stimulant use, thereby ruling out ephedra in the causation assessment. (page 32 of chapter 2	This case was reviewed by a cardiologist who made this judgment.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
methodology). In the case of AER 14530 (page 63), I would disagree that 20-30% stenosis would be significant enough to result in acute M.I. in a 43-year-old female smoker without a significant contributory effect from the ephedra alkaloids in Metabolife.	
On page 59 I suggest the authors be slightly cautious with their use of language such as "probably causal" and "no other possible explanation." The latter phrase is particularly troubling. What the mean is no other explanation that they could identify. Similarly, on page 69 they state that there are a certain number of cases of serious adverse events that "cannot be explained by causes other than ephedra use." While I do not deny that it is extremely likely that many cases of adverse events happen due to ephedra use, simply because we do not have in our hand an explanation of why an event occurred other than a particular explanation under consideration, does not mean that the particular explanation under consideration is the correct one.	We have revised this language.
Because of the paucity of large randomized trials, evidence concerning stroke and ephedra by necessity consists of analysis of case reports. "…An assessment of case reports is insufficient to warrant definite conclusions regarding causality." (p.110) Nevertheless, arbitrary criteria are used to define "probably cause": documentation that a stroke occured, that ephedra was used, and that there was exclusion of other potential causes. The definition may be too liberal. Of ischemic strokes in relatively young adults (i.e. those <50 years old), perhaps 20-35% are "idiopathic" despite thorough evaluation. The definition implies that all idiopathic strokes would be classified as "probably causal" if ephedra was used in any dose in proximity to the event. Given the frequency of idiopathic stroke, many (perhaps most) neurologists would consider "possibly causal" to be a better designation in this situation. Are there specific clinical circumstances in which the relationship of ephedra use and idiopathic stroke could be certain? Perhaps if acute, striking elevation of blood pressure were known to precede the stroke onset of of angiographic features characteristic of vasospasm were present in the abscence of migrane? Arbitrary to be sure, and not very helpful.	
The evidence for effectiveness supports the conclusion. Except for AERs, however, little evidence of toxicity is actually provided, and evidence of safety has been largely ignored. No evidence is provided to even suggest "a possible causal role of ephedrine-containing dietary supplement in rare, but serious events," let alone extremely common events such as heart attack and stroke. Even critics of ephedra have concluded that the clinical effects of pharmaceutical ephedrine, and the ephedrine contained in	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
herbal preparations, are indistinguishable. Gurley states that the increased incidence of ma huang toxicity does not stem from differences in the absorption of botanical ephedrine compared with synthetic ephedrine." Haller and Benowitz, in their most recent publication, conclude, "Botanical stimulants have disposition characteristics similar to their pharmaceutical counterparts..." The Cantox Report, and the Report of the Expert Panel of the EEC reached similar conclusions. Since there are no real differences, studies demonstrating the safety of pharmaceutical ephedrine and pharmaceutical ephedrine in combination with pharmaceutical caffeine, should not be excluded when considering the safety of herbal equivalents. The explanation most frequently offered for alleged cases of ephedrine-related stroke is drug-induced blood pressure elevation, this in spite of the fact that no clinical trial, of any duration, has ever demonstrated that a clinically significant effect on blood pressure exists. Indeed, the studies that have addressed this question. including the most recent paper by Drs Haller and Benowitz, have shown diminishing cardiovascular effects over time. In other words, if dangerous blood pressure elevations do not occur with the first dose ephedrine, they are even less likely to occur with prolonged dosing. These studies should be included in the RAND review of ephedra and should be used to address the question of potential increases in blood pressure and other safety issues. Ephedrine has been studied in more than 50 double blind, placebo-controlled clinical trials, some of long duration. A far from exhaustive literature search produced the attached list of peer-reviewed, published, clinical trials. Most have compared ephedrine to placebo, and to other sympathomimetic drugs used to treat asthma. However, others have involved smoking cessation, sexual function and athletic performance. Nearly a dozen of these trials involved caffeine/ephedrine combinations using does exceeding those found in herbal supplements. In total, more than 2000 individuals have been enrolled in these trials. In several studies there was even continuous cardiac monitoring in middle-aged patients with known heart disease; no effect was observed. No clinically significant episodes of toxicity were reported. Including these and other studies on ephedrine that have been excluded from the RAND review will increase the power of the safety calculations that can be derived from clinical data. One of the major limitations of the report was the composition of the TEP and the reviewers who made subjective assessments of the AERs. Given the importance placed on assessment of AERs, it is unfortunate that no pathologist was included in the view or on the panel. The lack of expertise is obvious from the comments made about	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
the individual AERs. The failure to provide information about any potential conflicts of reviewers also detracts from the study. Why were the findings of Expert Panel of the Ephedra Education Counsel not considered? The analysis of this panel was in some ways unique, as it is the only consensus opinion on ephedra safety. In addition, this panel conducted the most comprehensive review of the ephedra AERs to date, and yet the causality assessment, which conflicted the findings of the draft report, are not even mentioned. If RAND believes that the EEC review and analysis was, in some way scientifically flawed, then the reasons for that belief should be stated.	
The danger of drawing conclusions from AERs without a control group can be illustrated by examining data from randomized trials in which participants are blinded to whether they are receiving the study treatment or inactive placebo. In the placebo group of a recent randomized trial,2 there was an increase in ventricular couplets (extra heart beats) at the 4th week of the study 9from 3% at baseline to around 14% at week 4). This is, of course, not due to the placebo, which is inactive, but rather just spontaneous ventricular couplets that occurred by chance. However, if these participants had been given ephedrine alkaloids in an uncontrolled study (without placebo), this change could have been attributed, incorrectly, to the ephedra. That is, these could be AERs that were attributed to ephedra. In fact, in the controlled trial, a similar increase in ventricular couplets was not seen in the ephedra/caffeine arm. Another example is the 15-year-old female (case 12843) with Bland-White-Garland syndrome who died while playing soccer. This disorder has been associated with sudden death after physical exertion. In the absence of a unique pathologic process, it is almost never possible to establish a causal association on the basis of adverse event reports. There is nothing pathologically or diagnostically unique about the adverse events noted in the ephedra database (e.g., myocardial infarction, stroke) that allow one to distinguish a spontaneous event from one caused by use of Ephedra products. In fact, a review of all autopsy data from ephedra AERs by Dr. Grover Hutchins, a Professor of Pathology at the John Hopkins University School of Medicine and member of the Expert Panel of the EEC, concluded that "The pathology data available do not show any pattern consistent with ephedrine alkaloid-containing dietary supplements as a cause of death."5 Similarly, 10 participants in the ephedra/caffeine group (12% of these participants) withdrew because of cardiovascular symptoms (palpitations, elevated blood pressure, arrhythmias). If there were no control group, these also might have been attributed to the ephedra/caffeine combination. However, the same	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
proportion of participants in the placebo group (13%) withdrew for the same reasons. The withdrawal rate in the ephedra/caffeine group thus was consistent with the background rate of these events in a placebo group unrelated to ephedra use. A second limitation of adverse event reports is that other potential causes of the event are often present, making it extremely difficult to determine if an event truly is related to the exposure. As an example, it is well established that physical exertion can trigger myocardial infarctions and cardiac arrest (up to a 74-fold increase in the risk of sudden death, according to a recent report in the New England Journal of Medicine6). Therefore, a case such as the 38-year-old man with three vessel coronary disease and a dilated heart who died after jogging could very well have been related only to the physical exertion and not the ephedra. According to his wife, his heart problems had been known for at least five years. In addition, this man had been taking an ephedra product for one year without apparent ill effect. The third limitation of AERs is that their interpretation remains subjective. Even experts in the field will disagree about the possibility that an event may or may not be due to an exposure. The RAND report (Table 22) displays a comparison of the causality assessment between their review panel and a published report by reviewers for the FDA.3 In only two out of 20 cases that both groups reviewed was their agreement about the highest level of possible causality. For example, with respect to the 38-year-old discussed above, the Haller and Benowitz review recorded this event as "Definitely or probably related" to ephedra while the RAND report classified it as only "Possibly causal." Another review of AERs performed by Dr. Theodore Farber and Dr. Norbert Page, members of the Expert Panel of the EEC, reveals similar disagreements. The two "probable" cases reviewed by EPC were rated as "Low Possible" (case 12980) and "Improbable" (case 13418) by Drs. Farber and Page. The fact that the etiology of events can be debated simply illustrates the substantial limitation of case reports that lack a comparison control group. This echoes reviews done by FDA and its consultants in which agreement about causality was poor. For example, two consultants from FDA, Drs. Ricaurte and Stoll, reviewed 28 AERs related to neurological events. Dr. Ricaurte classified eleven cases as "attributable" while Dr. Stoll classified only five as "attributable."5 Only two of the consultant's findings overlapped - that is, there were only two cases that both Dr. Ricaurte and Dr. Stoll agreed should be categorized as "attributable."5 This disagreement is not a flaw of the reviewers, but rather a flaw of AERs. A fourth limitation of AERs is that ingestion of the substance	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
in question cannot always be substantiated. For example, case 10276 is a 32-year-old with an enlarged heart who was found dead in his truck. Although a product that contains herbal ephedra was found in his truck, so were several bottles of cold medications, including Nyquil. Toxicology revealed no ephedrine, but did identify pseudoephedrine and doxylamine, both components on Nyquil. Thus, there is no evidence that this person even ingested the herbal ephedra. A similar case (13096) revealed no ephedrine in a toxicology screen again suggesting that the man had not ingested ephedrine around the time of death. Equally importantly, this man died of a disease that appeared to run in his family (aortic dissection), including an 18-year-old niece. In summary, AERs cannot be used to assess causality. As stated by several authors with experience in interpreting adverse event reports for the FDA: "It is probably impossible to do comparative analyses employing ADE [adverse drug event] reports for drugs that have received extensive publicity in the mass media for an adverse event..."7 In fact, the FDA's Center for Drug Evaluation and Research pointed out, in a February 10, 2000 memorandum concerning ephedra products, that "it is possible that the reported serious adverse events are reflective of coincidental background spontaneous occurrences in the population and are not necessarily causally related to [the use of dietary supplements containing ephedrine-type alkaloids]." The RAND review notes this as well, stating that "The most important limitation [of their assessment of adverse events] is that the study design, that is an assessment of case reports, is insufficient to warrant definite conclusions regarding causality." They list this limitation as one of the "most important...gaps" in the current knowledge-base. The RAND review also states that "Disentangling the relative importance of the pre-existing condition and the ephedra use is not possible." They state further that "Continued analysis of case reports cannot substitute for a properly designed study to assess causality. A case control study would probably be the study design of choice." Their Technical Expert Panel also "judged that case reports alone would be insufficient to establish definite causality between ephedra use and serious adverse events." Because of these limitations, terms such as "probably causal" and "possibly causal" in AER reviews are potentially misleading (see, in particular, the "Conclusions" section of the "Structured Abstract," page vi, and the "Conclusions" section, page 112). They represent only reviewers' assessment of causality based on uncontrolled data and subjective assessments. Although these terms are often used in scientific publications, their use in the RAND report may suggest a level of evidence that does not exist from the	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
current data. These statements, therefore, should not be taken out of context. RAND's stated limitation that "Definite causality cannot be determined from case reports" must be kept in mind when interpreting this report. It is also critically important to remember that case reports can produce false signals of cause and effect.8,9 Most importantly, because of the limitation of AERs, it is unclear why the RAND report states, in the Structured Abstract, that "These [sic] is sufficient evidence to suggest a possible causal role of ephedra-containing dietary supplements in rare but serious adverse events, particularly cerebral hemorrhage." With respect to a possible causal role of ephedra in adverse events, RAND acknowledges that AERs are not sufficient to draw conclusions about causality, consistent with the known limitations of AERs discussed above. In addition, this comment is puzzling given that their "Conclusions" section (ages 112-113) does not state this at all. With respect to cerebral hemorrhage, the only comment in the body of the report on cerebral hemorrhage refers to a case-control study of phenylpropanolamine (PPA) and cerebral hemorrhage.10 However, the RAND report refers to this study only as an example of case-control methodology that could be applied in the future to ephedra. The report does not discuss the PPA study further. In fact, this study has been heavily criticized. Despite this, there was not a formal review of this study by RAND (and there were no members of the technical expert panel listed with expertise in epidemiology to perform such a review). In addition, PPA and ephedra have different chemical structures and different pharmacological activities. Finally, the RAND report does not mention that the PPA report also presented data on non-PPA, ephedrine-alkaloid containing products. These agents included medications that contained pseudoephedrine hydrochloride, phenylephrine, ephedrine, and epinephrine. In the report, there was similar prevalence of use of these products among those with and without hemorrhagic strokes. Although this is not a definitive analysis, it suggests that there was no association between these ephedra-containing products and hemorrhagic stroke. Therefore, the statement in the Conclusions section of the Structured Abstract of the RAND report is inconsistent with currently available scientific data. In summary, the RAND review supports the use of herbal ephedra and caffeine for weight loss, an effect that may have beneficial health consequences. The report also suggest, from controlled studies, that adverse events following ephedra use are, at most, rare. (The should not imply that the events are even causally related to ephedra us.) Most importantly, because of the reliance on AERs, the report cannot establish a causal effect of ephedra on	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
serious adverse events.	
Review of Anecdotal Adverse Event Reports - Limitation of Review. On pages 109-110, the report identifies potential limitations associated with the review of anecdotal adverse event reports. These limitations are buried at the end of the report, rather than being incorporated into the appropriate sections of the report (as per prior AHRQ reports, such as the Garlic Report - see Section IX, herein). Moreover, a number of limitations are not prominently identified, including but not limited to: a) The poor quality of the data and information contained in the anecdotal adverse event reports; b) Inherent problems associated with voluntary reporting systems; c) The number of individuals in the general population who experience the adverse events identified (i.e. background risk); d) The possibility that certain products may not have been standardized and/or manufactured according to Good Manufacturing Practices ("GMPs") - resulting in potential adverse events that have nothing to do with ephedra or caffeine when consumed in recommended amounts. Specifically, in the absence of standardization and quality control, it is conceivable that certain products may contain far more ephedra or caffeine (or other constituents) than indicated on the product label. This possibility must be considered when evaluating anecdotal adverse event reports. In the absence of identical product identity, any general conclusions regarding ephedra and caffeine are inappropriate and highly suspect based upon adverse event reports (see the AHRQ Garlic Report for an appropriate way to address this issue). In my opinion, RAND should strongly support immediate issuance by the FDA of dietary supplement GMPs and should endorse stringent quality control measures to ensure that all ephedra supplements contain what they are claimed to contain; e) The possibility that the consumer abused or misused a product by ingesting more than the recommended amount - or that the consumer ignored detailed product warnings and contraindications. RAND should emphasize the detailed warning label contained on the vast majority of ephedra supplements - and should acknowledge that there is little way to know from anecdotal data whether a consumer abused a product (either intentionally or more likely inadvertently); f) The possibility that the anecdotal adverse event reflects chance, coincidence , or confounding factors - including but not limited to the possibility that ingestion of a different product or substance led to the stated event. I I. Review of Anecdotal Adverse Event Reports - B. Ascribing Causality to Adverse Event Reports - 1. Reliance on Unpublished Article. In order to establish a framework for analyzing the adverse event reports, the draft report relies upon an unpublished article written by Cynthia	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
Mulrow, M.D. Reliance upon an unpublished, non peer-reviewed article to establish the framework for a critical portion of the report is entirely unacceptable. AHRQ studies have not, to our knowledge, ever relied upon an unpulished article to establish the framework for this type of analysis. In addition, reviewers such as myself have no way of analyzing the article - thereby defeating one of the primary reasons for review of the draft report. II. Review of Anecdotal Adverse Event Reports - B. Ascribing Causality to Adverse Event Reports - 2. Failure to Review Factors Critical to the Interpretation of Anecdotal Adverse Event Reports. Table 4 of the draft report summarizes the various methods researchers use to assess causality from adverse event reports. The following nine factors are identified: a. Temporal relationship. b. Dechallenge response. c. Rechallenge response. d. Could there be an alternative explanation? For example, dehydration or consumption of other toxic substances. e) Prior reaction to same substance. f) Dose response. g) Objective evidence of adverse event. h) Previous conclusive reports. i) Definition of substance. Despite the report's reference to these nine factors, it is my understanding that the report concludes that events are "probably causal" based upon a review of only two factors - a and g. The draft report does not explain why RAND believes only two factors out of nine can be used to ascribe degrees of causation to anecdotal adverse event reports. II. Review of Anecdotal Adverse Event Reports - B. Ascribing Causality to Adverse Event Reports - 2. Failure to Review Factors Critical to the Interpretation of Anecdotal Adverse Event Reports (cont'd). The draft report also indicates that three factors are used to determine if an event is "probably causal": a) Reasonable certainty that the adverse event occurred. b) Reasonable certainty that the patient took ephedra in a dose and timing compatible with the known pharmacology of ephedrine. c) An adequate evaluation must have been done to rule out other potential causes for the adverse event. The third factor (factor c, above) is exceptionally problematic from a scientific perspective. The report acknowledges that the third factor is subjective. Specifically, in an effort to rule out other potential causes, the report indicates that such a determination was made by determining if the subject had a pre-existing condition that was identified in the adverse event report.	
It is stated, "With regard to adverse events, it was recognized by EPC staff and the TEP that, even in aggregate, the number of patients included in randomized trials was likely to be few…. Because of this, it was recognized that case reports would have to be relied upon to assess serious adverse events." It was Dr. Shekelle who	The causality analysis has been removed from this revision. We revised the sentence to indicate EPC staff recognized assessing case reports was going to be required to meet the terms of the contract. Our notes from the TEP meeting are clear

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
advanced his position to this question, but it is not correct to state that this was recognized by the TEP. In my remarks I made it clear that the clinical trials were the most useful clinical data available, and that the FDA AER database was to flawed and incomplete to be able to draw any conclusions. Dr. Shekelle introduces Cindy Mulrow's criteria as Rand's position in the assessment of adverse medical events. My point was that these criteria would be useful in assessing adverse events in the course of controlled, conventional, pharmaceutical trials, not the poor quality of the voluntary AERs in the FDA database was shocking in contrast. We did the best analysis possible in the circumstances for the CRN report and found that even the most complete subset was not sufficient quality to draw any conclusions. Dr. Benowitz agreed that the AERs did not have all the elements of an AER analysis. Therefore, the statement on page 21of the draft report is a reasonable summary of the discussion, i.e., "Consequently our TEP judged that case reports alone would be insufficient to establish definite causality…" Dr. Shekelle also agreed that the FDA AER could not be used to assess causality. He said that the adverse event issue would be the hardest to deal with, because it is front page and gets wide attention, but he was not worried about disagreeing with the FDA. Stating that not all the AERs were true or false, he acknowledged that the gold standard was lacking to link exposure and outcome, there is no basis for a conclusion.	that the majority of the TEP agreed.
The assessment of probability/ possibility respecting causality between use of ephedra-containing dietary supplements and adverse health events seems shaky based as noted mainly on FDA case reports, "insufficiency documented". It is not clear to me where the margin is between probably and possibly. and whether there is a clear basis for location on either side of the margin.	Causality has been removed from this revision.
It is not clear why you require coronary angiography for cases of myocardial infarction. Clearly MI can be diagnosed on the basis of EKG and enzyme changes. Coronary angiography does address the severity of underlying coronary artery disease, but that does not address whether or not ephedra played a causative role. It is well known that coronary spasm is most likely to occur at the site where there is some underlying coronary artery disease. If ephedra can cause coronary spasm, a person with underlying coronary artery disease would be most vulnerable to this occurring.	We clarified that coronary angiography was required in cases of myocardial infarction in order to evaluate other causes, such as coronary artery disease, not to make diagnosis. In the presence of coronary artery disease, the occurrence of an MI could be classified no higher than a possible sentinel event.
The criteria in our causality algorithm may be too conservative. Requiring coronary angiography in cases of M.I. would exclude cases diagnosed by cardiac enzymes and ECG changes alone.	Angiography was required to assess the possibility of alternative explanations, not the presence of a myocardial infarction. The text has been revised to reflect this.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
The interpretation of co-existing cardiovascular disease with respect to causation is an important issue. It is quite likely that underlying cardiovascular disease would predispose to ephedra causing a serious cardiovascular event. This has been well shown to be the case for cocaine. I think this issue needs to be made quite clear, especially since this is the reason why many of the adverse events are classified as possible in your evaluation, while they were judged to be probable in the evaluation done by Christine Haller and myself.	In this revision we no longer deal with causality.
What is meant by "more than the minimal dose?" How do you know what a "minimal dose" is? Were any cases excluded because of this criterion? In the same figure on level 3, the question comes up again about the difference between probably causal and possibly causal. The box above possibly causal says "interaction with ephedrine likely." If you say that the interaction is likely, then why do you say it's possibly causal?	We no longer use this criterion or assign causality in this revision.
"A 41-year-old female has four stroke events over a 2-month period between December and February." This case seemed to have incomplete description as there was no mention of the product (Diet Phen) that the patient was taking and when and for how long (14-60 days).	This text has been revised.
According to Table 20, Ripped Fuel was also involved but there is no mention of the product in the description (which is probably important information for the reader).	This change was made.
If the date of death is May 1994, she cannot be admitted to the Emergency Room in December of 1994. Should it be 1993?	This typo has been corrected.
I found that Table 20, column titled as "Key Determinants of Causality" rather confusing, incomplete and unclear. Delete "Timing<24 hours" from all cases as this does not contribute any additional information but rather add confusion to the interpretation (reader may interpret a "no" to Timing <24 hours means ingestion did not occur within 24 hours, which is not the case as many times tox screen is "yes".) Change "Tox screen**:" to "Tox screen was done:" and eliminate "**Ephedrine/amphetamines found in toxicology screen" as this is not true in all cases. Add "Ephedrine alkaloid detected: Yes or No" to the column since tox screen may not include detection of ephedrine or its alkaloid.	This table has been revised to improve clarity.
Does Nature's Nutrition Formula Three contain ephedrine (see p. 60 case description)? Should this be included in the table? In the description on p.60, it stated that "Toxicology screen was negative for ephedrine...", however, table 20 indicated that ephedrine was found The footnote may be	Change made.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
misleading and it may be more appropriate to footnote tox screen with"Tox screen was done" rather than "Ephedrine/amphetamines found in toxicology screen".	
The case description stated that "Toxicology screen for cocaine, ephedrine and amphetamines was negative." However, the Table indicated that ephedrine / amphetamines were found (same problem as #15).	Change made.
The age of the patient is different (37 y.o. or 36 y.o as described on page 62)?	Change made.
Should the product be Metabolife (as described on page 63) or Metab-O-Lite as indicated in the table.	Change made.
Does Accelerator also contain ephedrine? Should this be included in Table 20? Footnote of tox. screen is inconsistent with the description on p.63 (same problem as #15).	Change made.
The ages are different between the table and the description on page 64.What is the tox. screen results for this patient?	Change made.
The ages are different between the table and the description on page 64-66.	Change made.
There is no description of the case on page 66 under Stroke, possibly causal section but it is found under Subarachnoid Hemorrhage on page 68. Yet, it is grouped together with all the other cases of Stroke (CVA) in table 20. Should this be separately described or should this be described under Stroke section?	The grouping of the cases has been changed to better improve clarity.
For AER 13672 described as "probably not causal" on page 62, There are toxicology results that showed 280 ng/ml ephedrine in the blood. These results were reported by the Medical Examiner on 2/12/02, which may be after the FDA report was finalized.	
Metabolife recently admitted (after this draft was issued) that it has received 13,000 complaints about its ephedra products. These should be included in your analysis.	These are now included.
p 47, question 18: what was the rationale for dividing the durations of use into the listed classes? Because tachyphylaxis generally occurs after about 14 days of continuous use of ephedrine, the evaluation of acute use or for acute effects is commonly limited to days 1-14, with durations longer than 2 weeks considered "chronic" use.	Because we wanted to distinguish dosing within 24 hours, we divided the categories in the Table in this fashion.. In the actual data, we recorded the exact time.
Regarding adverse event adjudication -- you make your reasoning clear for exclusion of individual events due to insufficient information or downgrading attribution due to pre-existing conditions; however, the point could be made more clearly that this may tend to underestimate the number of serious adverse events	We have emphasized in the text that our methods of case report analysis are conservative and may underestimate the number of events.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
A concern is that in the case studies there was no analysis and no sufficient emphasis place on the evaluation of dosage amount, which in many cases appeared to be excessive.	In the majority of case reports the dose was not even reported, preventing any dose analysis of the case reports.
To demonstrate how additional data could affect the outcome of the present study, consider case 13672, which was designated "probably not causal". This reviewer has access to additional data on this particular case, specifically that the decedent did not have a toxicological examination that revealed a postmortem ephedrine blood concentration of 280 ng/mL and a pseudophedrine level of 100ng/mL. Given this additional information, is the causality level assigned to this specific case still valid? More thorough investigations like those pursued during the discovery process of specific lawsuits almost always yield additional information that would likely modify the causality assessment of specific cases submitted via the MEDWATCH program. Also, were those case reports described in the medical literature used in the case report assessment? They do not appear to have been utilized.	We did include the medical literature case reports in this revision. The additional information about a specific case, as provided by this reviewer for this case and other reviewers for other cases, we unfortunately cannot include or assess in our report, as we have no access to the original information.
I wasn't clear as to why only case reports documenting death, myocardial infarction, and/or cerebrovascular accidents were evaluated. To me this made the comparison to the Haller and Benowitz study less meaningful.	We have included additional case reports in this revision. We have deleted the comparison to the Haller and Benowitz study since we no longer assess causality.

Appendix 3. Reviewer Comments (continued)
Second Review of Safety Analysis including Metabolife Data

Reviewer Comment	Rand Response
Meta-analysis. This study was sent to me for information purposes only. The analysis is well done and meets the highest standard. The efficacy of Ephedra is very modest in terms of weight loss. Treatment causes significant adverse effects with RRs ranging from 2 to 3. Statistical power is inadequate to rule out severe adverse events occurring at a rate of 1.1/1000. The severity of the adverse effects can not be determined (my assumption). Likewise, the dose-response relationship can't be estimated.	No response
Metabolife analysis. The database is extremely messy and does not allow many meaningful analyses and conclusions. Your approach in terms of coding rules, data extraction and event classification is good. If a pharmaceutical company had kept records in this sloppy way my assumption is that it would be in deep trouble with the FDA. You conclusion is weak in my view. I would say that the "Findings are consistent with an increase in rare serious adverse events. What troubles me is that the population is so young. I would not expect serious cardiovascular events occurring so often. I realize that we don't have a denominator so any attempt at even guessing what the event rates might be are probably too speculative. In summary, you have from an analytic point of view done what can be done.	No response
Evidence Report. This is another well-done study. My interpretation is colored by two facts. When people complete a MedWatch form they suspect an association. There is a marked underreporting ranging from 90 to 99 %. This means that what appears to be rare may not be very rare. The temporal relationship between use of Ephedra and the occurrence of an event may exist even if it isn't documented. Again, I think your conclusion is too mild. I am fairly convinced that Ephedra causes serious events but I can't give a rate. Moreover the benefit-risk ratio is unfavorable (minor benefit for sure), so I would question the wisdom of leaving the compounds) on the market. My position is also influenced by the age of the victims.	We added to the limitations that MedWatch may underestimate the number of events.
The AERs presented in this report appear to be consistent with the known pharmacologic actions of ephedrine. Would it be appropriate to include a statement to this effect in the report?	We included such a statement.
When the summary text under Results and Conclusions is updated to include the analysis of the ephedrine AERs, care should be taken to present results for ephedra and ephedrine separately.	This was done.
For the case descriptions of the cerebrovascular accident/stroke events, it would be helpful to include the individual's functional status in the text (it is already included in the Table 20).	We included this information to the extent that we identified it in the source documents.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
We've seen a comment from NCCAM regarding the tables. We agree with that comment.	These tables are now incorporated into this revision.
Page 5, paragraph 3, last line: "Subject" should be defined. Page 7, paragraph 2: Second sentence: Text would read better as follows, "…seizure, only those cases described as generalized toxic-clonic seizures underwent further review." Sentence beginning at end of line 5: Text notes requirement that there be documentation that the individual had consumed ephedra or ephedrine within 24 hours of the event, but that this was not a requirement for psychiatric events. It would be helpful to explain why this decision was made. "Sentinel case" was not defined previously or used subsequently in the report. If it means "sentinel event", should change wording. Page 24, paragraph: Should "doses" be changed to "dosage?" Page 27 Paragraph 1, under FDA Cases Ephedra: The dates aren't correct. It appears as if the patient was taken to the hospital in December 1994 where she signed out AMA even though she had died in May 1994. What is "chlophoramine?" Paragraph 2, line 3: Change "toxicology" to "toxicology screen." Page 28, Paragraph 3 (case# 12722): Text doesn't mention ephedra exposure – what product was used? Paragraph 4 (case# 12843): Text doesn't mention ephedra exposure – what product was used? Paragraph 7 (case# 14638): Text notes that individual had been taking Hydroxycut for seven days, but Table 20 (page 52) says 2-13 days. Page 29 Paragraphs 2 and 3 (case# 44): Text doesn't mention ephedra exposure – what products were used? Paragraph 5: Case# 258 is not included in Table 20. Paragraph 6 (case# 13672): Change "rain" to "run." "Soldier" does not indicate gender. Although from the text the individual is apparently a male, wording should be changed. Should indicate whether the toxicology screen looked for ephedrine or that ephedrine was not mentioned in the report. Paragraph 7 (case# 1859087): Text says this individual was taking Max Alert, but Table 20 says the product is unknown Page 30: Paragraphs 5 and 6 (case# 13806 and case# 14465) are not included in Table 20. Paragraph 6 (case# 14465), last line: Change "not conclude anymore" to "come to no other conclusions." Page 31: Paragraph 1: Product names in text and Table 20 do not match. Paragraph 2: Product names in text and Table 20 do not match. Indicate whether the toxicology screen looked for ephedrine or that ephedrine was not mentioned in the report. Paragraph 6: Text notes that this individual was taking E'ola Amp II Pro drops for 12 days but Table 20 indicates that he took them for 2-13 days.	We corrected typographical errors. We made suggested changes in language. We stated whether ephedrine was looked for in the toxicology screen. We made the product names match, e.g. "Ripped Fuel (Twin Labs)" was changed to "Ripped Fuel." The reviewer is incorrect about cases being in the text but not in Table 20 (now table 22); all were present.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
Page 32: Paragraph 2 (case# 9504): Product names in text and Table 20 do not match. Paragraph 3 (case# 10009): Is (was) there such a product as "Metabolift?" Paragraph 4 (case# 13009): Text indicates that the individual described is a male, but Table 20 notes that it was a female. Indicate whether the toxicology screen looked for ephedrine or that ephedrine was not mentioned in the report. Paragraphs 5 and 6 (case# 14114 and case# 14530): Product names in text and Table 20 do not match. After case# 14530, the last one in the text under "Myocardial Infarction", Table 20 continues with many more case reports of MIs (pages 64-66) and then lists "other cardiac " (pages 67-71) starting with three "possible sentinel events." Why aren't the descriptions in the same order in both text and table? Page 33: Paragraph 1: Indicate whether the toxicology screen looked for ephedrine or that ephedrine was not mentioned in the report. Paragraph 2, line 1 (case# 11062): Paragraph 2: Text indicates individual was 44 years old, but Table 20 says she was 42. Insert "was" between "and" and "a" in "was taking Power Trim and a cigarette smoker." Paragraph 4: Indicate whether the toxicology screen looked for ephedrine or that ephedrine was not mentioned in the report. Page 34: Paragraph 1: Product names in text and Table 20 do not match. Paragraph 5: Sedimentation is misspelled. Page 35: Paragraph 1: Product names in text and Table 20 do not match. Line 2: editorial - change "here" to "her". Line 7: editorial – change "with embolus" to "with an embolus" Paragraphs 3 and 4 (case# 10094 and case# 12713): Product name in text and Table 20 do not match. Paragraph 6 (case# 515): Text doesn't mention ephedra exposure – what product was used? Page 36: Paragraph 2: Text indicates individual is 25 years old while Table 20 indicates she is 26. Paragraphs 2, 3, and 4 (case# 14378, case# 14434, and case# 14553): Product names in text and Table 20 do not match. Paragraph 5: Thoracic is misspelled Page 37: Paragraph 2: Product name in text and Table 20 do not match. Paragraph 3: Delete either "other" or "additional." Paragraphs 3 and 4 (case# 13829 and case# 13905): These cases are not listed in the same order in the text and in Table 20 making them difficult to find (they are located on page 78). Paragraph 4 (case# 13905): Text notes individual is a female of unknown age while Table 20 indicates she is 36 years old. Paragraph 6: Text notes that individual was taking 40-60 mg of ephedrine for 10 years. Was this 40-60 mg per day? Some indication of amount per unit time should be provided, or a note should be made that the information is not available. Page 38, paragraphs 1, 2, 3, and 4 (case # 12851, case# 13031, case# 13643, and case# 13793): These cases are	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
not included in Table 20. Page 39: Paragraphs 1, 2, 3, and 4 (case# 110, case# 297, case# 260, and case# 13945): These cases are not included in Table 20. Paragraph 1 (case# 110): This case was classified as a "possible sentinel event" but is listed in the middle of a list of cases with "insufficient information." Paragraph 2 (case# 297): Change "taken" to "taking." Should indicate whether or not there was any information on how long he had been taking Herbalife supplements? This case was classified as a "possible sentinel event" but is listed in the middle of a list of cases with "insufficient information." Paragraph 3, last line (case# 260): Delete "intake." Paragraph 5 (case# 13062): Product name, duration, and dose are not included in the text, but are given in Table 20 and should be included here. This case was classified as a "possible sentinel event" but is listed in the middle of a list of cases with insufficient information. Page 40: Paragraph 3: "5am in the morning" is redundant. Paragraphs 3 and 4: Product names in text and in Table 20 do not match. Page 41: Paragraph 2 (case# 10432): Product name in text and Table 20 do not match. What is "encepholophy?" Change "focality" to "focal nature." Paragraph 3 (case# 11062): Product name in text and table do not match. Change "taking" to "taken." The last line notes that because of the possible structural abnormality, this event was classified as a possible sentinel event. However, it is not clear from the text what the possible structural abnormality was. Paragraph 4 (case# 11649): Product name in text and Table 20 do not match. Indicate whether or not there was information regarding how long this individual had been taking Metabolife prior to the event. Page 42: Paragraph 1 (case# 13408): Product name in text and Table 20 do not match. Paragraphs 1 and 4 (case# 10874 and case# 11675): Indicate whether the toxicology screens looked for ephedrine or that ephedrine was not mentioned in the report. Paragraph 2: Text indicates individual in case# 14275 was 38 years old, but Table 20 says she was 30. Paragraph 3: Text indicates individual in case# 11105 was 31 years old, but Table 20 says she was 30. Paragraphs 3 and 4 (case# 11105 and case# 11675): Product names in text and Table 20 do not match. Pages 43 Paragraph 2 (case# 9747): This case is not included in Table 20. Paragraph 3 (case# 9509): Product name in text and in Table 20 do not match. Paragraph 5, last sentence (case# 13809): Text indicates the individual described was intent on doing harm to others, but Table 20 describes her as suicidal. Suggest changing "alleviated" to "subsided." Page 44: Paragraph 2 (case# 1855921): Text notes that this individual was taking Minithin but this information is not	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
included in Table 20. Text notes that there was no history of other drug use, but Table 20 says this individual had a history of substance abuse. Paragraphs 3 and 4 (case# 48 and case# 238): Change "psyc" to "psychiatric." Paragraph 4 (case# 238): Provide composition of Tedral as is done for Bronchipax in paragraph 5. Also note that this drug is listed as "Bronchi Pax" in Table 20. Paragraph 6 (case# 9751): Product name in text and in Table 20 do not match. Should note in Table 20 that problem resulted from discontinuation of product use as described in the text. Page 45: Paragraph 3 (case# 12372): Product name in text and Table 20 do not match. Last line: Change "classified as" to "classified it as." Paragraph 4, first line Case# 13005): Change "also" to "used." Paragraph 4 (case# 14436): Delete "(tid)." Paragraphs 5, 6 and 8 (case# 14436, case# 14528, and case# 79): Product names in text and in Table 20 do not match. Paragraph 6 (case# 14528): Clarify "very soon" vs. "approximately 1 week after." Paragraph 7 (case# 1682426): End of line 3: Change "in residential" to "in a residential." Last line: "Note" is misspelled. Page 46: Paragraph 1 (case# 79): Text notes that product name is not given, but that investigators contacted the manufacturer – is the name of the manufacturer known? Paragraph 2: Should term "causality" be used here after the discussion about not trying to determine causality on page 6? Page 63, Table 20, row 6, and Page 73, Table 20, rows 5 and 6: Care should be taken to provide full product name. E'ola is the manufacturer name and E'ola makes some diet products that are laxative-based which would be inappropriate for inclusion in this report.	
Page 72 and elsewhere in Table 20: What does "implicit review" mean?	This was defined in the report.
Page 91 and elsewhere: Replace "psyc" with "psychiatric."	This change was made.
Page 103: Paragraph 1: Use of term "causality" should be reconciled with discussion on page 6 regarding the intent of the report. Last two sentences would be more accurate if changed to: "Definite causality for adverse events cannot be determined from case reports. When an adverse event is very serious it may be infeasible or unethical to conduct a de-challenge/re-challenge test for causality."	Causality has been removed from this revision.
Given the short time frame RAND has to fulfill its contractual obligations, the peer review process also has necessarily been severely time-constrained, not to mention coincidental with the year-end holiday period. This is regrettable but I have had the opportunity to review these drafts and reflect upon what they say in general and how it is said. I have not had sufficient time to review the details of the reports and tables for accuracy, which is almost certainly true for the	We indicate in the report that this section did not receive the same level of peer review as the other portion of the report.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
other peer reviewers. Any claim that the Metabolife Report in particular has been "peer-reviewed" should be qualified, as a true peer review was not possible under the circumstances.	
The Metabolife Report spends a large proportion of its content describing the poor quality of the reports, limitations of the records and methods, the short time frame, the many thousands of files, etc., making it difficult to see a scientific value to this review under these circumstances.	We were requested to do this review as part of the contract.
A major criticism of the Metabolife Report is the disconnect between the final limitation listed on p7, which clearly states that ". . . case reports are in general not considered sufficient evidence to draw conclusions about causality", and the overall tone of the report. Any reader will be led to believe that this report links the occurrence of adverse events to the consumption of ephedra, despite the limitations listed at the end. Even though the words "possible" and "may" do appear in the report, the terms and phrases used in the methodology, the detailed and repetitious descriptions of the case reports, and the results, are all written with such a factual tone that there can be no doubt that this report will be interpreted to mean that these events were caused by ephedra. The report should be rewritten to state the study's major limitations (p7) at the beginning, i.e., that these case reports cannot be ". . . considered sufficient evidence to draw conclusions about causality". Then the report should state at the outset in clear terms that the purpose of the analysis was not to establish or prove that there is a risk of serious adverse events, since AERs are not suitable to that task. The purpose was to determine whether or not this database might be useful to "generate a" (rather than "support the") "hypothesis that ephedra may cause rare serious adverse events", to quote conclusions on p7. The introduction should also state the fact, which is not a conclusion of the study but was included as the final statement of conclusions (p7), that "A hypothesis-testing study, such as a case-control study, is necessary to prove or disprove this hypothesis". Each statement in all sections should be carefully examined and rewritten if found to be interpretable as drawing a link between ephedra and effects. Phrases such as "instances of serious adverse events such as death, heart attack, or stroke" are repeated several times which undoubtedly will lead to the impression that these are caused by ephedra. It should be made more clear that the Metabolife Report, as well as the review of the FDA AERs and published case reports, are part of the effort to explore the hypothesis that ephedra may cause rare serious adverse events. Common sense and a rudimentary understanding of pharmacology will lead to the conclusion that this clearly is possible, depending principally upon the dose. Anything at a	RAND did not generate the hypothesis that ephedra causes serious adverse events, that hypothesis was already generated and one reason why our report was commissioned. Certainly the existence of serious adverse events in otherwise healthy young adults must be considered “support” for this hypothesis, just as the lack of such events would be considered a lack of support. It is not proof of a causal relationship, and we say so, repeatedly. We also note that the concern about out report being overly suggestive that the case reports imply a cause and effect relationship is not shared by numerous other reviewers, who believe just the opposite, that our report downplays the possibility of a causal relationship.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
sufficiently high exposure can cause adverse events. Continued analysis of deficient and flawed adverse event databases cannot and will not lead to any conclusions about causality, but hurried evaluation, suggestive language, and imprecise wording can lead to perceptions of cause and effect that are not scientifically supported. Turning to the revised RAND Report, it is difficult to comment on the "Safety Assessment Excerpt" without seeing and understanding how it is used and referenced in the rest of the report. Most of the comments that I made previously still apply to this draft because 99% of the safety assessment deals with adverse event reports which are flawed and inconclusive. Some limited peer review of the introduction and conclusion sections of the completed report should be permitted to assure that the wording of these sections avoids the continuing problem of implications that a cause and effect relationship can be established from the number of AERs, or the exhaustive treatment given to the AERs, in the report. The language in the previous draft has been changed to reflect the fact that the "causality scale" leads to erroneous and exaggerated conclusions. Nevertheless, the new scheme of classification, using the terminology "sentinel" instead of causal, is still suggestive that these reports can be used for interpreting cause and effect. To help avoid this problem, there should be added to the explanation of the term "sentinel" on p7 of the Metabolife Report and on p6 of the Safety Assessment Excerpt that adverse events, even serious events, are commonly idiopathic in etiology, and that therefore the lack of any known cause combined with known consumption of ephedra is not meant to imply that ephedra was the cause -- the intent is simply to show which events could potentially have been caused by ephedra, given the lack of a known cause, with the understanding that a cause and effect relationship for ephedra cannot be established from such reports. I have not been able to review in any detail the descriptions of events categorized in the Metabolife Report or in the Safety Assessment Excerpt. Nonetheless, a cursory review indicates that the criteria established for "sentinel" events in particular has not been met in a significant number of these cases, and these reports should be reviewed with this in mind. For example., RAND did not have access to the results of the autopsy in the first death listed as a sentinel event on p26, and availability of autopsy results is appropriately listed as a criterion for qualification as a sentinel event on p7. Also, there are a number of cases described as sentinel events where the individuals are also described as long-term smokers, alcoholics, or drug abusers. These and other cases do not appear to meet the criterion that sentinel events be idiopathic when these conditions are known to be risk factors for events at issue.	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
The revised report on p7 states that the Office of Dietary Supplements had given RAND a key question concerning "the relationship between dose and the likelihood of serious adverse events". The authors, however, ". . . do not believe such an analysis is justifiable on the case report evidence. . . ." It seems to me, if evidence is insufficient to be evaluated for dose relationship, then any such evaluation is unjustifiable, which speaks to the point I made in my previous review that reliance on flawed case reports can only lead to flawed analysis and conclusions. Similar to previous comments on the Metabolife report above, the limitations and lack of ability to draw conclusions from AERs should be stated clearly up front in the Safety Assessment Excerpt. The language used to describe the large number of reports clearly suggests causality, even if not intended. The preponderance of the description in the safety section leads the reader to conclude early on that ephedra must be responsible for these effects. The very brief description of controlled trials is dismissive of strong evidence for ephedra safety, and the extensive toxicology database is completely ignored. Therefore, the safety section continues to be unbalanced by the absence of objective evidence in contrast to the voluminous treatment given to the case reports. I agree with the statements in the revised report (p6) concerning the variability and subjectivity of interpretation of case reports. This is a principal reason for my objection to their consideration being the centerpiece of the report's safety assessment. It is an important exercise, and RAND has done as thorough a review as could be expected. It is extremely important, therefore, that readers of this report not be led to an impression that the repetitive descriptions of large numbers of case reports can be interpreted as evidence for cause and effect. Clearly this is and will be the message unless the introduction, methods and language throughout are consistent with the messages about limitations, insufficiency for causality, and the need for a conclusive study of a different kind, i.e., a case-controlled study to add to the existing objective clinical evidence. As a final point in this regard, the "Conclusions" section on p103 of the Safety Assessment Excerpt should be revised to remove the implication that RAND has concluded that the case reports are useful to establish causation or that the case reports establish that there is in fact a risk of serious adverse events. The reports generate a hypothesis, and whether a risk of serious events exists as well as the estimate of the level of any risk needs to be determined through scientific studies, not review of additional case reports. In particular, the sentence beginning with "For rare outcomes" in the first paragraph should be revised to make	

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
it clear that the review of case reports was to assess whether these reports generate a hypothesis that ephedra might cause rare outcomes.	
In addition, the third bullet should be revised to avoid the implication that the lack of other identifiable causes combined with ephedra consumption establishes causation.	We added this important qualifier.
RAND's conclusion that further analysis of case reports is pointless is the key to moving forward with a scientific evaluation of ephedra and the resolution of a controversy that has been created by over-focusing on case reports. This point should, therefore, be made in clear terms at the very beginning of the completed Ephedra Report.	The “Conclusions” is the appropriate place for this conclusion.
We would like to see summary tables of the sentinel and possible sentinel events by ephedra use, by ephedrine use; by gender; by broad age groups; by category of AER. The long tables listing each event are not sufficient.	These tables are now added.
As to the adverse consequence conclusions it would seem appropriate to summarize the events for ephedrine as they are 'bulleted' for ephedra. Right now, it looks as if there is no conclusion on the sentinel and possible such events for ephedrine.	This change was made.
A recommendation is made for scientific studies of ephedra risk. No comment is made on whether it would be appropriate to also do this for ephedrine. For the present data, one could argue that PPA-like case-control studies should be generated for other ephedra and ephedrine products.	These changes were made.
Adverse Event Data from Randomized Trials. Methods. RAND identified 44 randomized, controlled studies, and a pooled meta-analysis was conducted of the risk of adverse events in treated vs. placebo groups for the most commonly reported adverse events. Risk was significantly elevated for psychiatric symptoms (OR 3.24, 95% CI 1.67-6.58), autonomic hyperactivity (OR 2.91, 95% CI 1.84-.70), nausea and vomiting (OR 2.37, 95% CI 1.51-3.78), and palpitations (OR 2.11, 95% CI 1.16-4.02). The risk was elevated, but not significantly, for hypertension (OR 1.86, 95% CI 0.39-11.74). Tachycardia was reported in only one study. The methods used are standard; the analyses appear appropriate.	No response
Adverse Event Data from Randomized Trials. Methods. The subgroup analyses of adverse events of ephedrine (+) caffeine were said to be “similar to the main analysis”. This data may be important and should be presented in greater detail. The reason for this is that caffeine can potentiate the CNS stimulant effects of this class of drugs (sympathomimetic aminies).	The results are the same because the ephedrine plus caffeine studies contribute the vast majority of the data to this analysis. So we have said as much as we can about this.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
Adverse Event Data from Randomized Trials. Methods. A second recommendation is to conduct the pooled analysis combining similar adverse event groups in an attempt to reanalyze for a dose-response effect from ephedra or ephedrine (+) caffeine. For example, it would make sense to combine palpitations, tachycardia, and hypertension in such an analysis, since all are cardiovascular, sympathomimetic events.	We did this analysis and included it in the results.
Adverse Event Data from Randomized Trials. Methods. I would also like to see a pooled analysis of headache, and add this to Table 17. The reason for this is that headache may be a prodrome to more serious neurologic events, and was present in all three cases I reported at the 1996 meeting of the American Academy of Neurology.	This adverse event analysis was added.
Adverse Event Data from Randomized Trials. Potential for Bias. Since this data is a meta-analysis, there is little opportunity for bias.	No response
Adverse Event Data from Randomized Trials .Clarity of Reporting. The writing is clear. The report (I) would flow better if the meta-analysis section was stand-alone and separate from the case report analysis.	It is separate in the final version of the report.
Adverse Event Data from Randomized Trials. Conclusions. The conclusions are to the point. However, I find the meta-analysis conclusion somewhat lacking in methodologic content and discussion. To be more useful, expansion of the author's critical point in ¶1, page 25, should be added to the Conclusion section (p.103). The conclusions would more properly read: 1."There is sufficient evidence"…. (same). 2. Safety data from relatively small clinical trials of ephedra/ephedrine are unlikely to reveal rare but serious adverse events, those that may occur at a rate of less than 1/1000. Thus, such data cannot be used to conclude that ephedra/ephedrine does not cause such serious adverse events. In addition, it is likely that, in some of these trials, differential drop out of treated patients related to a higher rate of milder adverse events could have removed subjects at higher risk for more serious events.	We reworded this to try and improve clarity.
Adverse Even Data from Reported Cases. Methods. It is important to point out that the authors utilized a very much more conservative method to identify the likelihood of association with ephedra/ephedrine than that reported by both Haller and Benowitz5 for cardiovascular and central nervous system events, and by Samenuk et al for cardiovascular events. The authors should point out the differences with these studies, and how these differences may have led to different counts of adverse events in the main categories in Reports I and II. A table highlighting the differences with Haller and Benowitz would allow clearer comparison of categories. One important difference is that	Since we dropped a "causality" assessment from this revision, we don't think such a comparison is valid. We do acknowledge in the limitations that our methods are more conservative than those used by some other groups.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
RAND dropped all cases with any alternative explanation or competing cause to "probably not related". Haller and Benowitz, however, considered events at least possibly related, even in the face of co-existing or pre-existing condition, if those conditions themselves could be severely exacerbated by ephedra alkaloids (e.g., hypertension, some psychiatric conditions).	
Adverse Even Data from Reported Cases. Methods. One table should summarize all of the reviewed cases by adverse event type (e.g., death, seizure) and by author's conclusions regarding category (sentinel, possible sentinel, probably not related, and insufficient information). For example, of 41 reported seizure cases, only two were deemed "sentinel" cases. This may highlight the insufficiency of available data with which one may judge likelihood of association. For example, I have detailed knowledge of seizure case 13408. Even with the author's criteria, this case should be classified as "sentinel".	These tables are now included. We acknowledge that limitations of the source documents limit our ability to draw conclusions.
Adverse Even Data from Reported Cases. Potential for Bias. With the conservative approach described under "Methods", there is potential for serious misclassification of cases, primarily in the direction of "probably not related" or "insufficient data". It is much less likely that misclassification substantially went in the other direction.	This limitation was acknowledged in the appropriate section.
Adverse Even Data from Reported Cases. Clarity of Reporting. As mentioned above, at least one or two other summary tables would be helpful to the reader.	These tables are now included.
Adverse Even Data from Reported Cases. Conclusions. The conclusions reached on p.103 regarding the case report assessment are not very helpful in moving things forward on this issue. There is an underlying assumption that, if causality cannot be proven from passively reported cases with poor documentation, then it may take a case control study to do so. There are several problems with these conclusions: (1) If ephedra were an FDA-approved drug, its use would have likely been banned related only to the sheer number of serious adverse events reported. Even if one accepts only the "sentinel" and "possible sentinel" events reported here, or those reported by Haller and Benowitz, or the cases from Texas or Rochester6, the likelihood of association, to most clinicians, would be overwhelming. (2) There should be substantial discussion added to the report related to other converging lines of evidence one would normally wish to include in an assessment of causal relations. These would include: (a) Expected actions of sympathomimetic amines, including effects on the peripheral vascular system, and the biologic plausibility of association with milder and severe adverse events. (b) A summary of the extensive literature on the potential for the "look alike" drugs such as PPA to cause similar serious	We note there is a great deal of controversy among experts about whether case reports are sufficient to conclude cause and effect relationship with serious adverse events. The other kinds of evidence cited by this reviewer were outside our scope. We do think a case control study is possible, and that the controversy is likely to continue to rage until such a hypothesis-testing study is performed.

Appendix 3. Reviewer Comments (continued)

Reviewer Comment	Rand Response
adverse events. For example, both PPA and ephedrine are known to be associated with angiitis. (c) Evidence from animal studies or basic neuroscience studies related to adverse events of ephedrine. (3) If one of the problems relates to poor reporting to the FDA or from the manufacturers, it would seem that, at a minimum, clearer reporting standards should be established. (4) It would be extremely difficult to conduct the type of case control study recommended. The serious events are rare, and among the major event categories (e.g., seizures), ephedra is not likely a frequent cause. I have thought about how to conduct such a study, either via emergency departments or poison control centers. However, there would be serious methodologic issues in proper case and control specification. Can we really afford to wait for such an imperfect study to be conducted? Is there really any justification whatsoever not to ban unfettered use and marketing of these sympathomimetic amines in pharmacologic doses?	