Appendix A

Appendix A

Epidemiology Expert Panel Meeting Report
Review of Yale Hemorrhagic Stroke Project
April 19, 2000

Curricula Vitae of Panel Members

Philip Gorelick, MD, MPH., FACP, Professor of Neuorological Sciences, Rush Medical College

Lewis Kuller, MD, DrPH, MPH, Chairman, Department of Epidemiology, University of Pittsburgh School of Health

Robert Wallace, MD, Chairman of Preventive Medicine, University of Iowa

Noel Weiss, MD, DrPH. (Panel Chair) Chairman of Epidemiology, University of Washington-Seattle

Epidemiology Expert Panel Meeting Report
Review of Yale Hemorrhagic Stroke Project

OVERVIEW:

The Consumer Healthcare Products Association (CHPA) requested that a panel of epidemiologic experts meet to give their opinion on the results of an epidemiology study, The Hemorrhagic Stroke Project (HSP), conducted to determine the relative risk of having a hemorrhagic stroke event coincident with taking phenylpropanolamine either as a cough/cold medication or as an appetite suppressant. CHPA is the trade association that represents the nonprescription drug industry. This panel was convened under the express condition that it would be independent from CHPA and the pharmaceutical industry and be free to express its opinions and conclusions.

The members of the panel represented expertise in the design and conduct of case-control studies involving cardio- and cerebro-vascular diseases, neurology and cardiology. The panel consisted of:

Dr. Philip Goerelick, MD, MPH, FACP (Rush Medical College)
Dr. Lewis Kuller, MD, DrPh, MPH (University of Pittsburgh)
Dr. Robert Wallace, M.D. (University of Iowa)
Dr. Noel Weiss, M.D., Dr. P.H. (University of Washington, Chair of Panel)

Prior to attending the panel discussion, we were provided with comprehensive materials related to the design, conduct, analysis and interpretation of the HSP. These materials included the protocol, interview manual, interim data reports, draft HSP study report, case summaries, and the appendix to the letter sent by CHPA to the investigators in response to their request to industry for comment on the draft report. Also provided was the "Points-to-Consider" document prepared by CHPA epidemiologic and statistical consultants.

Our objective was to discuss the results of the HSP and to present an objective report on our interpretation of the results from this study. It should be noted that the CHPA sponsored the panel in the interest of providing independent expert advice to the manufacturers and distributors of phenylpropanolamine (PPA) containing products.

Although not every member of the panel was in full agreement on every issue, the deliberations are summarized in the attached Appendix. Overall, based on the analyses of the data that were available to us, we did agree that:

- This study had several methodological issues that could have confounded the results.
- Hemorrhagic stroke was a rare event among users of PPA.

- The results of this study, by themselves, are not sufficiently compelling to drive a public health decision regarding reported PPA use and the subsequent development of hemorrhagic stroke.

PANEL DELIBERATIONS:

The following seven questions related to the design, conduct, analysis and interpretation of the study were the focus of the panel deliberations:

1. What is the likelihood that uncontrolled or uncontrollable confounding is a plausible explanation for the study findings?
2. What is the likelihood that uncontrolled or uncontrollable bias is a plausible explanation for the study findings?
3. What is the likelihood of the study findings being affected by information bias?
4. What is the likelihood that chance is a plausible explanation for the study findings?
5. Were the analyses conducted appropriately?
6. Does the study demonstrate a valid statistical association between PPA and hemorrhagic stroke?
7. Are there other aspects that require consideration in evaluating the study report?

(The Appendix provides detailed comments relative to these questions.)

DISCUSSION:

We recognize the difficulty and complexity in carrying out studies of this type and agree that the investigators used best efforts in the conduct of the study. Nonetheless, numerous methodological issues and concerns limit the interpretability of the study. Of concern to us were the marked differences in characteristics between cases and controls. The fact that the small number of exposed cases limited the ability to statistically control for these variables in this study greatly increased the possibility that chance, bias and confounding remain plausible alternative explanations for any apparent association between PPA use and hemorrhagic stroke.

Importantly, the findings demonstrate that, even if real, the population risk associated with the use of PPA and hemorrhagic stroke would be exceedingly small. One might even question the clinical implications of such relative risk values even if they were from a randomized, prospective study. We all agree that the small number of cases precluded adequate controlling in the statistical analysis for known confounding factors. We also have concern that since the overall finding for the primary hypothesis in the study- any PPA exposure- was null, selective emphasis on particular subgroups with smaller numbers might well be misleading

While one cannot eliminate the possibility that the HSP provides a signal, as a stand-alone study, these data are not sufficiently informative to draw any definitive conclusions. It is quite possible that all of the effect could be attributed to confounding and selective emphasis on particular subgroups. Therefore, any presentation of the results should include a detailed discussion of the possible role of confounding, bias, and chance as plausible alternative explanations of the findings.

CONCLUSIONS:

We emphasize that this study represents a significant undertaking and the investigators made strong efforts to control for many variables. Importantly, there were very few cases of hemorrhagic stroke in PPA users. The small number of cases in conjunction with the large number of potential confounders makes a robust statistical analysis impossible to accomplish. A single, case-control study with results of this type, can, at best, provide a signal of an association. Nonetheless, an alternative conclusion of no association is plausible as well. Although this panel is not qualified to render a public health decision, given that we have not reviewed the entire safety database on PPA, we believe that this study, by itself, does not suggest that use of PPA is creating an imminent public health concern. It could at best be used as only supportive evidence if there are other scientifically valid confirmatory data available. In addition to the ambiguous epidemiological data relating PPA and hemorrhagic stroke, the HSP report offered no plausible pharmacological mechanism that might underlie a causal relationship. We remain interested in assisting the investigators, sponsors, or FDA with the review and interpretation of this study, if requested.

Signed:

Philip B. Goerelick, MD, MPH, FACP

Robert Wallace, M.D.

Lewis Kuller, MD, DrPh, MPH

Noel Weiss, M.D., Dr. P.H.
(Chairman)

Attachment

APPENDIX I:

Some points to consider relative to the study design, execution and interpretation are summarized below:

I. Rationale for HSP

A. Signal Strength for hypothesis generation

1. The anecdotal case reports that preceded the design of the study should not have biased the design or execution of the study. The decision to use one-sided confidence intervals based on an expectation of risk was not warranted.

II. Methodology Issues

A. Identification of cases and matched controls

1. Cases were enrolled from hospital networks including tertiary care centers, whereas controls were selected by random digit dialing (RDD), which might account for several observed differences between the two populations, including PPA exposure and socio-economic status (SES).

2. The general method of case surveillance employed was reasonable. However, there were a significant number of strokes that could not be studied because of morbidity and mortality.

B. Participation Rates

1. Large differences exist in participation rates between cases and controls. This is a potential bias that was not accounted for in any way in the study report.

2. A large number of potential cases died or for other reasons could not participate. Only 61% of the identified case population were considered eligible and, of these, only 77% were actually enrolled.

3. Control response rate of 30% raises questions about validity and may produce more disparity between cases and controls. A question is whether the RDD procedure was flawed, as 150 needed to be called to get one control (normally expect 25 per enrolled control).

C. Comparability of Cases and Controls

1. RDD matching of controls was ineffective in controlling for SES.

2. Cases differed from controls in race, SES, caffeine exposure, history of hypertension, family history, as well as alcohol, nicotine and caffeine consumption. Inadequate or inappropriate control for these confounders could easily explain any positive association with PPA use. It needs to be emphasized that the small number of cases simply does not allow for appropriately controlling for these variables.

3. SES differences may explain differences in who gets the disease as well as who uses certain products. Particular concern was raised with respect to educational differences that might result in residual confounding sufficient to invalidate the analysis. In other words, we question whether this was truly a population based study.

4. There is some question regarding the geographical diversity of the cases. It would be helpful if the location of the cases by site be identified in the final report to determine if there was heterogeneity by site.

5. Controlling for body mass index (BMI) differences was not adequately addressed in the statistical analysis. While BMI appears to be similar between cases and controls, there may be larger differences in patients with aneurysms and intracerebral bleeds.

6. Heterogeneity in cases may make interpretation more difficult as the risk factors for aneurysm may be different between arterio-venous malformations and intracerebral hemorrhage.

D. Recall Bias/Interview Quality

1. Exposure estimation by self-report is subject to limitations. Cases were asked about drug use immediately prior to a catastrophic event, whereas controls were asked about drug use prior to an arbitrarily chosen day some days beforehand. The fact that cases have had a catastrophic event may bias them towards a greater awareness of previous product use. Controls not only did not have such an event to trigger their recollection, but they also appear to have had different recall periods.

2. Compared to the hospital cases, the non-hospital setting in which controls were interviewed may have influenced their response.

3. A large number of cases (44%) demonstrated some degree of aphasia, possibly limiting their validity and reliability. It appears that the differences in interview quality between the cases and controls could have been substantial. There is also some question regarding the number and quality of assisted interviews. It would be useful to perform an analysis including only subjects who the interviewer considered reliable.

E. Misclassification

1. Because of the small number of PPA users, even a modest degree of misclassification of product use by cases or controls could dramatically alter the findings.

2. The existence of numerous branded and generic products containing PPA could lead to confusion. Furthermore, many of the branded products, while carrying the same trade name, may or may not contain PPA. We all agree that the investigators appeared to have done their best to avoid this confusion; but nevertheless errors could have occurred.

F. Stroke Subtypes

1. Arterio-venous malformations (AVMs) and pressure-related cerebrovascular anomalies are different diseases. Combining them in the analyses may over emphasize a risk.

2. AVMs and intacerebral bleeds should be analyzed separately.

3. 2/3 subarachnoid hemorrhage (SAH) vs. 1/3 intraparenchymal hemorrhage (IPH) distribution in the cases is opposite to the case report experience or the SAH/IPH distribution in the general population (18-49 age group) from various health databases. This finding is difficult to interpret and again brings up the issue of this study being truly population based.

4. Three out of the six appetite suppressant cases had underlying aneurysms. This was not adequately addressed in the study report.

5. Six of eight cases in the "first use" analysis represented subarachnoid hemorrhage leaving only two cases classified as intracerebral. Two cases are truly insufficient to address an effect of PPA in this condition.

G. Prodromal Symptoms

1. Headache should be examined as a potential confounder since all subarachnoid cases were preceded by a headache, whereas controls had no prior headache. It is possible that headache could have contributed to the use of PPA-containing products.

2. Exclusion of cases with sentinel symptoms and alternate index dates changes the outcome events from 8 vs. 6 to 5 vs. 4. The study report should thoroughly discuss the association between sentinel symptoms and product use.

3. Seasonality of cases should be examined as cold/allergy symptoms and associated coughing could be an independent risk factor.

H. Other Drug Use

1. Self-reported cocaine use may be underestimated. Multiple drug use should be examined. Excessive alcohol and illicit drug use are likely to occur concomitantly and to be associated with lower SES and less geographical diversity. As such, geographical representation of cases should be further explored with respect to alcohol and illicit drug use.

2. Caffeine was significantly more prevalent among cases but not controlled for in the analysis.

3. Controls are more likely to use NSAIDs and other non-PPA stimulants than cases. Little consideration has been paid to additional ingredients in cold/allergy products as well as concomitant use of other products. A discussion of the possible role of other drugs in either a protective or detrimental role should be discussed (eg., NSAID effects on coagulation).

4. A higher proportion of controls took PPA-containing products during the 3- to14-day period than cases.

5. All other drug use during the one-day window should be evaluated.

III. Statistical Data Handling

A. Unusual findings with respect to adjusting for confounders

1. In several analyses, the strength of association between PPA and stroke increases when confounders are controlled for. One would expect just the opposite.

B. Residual confounding

1. Confounders could not be adequately controlled for in the analysis because of the small numbers.

2. Chi-squared analyses should be presented by level of confounder to provide a statistically appropriate indicator of the level of such adjustments.

C. Exact rather than asymptotic methods of analysis would have been appropriate.

1. Numbers are too small for asymptotic methods to be used for appetite suppressants.

2. Asymptotic methods were used to analyze data when appropriate methods failed to yield interpretable results.

D. One-sided confidence intervals are not appropriate.

E. The possibility of confounding being responsible for the observed association cannot be eliminated.

F. Association is only observed in subgroup analyses. It is misleading to overemphasize the extremes, particularly when they are inconsistent. For example, the "matched" odds ratio of 3.14 for the "first use" subgroup analysis (based on only eight cases and five controls) involved use of PPA for cough/cold only (and had a two-sided p-value of 2 x 0.029=0.06). By contrast the "any use" subgroup analysis found no consistent association with use of cough/cold PPA (odds ratio 1.23, 1p=0.245), and the apparently extreme ratio of 15.96 was for appetite suppressants, based on 6 cases vs. 1 control.

G. Emphasis on subgroups by time may similarly be misleading. "Current use" on index or prior day (21 cases vs. 21 controls) had a matched odds ratio of 1.61 (1p=.078; 2p=0.16), use on day 2 or 3 had an estimated odds ratio of 1.0 (6 cases vs. 12 controls), and use on days 4 to 14 had a crude odds ratio of 0.67 (11 cases vs. 33 controls). These numbers and matched odds ratios should be given explicitly in the tables (not just available by subtraction and footnotes) and are compatible with differential recall.

IV. Interpretation

A. Overall risks are not significantly elevated. Increased risks are only observed in subset analyses that are limited by small numbers, and not clearly significant when allowance is made for multiple comparisons.

1. The apparent finding that PPA use is protective if not taken within the one-day window is confusing.

2. It is noteworthy that all first-use cases were in cold/allergy products despite higher odds ratio for appetite suppressants.

3. Potential for prodromal symptoms to lead to use of cough-cold/allergy products.

4. Seven of eight "first-use" (cough cold) and five of six appetite suppressant cases were non-black females. Generalization to men or black women might, therefore, be inappropriate.

B. No consistent pattern of use, timing of exposure, or type of product use provides insight into a possible biological mechanism. More emphasis should be placed on physiology and metabolism of the drug in the final report.

C. Data regarding appetite suppressants is difficult to evaluate based on small sample size and lack of biological plausibility.

D. It is unlikely that a transient rise in blood pressure associated with PPA use explains the association seen in the HSP. However, alterations in the vasculature might be expected with chronic alterations in blood pressure.

E. One-sided tests are not appropriate, given the hypothesis being tested. In fact, the one previous analytic study reported a RR=0.59.

F. Lack of consistency in findings and unusual pattern of the data limits interpretability. Nonetheless, the study demonstrates that the population risk of a hemorrhagic event is extremely low.

V. Further Analyses

A. Analysis should be restricted to populations in which data are available.

1. Analysis should be repeated restricting inclusion to white women, as African Americans and men contribute no meaningful data to the overall analysis.

2. Analysis should be restricted to SAH, as there were no cases of intracerebral hemorrhage (ICH), which may be related to high fatality rate of ICH.

B. Stratification based on potential confounders

1. Stratify based on history of prior headache. (Not regression analysis because of small number of cases.)

2. Stratify by heavy versus light caffeine consumption.