# Yale University

*Walter N. Kernan, M.D.*
*Professor of Medicine*

*Pamela Doll*
*Senior Administrative Assistant*

*Telephone: 203 764-7000*
*Fax: 203 764-7552*

*Yale University School of Medicine*
*Suite 515*
*Chu Street South*
*New Haven, CT 06519*

7147

7 MAY 16 A9:34

May 11, 2007

Charles Ganley, M.D.
Director, Office of Non Prescription Products
Food and Drug Administration
Division of Dockets Management (HFA-305)
5630 Fishers Lane, Room 1061
Rockville, MD 20852

RE: Docket 1976N-0052N (nasal decongestant)
Docket 1981N-0022 (weight control)

Dear Dr. Ganley:

We are responding to the citizen's petition filed by Wyeth Pharmaceuticals critiquing the methods and results of the Yale Hemorrhagic Stroke study. The Wyeth critique is presented as the basis for their allegation that the HSP study is flawed epidemiology and cannot be relied upon as the basis for assessing the potential risks of phenylpropanolamine (PPA).

Valid and credible epidemiologic research depends upon methods of design, conduct and analysis that minimize the occurrence of chance and bias as explanations for a study's results. An essential element of valid epidemiology is the development of a sound research protocol and the use of unbiased methods in the conduct and analysis of the study. The HSP study developed a rigorous research protocol that was approved by the pharmaceutical sponsors of the study (including Wyeth Pharmaceuticals), the FDA, and the independent board of Scientific Advisors (chaired by Professor Louis Lasagna). The investigators presented the ongoing conduct of the study to the Scientific Advisory Group, including requests to implement any modifications to the protocol for either design or analysis. Central to the methods of the study were explicit criteria for case and control eligibility, subject exclusion criteria, definitions for classifying brain hemorrhages and PPA exposure, and methods of analysis. Whenever decisions were made about case or control eligibility, focal time in the study (to classify exposure to PPA or other potential drugs or substances), or any relevant methodologic issue the investigators strove to maintain adherence to the principle of blinding. This strategy helped to ensure that any errors in the research would not be systematic and would not introduce bias into the estimation of the odds ratio for the association between PPA and hemorrhagic stroke.

Wyeth Pharmaceuticals, its attorneys and their epidemiologists, operated under a completely different set of procedures and principles. Wyeth's epidemiologic consultants



1976N-0052N

C 252

knew the results of the research when they undertook their retrospective review of the study. They were aware that whenever they attempted to cast doubt on the credibility of a case subject classified as exposed to PPA, their attempt would have the effect of weakening the association between PPA and hemorrhagic stroke. They conducted their review without a defined protocol and without any evidence that they took necessary steps to "blind" themselves to PPA exposures. They restricted their post hoc analyses only to those that might weaken the PPA-stroke association. Thus, in the Wyeth analysis, exposed cases are closely examined for eligibility, but not unexposed cases. Included cases are subject to careful assessment of their eligibility and exposure status, but not excluded cases. And little effort is made to closely examine the included control subjects who are the basis for the comparison in this case-control study. Although the HSP investigators provided the Wyeth attorneys and epidemiologists all of the data collected for both included and excluded subjects, as well as all of the data for case and control subjects, the Wyeth attorneys and epidemiologists present a highly selected sample on which to base their complaints of flawed epidemiology.

These decisions are understandable in light of the apparent purpose of this post hoc review. But no one should think that such a review, with all of its obvious and insurmountable sources of bias, offers the basis for a credible epidemiologic assessment of the HSP study. We do not suggest that Wyeth or its epidemiologic consultants consciously introduced bias into their review and analysis. However, rigorous epidemiologic methods were never intended to prevent conscious bias alone. Rigorous methods are used to ensure that inadvertent and unconscious sources of bias do not arise as will occur if the end results of research are known by reviewers or if PPA exposure status is apparent to reviewers who are also deciding case or control eligibility. Indeed, the methodolgic weaknesses of the Wyeth review are so deep and their analysis so confounded by these sources of bias as to raise questions about whether any of the claims made in this citizen's petition warrant a specific rebuttal. However, knowing the importance of open debate and criticism for the scientific method, the HSP investigators have chosen to provide a point by point response to the allegations in the Wyeth petition.

## EXECUTIVE SUMMARY

The petitioner, Wyeth Pharmaceuticals describes itself as a former manufacturer of products that contained PPA. As a defendant in lawsuits brought by plaintiffs that allege injury as a result of exposure to such products, the petitioner alleges that our study demonstrating an association between PPA and increased risk for hemorrhagic stroke is "irreparably compromised." In attachment 1 to its Citizen Petition, the petitioner describes the basis for its allegation.

If the petitioner's erroneous allegations were left uncorrected, we believe the FDA and the public would be left with a false understanding of the HSP. The petitioner's allegations are largely based on restricted and misleading interpretations of clinical information, misinformation, and dogmatic methodologic reasoning. In some instances, the petitioner has advanced misinterpretations of comments by the investigators. In other instances, the petitioner has identified possible shortcomings of the HSP, but failed to accurately assess the small impact of these on the overall reliability of the research. Our purpose, therefore, is to reply to the allegations of the petitioner, providing information

that we believe will help the FDA and the public arrive at a more confident interpretation of the HSP results.

The petitioner is critical of the HSP in eight broad areas. Here is a brief summary or our replies:

**1) Allegation: Ineligible cases were included in the HSP**

The petitioner lists two case subjects that it alleges did not have events that meet the HSP case definition for hemorrhagic stroke. In both instances, the petitioner relies on restricted interpretation of clinical information and has, in our opinion, come to an incorrect judgment. In our reply, we reiterate the basis for the original decision to admit these case subjects. In both instances eligibility was confirmed by investigators blinded to PPA exposure status who were acting as part of a quality assurance program prior to data analysis.

The petitioner does not describe its methodology for reviewing the eligibility of HSP case subjects, but it provides no indication that all case subjects were reviewed by clinicians who were blinded to PPA exposure status. Assuming that no such comprehensive, blinded review was conducted, the petitioner's review must be understood as incomplete and potentially biased.

**2) Allegation: Exposure status was misclassified**

The petitioner lists four case subjects which it believes were erroneously classified as exposed to PPA containing products. Each of these case subjects reported consumption of a product containing PPA within the HSP exposure window preceding his or her focal time. We review the timetable for each case subject and explain the basis for our original exposure classification decisions, which we confirm are correct.

**3) Allegation: The HSP investigators modified the protocol and analyses during and after the study with knowledge that the changes affected the outcome.**

As in most large and complex clinical research, the HSP protocol was modified and improved in response to research experience and out of necessity to respond to practical and analytic realities. The number of modifications to the HSP protocol was few and those having a potential significant effect on the research were reviewed by the scientific advisory group prior to implementation. Of the four modifications identified by the petitioner, two were not modifications at all. Review of focal time and reclassification in the event of error was part of routine quality assurance. Our use of an analytic rule to examine potential confounding variables was consistent with our protocol and good research practice. Of the two other modifications, one involved selection of a dichotomous definition for education and it had no effect on the main findings of the study. The other involved a change in the definition of first use in response to compelling field experience.

In as much as the petitioner is suggesting that the investigators made changes aware of the direction and with intent to effect the overall study results, the allegations are false and misleading. Protocol changes were made rationally without knowledge or consideration of specific effects on the research outcome.

**4) Allegation: Blinding was compromised**

The petitioner correctly indicates that the HSP interviewers were not blind to the fact that the study was designed to examine the association between PPA and risk for hemorrhagic stroke. In response, we point out that blinding of interviewers is just one way to protect against biases that can affect case-control research. The HSP employed several strategies to reduce bias, including complete case ascertainment, a highly structured interview, and rigorous exposure verification procedures. It is extremely unlikely that bias resulted from interviewer awareness of the HSP purpose.

The petitioner also points out the participants in one of four research sites signed a consent form that announced the purpose of the study as involving PPA. We provide evidence that this advertisement did not affect the overall study results and point out that the advertisement was non-differential. That is, controls and cases were equally aware of the language in the consent form.

**5) Allegation: Confounding was not adequately controlled**

The petitioner argues that the association observed in the HSP between consumption of products containing PPA and increased risk for hemorrhagic stroke might be explained by factors other than PPA. They argue that the association may be a result of cocaine use, caffeine use, vascular lesions, trauma, tobacco use, alcohol use, menopause, low body mass index, and nicotine in pharmaceutical products. We reiterate our practices to manage potential confounding and explain why we believe that the listed factors were either 1) unlikely to explain the observed association between PPA use and increased risk for hemorrhagic stroke or 2) appropriately excluded from consideration as confounders. The HSP protocol was designed to address confounding and the protocol was followed through all stages of the research, with minor exceptions (e.g., not all cases and controls were matched on race). We correct factual and inferential errors made by the petitioner.

**6) Allegation: The HSP failed to obtain an adequate number of controls**

In its most surprising criticism, the petitioner argues that the observed rate of exposure to PPA-containing products in the control group is evidence of selection bias. The basis of their argument is that the observed rate of exposure to PPA-containing appetite suppressants (0.1%) is less than that anticipated from survey market research performed before the start of the research (0.64%). In our response, we review the purpose of contemporaneous controls in case-control research and argue that neither historical controls nor consumer market surveys are adequate substitutes for rigorous epidemiologic design and analysis. To support our view that case and control subjects in the HSP are drawn from the same underlying population, we provide evidence that they were similar in propensity to take non-PPA medications. We also describe later market research from 1998 that is consistent with the HSP exposure prevalence.

**7) Allegation: HSP deviated in other ways from the protocol and good epidemiological practices**

The petitioners assert that HSP interviews were not done properly, matching was conducted incorrectly, and the study was not population based. In our reply, we correct several factual and inferential errors of the petitioner. For example, the petitioner mischaracterizes the quality of interview data, misrepresents the control matching procedure, and overestimates any potential effect of referral bias. We explain why we

believe that none of these criticisms have substantial merit and are not likely to explain the results of the HSP.

**8) Other Allegations**

The petitioner is critical of the HSP investigators for their use of a one-tailed test of statistical significance in its final report to the FDA in May, 2000. In response, we point out that the one-tailed test was specified in the HSP protocol and approved by the sponsors and Scientific Advisory Group, and that sample size estimates were based on a one-tailed test.

Finally, the petitioner is critical of the HSP investigators for combining subarachnoid and intracerebral hemorrhage in the same analysis. We respond by pointing out that this plan was approved at the start of the research by representatives of industry, the FDA, the Scientific Advisory Group, and the investigators. The decision was based on clinical evidence at the time that PPA may be associated with either type of event.

In the eight areas of criticism above, the petitioner has identified most major potential problems that can affect case-control research. They have suggested that case and control subjects came from differing underlying populations as a result of selection and referral bias; that exposure to PPA was unequally uncovered in case and control subjects as a result of classification bias; that factors other than PPA explain the observed association between consumption of PPA-containing products and increased risk for hemorrhagic stroke as a result of confounding; that errors in eligibility classification resulted in the improper inclusion of exposed cases; that errors in focal time assignment resulted in improper exposure classification of some cases; and finally that the investigators manipulated the protocol and statistical testing to arrive at a specific research finding. In short, the petitioner has evoked most conceivable problems that, in theory, are capable of affecting case-control research to suggest an explanation for the HSP findings other than a true association between consumption of PPA-containing products and increased risk for hemorrhagic stroke.

The type of post hoc review conducted by the petitioner is probably unprecedented in the history of case-control research. Anyone who reads the petition, however, should be aware of fundamental way that this investigation differs from traditional peer review of science. First, the petitioner had access to all original research documents, drafts of publications, years of email correspondence among investigators, and days of deposition testimony by the investigators conducted by lawyers for the petitioner and other interested parties. This unprecedented access to data, were it interpreted without bias, might be regarded as a strong point of the petitioner's review. Second, the petitioner used many neurologists, lawyers, and epidemiologists to sort through material related to the HSP. Third, the review was conducted without the usual protections in peer review against conflict of interest.

Despite the intense scrutiny, after considering each of the petitioner's allegations, we remain confident that the findings of the HSP are not likely to be explained by bias, confounding, or error. The findings are not explained by misconduct. Our consideration of the petition itself has uncovered that the petitioner operated under false assumptions, misinformation, and mischaracterizations of the research process that appeared to have affected its review process. With the level of scrutiny applied by the petitioner to the

HSP, it would have been surprising if some research errors were not uncovered. All research is prone to error at various steps, but primarily in field execution and data transfer. None of the errors uncovered by the petitioner, however, explain the HSP findings. In summary, the HSP has received intense scrutiny by the petitioner, but the results hold up.

In science, the most important evidence in support of any finding is duplication by an independent group of researchers. Since completion of the HSP, a group of investigators in Korea completed a case-control study that examined the association between consumption of PPA in cough/cold remedies and risk for hemorrhagic stroke[1]. Case subjects were recruited from 33 hospitals in Korea during 2002-2004. Eligibility criteria included age 30 to 84 years, absence of brain lesions that would increase the risk of hemorrhage, absence of a history of stroke or trauma, and an ability to complete an interview within 30 days after the stroke event. Control subjects, two for each case subject, were selected among patients who visited the same institution as the case (hospital control) and siblings, friends, or neighbors of the cases (community control). Control subjects were matched to case subjects for age and sex. During a structured interview, a subject was counted as exposed if he or she reported use of a PPA-containing cold remedy within two weeks prior to the focal time. Based on analysis of data from 940 case subjects and 1880 control subjects, the unadjusted odds ratio for the association between consumption of a PPA-containing product and risk for stroke was 2.46 (95% CI 1.15 to 5.24). The adjusted OR was 2.14 (95% CI 0.94 to 4.84). The finding for an association between PPA and risk for stroke was evident in women (adjusted odds ratio=3.86, 95% CI 1.08, 13.80), but not in men (adjusted odds ratio =1.36, 95% CI 0.45, 4.15). The main findings in the HSP were for an association with PPA in appetite suppressants in women who were exposed within over a 3 day window (odds ratio=16.58, 95% CI 1.51, 182.2); an association between PPA in cough-cold remedies was suggested in an analysis of first use exposure, but did not reach statistical significance according to a two-tailed test (odds ratio 3.13, 95% CI 0.86, 11.46). By comparison, the study by Yoon et. al. did not test the association between PPA in appetite suppressants and risk for stroke, but demonstrated an association between consumption of PPA in cough-cold products and increased risk for stroke over a 14 day exposure window.

## RESPONSE TO ALLEGATIONS BY WYETH

### 1. Allegation: Ineligible cases were included in the HSP

In selectively reviewing the eligibility of exposed, enrolled HSP participants, Wyeth has engaged in an exercise that may lead to a biased manipulation of the HSP findings. In particular, the company's methodology can only lead to the exclusion of exposed case subjects and reduce the observed odds ratio for the association between PPA and risk for hemorrhagic stroke. A more unbiased, scientifically defensible review would include a review of all case subjects, including potential case subjects who, upon original determination of eligibility by HSP investigators, were excluded from the research.

Despite our objection to Wyeth's methods, we reply here to the company's specific claims.

**1.1 Case 62-0094**

According to her medical records, this 36 year-old woman had a witnessed, generalized seizure in a hospital emergency department waiting room. Upon evaluation immediately after the seizure, she complained of a severe headache. There was no reported history of head trauma before or during the seizure. The examining physician in the emergency department indicated that there was no evidence for head trauma. A CT scan showed a subarachnoid bleed in the right Sylvian fissure. An arteriogram showed no aneurysm. The day after her hospital admission, a neurosurgery resident wrote, "probable traumatic SAH."

When interviewed by an HSP researcher, the patient indicated she was sitting in the emergency room of a hospital, keeping company with a friend who was waiting to be seen. She next remembered her friend picking her up from the floor when she had a severe headache. She remembered no other details.

Wyeth alleges that this case subject should have been excluded as a traumatic brain hemorrhage. Upon blinded review by HSP investigators prior to the close of the study, however, the patient was determined to meet the HSP eligibility criteria. That is, a decision was made that the event was a primary hemorrhage rather than a seizure followed by a traumatic hemorrhage.

**1.2 Case 20-0092**

According to history we obtained from the case subject, she took a decongestant (later confirmed to contain PPA) at 12:30pm on 10/23 for congestion. At 1:30pm, while driving to work, she developed a "massive headache." At 2:30pm, emergency services were summoned to take her to the hospital. According to emergency department records, she was having mild headaches at home in the morning before going to work for which she took "Sudafed" (which did not contain PPA). Her physical examination in the emergency department was remarkable for a blood pressure of 188/96, but no focal weakness. A head CT was normal. On arrival back in the emergency room after the CT scan, her condition was unchanged. About two hours later (~6:40pm), however, she was noted to have left sided weakness and slurred speech and heparin was started in response. An MRI the following day at 4:57 pm showed a hematoma in the area of the right basal ganglia.

In the field, the HSP interviewer originally determined that the focal time was 1:30pm. In reviewing this case without knowledge of exposure status, the investigators at the HSP Coordinating Center concluded that the focal time for this patient was 6:40pm after the CT scan when she developed hemiparesis. We knew the focal time could not be before the time of the CT scan that showed no hemorrhage.

Wyeth argues that this patient did not meet the HSP definition for hemorrhagic stroke and, if she did, she should have been counted as unexposed. The petitioner's arguments reflect a misunderstanding of the HSP protocol and the pertinent clinical events in this case. In arguing that the patient did not meet the HSP definition for hemorrhagic stroke, the petitioner states that "the treating doctors ruled out hemorrhage on October 23 and diagnosed an ischemic event – a transient ischemic attack." The hemorrhage that was

eventually detected, the petitioner argues, occurred as a result of heparin therapy which was an exclusion factor for HSP. We agree that the patient did not have evidence for a hemorrhage at the time the CT was performed on October 23. Based on her clinical course, however, the hemorrhage occurred after the CT scan when she developed persistent, focal neurological deficits consistent with a stroke. Although the initial diagnosis after the CT scan on October 23 by her neurologist was an ischemic event, a hemorrhage was demonstrated on an MRI from October 24. The HSP investigators cannot be completely certain that the hemorrhage occurred before heparin administration, but during the review (blind to PPA-exposure status), they believed that this temporal sequence was of sufficient probability to properly include the patient in the study. Among the factors that influenced the decision to include the patient were the clinical and radiographic data available at the time eligibility was determined. We are seeking additional clinical records to inform a review of the decision.

Regarding the petitioner's claim that this patient, if classified as a hemorrhage, should have been counted as unexposed, we believe this reflects a misunderstanding of the HSP protocol. HSP patients were classified as exposed if they consumed a PPA-containing product on the index date before focal time or in one of the preceding 3 days. The company seems to argue that this patient's focal time was in the morning of her index date. This argument, however, is not consistent with the clinical evidence. The patient reported taking the PPA-containing medication at 12:30pm for congestion. The mild headache she experienced in the morning would not define a focal time as specified in the protocol (i.e., development of symptoms compatible with hemorrhagic stroke that leads a patient to seek medical attention). Symptoms that prompted her to seek medical attention occurred later at about 1:30pm. The focal time was eventually set at 6:40pm when she developed persistent, focal deficits after a CT scan showing no hemorrhage. In this unusual case, we have imaging data to help in the establishment of a more precise focal time.

## 2. Allegation: Exposure status was misclassified

### 2.1 Case 71-0039

According to history obtained by HSP researchers, this subject developed a headache sometime after 4pm on 10/31. The headache seemed "normal" at 3/10 intensity until approximately 9pm on 11/5 when it got a "little stronger," prompting the subject to take a medication (later confirmed to contain PPA) before visiting a relative. About 15-20 minutes after going to sleep that night (at approx. 10pm), the patient awoke with a severe headache, paresthesias in the right arm, loss of hearing, and blurry vision. Her husband took her to the hospital. Wyeth Pharmaceuticals alleges that this case subject had a subarachnoid hemorrhage prior to consumption of PPA; that is, they would classify 9pm on 11/5 as the focal time.

The patient's research record (excluding exposure information) was reviewed for data quality in 1998. During this review, we sought to apply the HSP definition for focal time, which was the time that marked the onset of symptoms plausibly related to hemorrhage that caused the patient to seek medical attention. The HSP interviewer had identified the focal time as 9pm on 11/5. In 1998, the reviewer recognized that the

researcher had not accurately applied the criteria for focal time; it was the symptoms after onset of sleep, not the symptoms at 9pm, that had prompted the patient to seek medical attention. The case was reviewed with two other investigators who were blind to exposure data and who concurred with the decision to change the focal time.

**2.2 Case 20-0092**

Wyeth argues this patient (described in section 1.2) should have been counted as unexposed because of "the presence of a sentinel headache." We agree that this patient probably had a sentinel headache that could reasonably have been classified as occurring either on the morning or at 1:30pm on 10/23. Medical record entries indicate that the morning headache was described by the patient as mild and did not prompt her to seek medical attention.

The Wyeth petitioners appear to misunderstand the HSP protocol. The HSP did not determine exposure status based on sentinel symptoms. Sentinel symptoms were defined as headache symptoms occurring before the onset of symptoms that caused a patient to seek medical attention. Sentinel symptoms were typically mild and transient, distinct from the symptoms that caused patients to seek medical attention. Recognizing that mild, transient headache is a common symptom among healthy persons and that it does not necessarily indicate hemorrhage, the investigators did not use sentinel symptoms to define focal time. Throughout the HSP, the investigators used standard definitions for hemorrhagic stroke and focal time and applied them uniformly. For this case subject, the symptoms that caused her to seek medical attention began after she consumed a product containing PPA and she was appropriately classified as exposed according to the HSP protocol.

We excluded this case subject in a secondary analysis that excluded all subjects with a definite or possible sentinel headache and their matched controls. The findings for an association between consumption of an appetite suppressant containing PPA and increased risk for hemorrhagic stroke among women remained.

**2.3 Case 18-0025**

Wyeth Pharmaceuticals alleges that this non-English speaking patient should have been classified as non-exposed, arguing that the patient consumed a PPA-containing product only after the onset of her symptoms. Wyeth cites data from the hospital records, including a 12-line "personal history" indicating that the headache began on the morning of hospital admission. Our own interviewers, however, used a translator to obtain an independent history of the present illness. Unlike the Wyeth reviewers, our interviewer was unaware of the patient's exposure history as she recorded the history. The patient told our interviewers that she was in her usual health until about 7pm when she suddenly developed the worst headache of her life. In assigning focal time, our standard procedure was to rely on the history we obtained directly from patients. This history was obtained by trained researchers who conducted an inquiry specifically constructed to gather the information needed to assign a focal time. History recorded in medical records, when available, was reviewed and considered. When a researcher was unsure of the history for any reason, he or she could contact an investigator before proceeding with the interview. Our standard procedures were followed in this patient; we believe our history is correct and that the exposure was not misclassified.

**2.4 Case 20-0297**

Wyeth alleges that this non-English speaking patient had a headache before consuming a product containing PPA and, therefore, she should be counted as non-exposed. The company bases its allegation on hospital records. A nursing note from 2:30am on 7/4/97 in the emergency department notes, "c/o 'bad headache' which began one hour ago – took Excedrin and imetapp without effect." A neurosurgery resident reported that, "patient has had a head cold for which she took Exedrin and 'cough medicine' today after which her headache reportedly became worse so brother brought her to hospital." A neurology resident later reported that she "presented to ER at ~12am secondary to headache not relieved by Anacin and Dimetapp."

When interviewed by HSP researchers on 7/8/97 (brother translating), the patient indicated that she was in her usual state of health until 3 days prior to admission when she developed symptoms of a respiratory illness, including a cough. While watching TV at home at 11:30pm on the evening of admission, she developed the "worst headache of life." She reported no headache to the interviewer prior to 11:30pm. She felt like she was going to die and was taken to the hospital. A CT scan showed a left basal ganglia hemorrhage. This patient was assigned a focal time of 11:30pm based on the HSP criteria. The interviewer ascertained and verified use of a product containing PPA in the evening of the index day for "cold" symptoms. The patient did not mention headache as a reason for taking the medication despite specific inquiry on her reasons. In assigning focal time for this patient, we followed our protocol and relied on information obtained directly from the patient by a trained interviewer. We believe the protocol was followed correctly for this patient and that the focal time was appropriately and correctly determined.

**2.5 Case 31-0001**

Wyeth argues that this case subject should have been classified as non-exposed because she last consumed a PPA-containing product 84 hours before her focal time. This patient's exposure status was assigned strictly according to the HSP protocol, which defined the window of exposure as the index day before the focal time and the preceding three calendar days. This case subject reported that she consumed a PPA-containing product at 11:00am on 2/22/95. Her focal time was 11pm on 2/25/95.

We disagree with Wyeth that this case subject should have been classified as non-exposed. To classify this patient as non-exposed would have represented a deviation from the research protocol which had been approved by the FDA, industry representatives, and the Scientific Advisory Group that oversaw the HSP.

One of the experts for Wyeth, Dr. Joseph L. Lyon (Affidavit, page 5, Tab 39 of Citizen Report) reports discrepancy in the time this patient last used a PPA-containing produce prior to her focal time as recorded in her research form. In fact, there was no discrepancy. Her last use was recorded as 2/22/95 in both relevant portions of the form.

**3. Allegation: The HSP investigators modified the protocol and analyses during and after the study with knowledge that the changes affected the outcome.**

### 3.1 Reclassification of focal time

Wyeth asserts that in "at least one instance" a case subject was re-classified from not exposed to exposed to a PPA-containing product by an unblinded investigator.

For purposes of quality assurance, an investigator at the HSP Coordinating Center (primarily Walter Kernan) reviewed medical records and research forms for case subjects (i.e., forms 01=HSP Case Identification, 02=HSP Event Evaluation, 02c=HSP Event Evaluation from subject, 03c=HSP Subject Screening for case). The investigator performing the eligibility review did not review forms 04 (Subject Interview) and had no access to data on verified exposures. For a very small number of cases, however, reported (non-verified) medication uses were noted in medical records or (inadvertently) on forms 02 or 02c. If focal time was changed for a case subject, it was also changed for the matched control.

Wyeth quotes selectively from an email message from Walter Kernan to Catherine Viscoli regarding case 71-0039 (described in section 2.1) in which Dr. Kernan wrote, "I believe that Tuesday 10/31 should probably be the FT {focal time} and ID." Changing the focal time to 10/31 would have resulted in this patient being reclassified from exposed to unexposed. The full text of the message indicates greater uncertainty regarding the correct focal time and a commitment on the part of the investigators to manage the uncertainty responsibly: "Others will probably argue that the primary index date was 10pm 11/5 and that the sentinel date is 10/31 around 4pm. The safest thing to do may to to (sic) use 11/5 as the primary and 10/31 as the sentinel date, resolving ourselves to do two analyses with the two different index dates/focal times." The proposal to undertake an analysis using alternative focal times amounts to a sensitivity analysis that accounts for relevant uncertainty. This case subject was excluded from the secondary analysis of cases without sentinel syndromes reported in the main HSP publication. The petitioner has called attention to Dr. Kernan's acknowledgement in the email described above that changing the focal time from 9pm (time assigned by the study coordinator who interviewed the subject) to 10pm would increase the likelihood that the patient would be counted as exposed to a PPA-containing product. Dr. Kernan's comment summarizes the obvious: If you advance the focal time you increase the opportunity for exposures to occur. The petitioner references this acknowledgement in a manner that implies that the investigators were attempting to manipulate the results of the study rather than seek a correct classification of focal time. Wyeth's position fails to consider the facts of the case history.

The petitioner asserts that the HSP investigators who reviewed case subject 71-0039 were not blinded to "the data for this case." The petitioner's assertion is not correct. Dr. Viscoli was unblinded to exposure data, but she did not participate in focal time or eligibility adjudication decisions. The two physician-investigators who adjudicated the focal time, however, were blinded to data related to this case subject's exposure to PPA.

### 3.2 Modification of definition for "first use"

In the original HSP protocol, participants were to be considered exposed to PPA as a 'first use' if they reported no other use within 7 days. At a meeting of the HSP Scientific Advisory Group on May 6, 1997 after 335 subjects had been enrolled, the interval defining first use was changed to 14 days. This change was the result of unanticipated uncertainties in exposures created by data collection procedures.

If an HSP subject reported use of a medication in the three days prior to their focal time, an interviewer would ask, "did you take this medication at any other time during the two week period before the {index date}." If the subject answered affirmatively, the interviewer would probe for the days on which the medication was used. Our intention in phrasing the interview in this way was to acquire as much information on drug use as possible. In practice, however, we found that some subjects could report use during the 2-weeks with certainty, but had difficulty recalling dates with enough specificity to define their exposure over the initial 7 days of the period.

Changing to a 14-day period was necessitated strictly by field experience with this method of data collection. In discussions of the proposed change with the SAG in May 1997, it was noted that use of the longer wash-out period to define 'first use' would lower the statistical power to detect an elevated odds ratio (because fewer patients and control subjects would be counted as exposed to a first use), but could lead to a more accurate estimate of the odds ratio (because a 14-day period more closely approximates first use than a 7-day window). In the end, the investigators and members of the Scientific Advisory Group were willing to give up statistical power (to detect a significant association between PPA use and increased risk for hemorrhagic stroke) for enhanced accuracy. The minutes from the May meeting 1997 meeting incorrectly report Dr. Samy Suissa's opinion on this topic; he was reported to say that shortening the window defining first use would enhance reliability (accuracy) when, in fact, he actually said it would reduce accuracy. The ideal definition for first use is no use ever prior to a defined use. Because people cannot be expected to reliably recall their exposure history since birth, this ideal definition is impossible to implement. Epidemiologic investigators are required, therefore, to specify pragmatic time intervals of non-use to define first use. They seek a balance between the ideal and that which is achievable within a person's ability to recall exposures. Longer intervals more closely approximate the ideal definition. Dr. Suissa's actual comment during the May meeting reflects this fact.

The authors of the Citizen's Petition argue that the interval definition for first was changed "after reviewing interim data and knowing that the change would double the 'first-use' odds ratio and increase the likelihood of finding a statistically significant association." In fact, the first interim analyses (in which first use could be calculated using both 7 and 14 days exposure windows) were presented only to the Chairman of the Scientific Advisory Group and were not discussed with the SAG during its deliberations on the practical need to revise the first use definition.

### 3.3 Adjustment for the effect of education

In the Citizen Petition, Wyeth accuses the HSP investigators of testing several different strategies for classifying educational attainment and selecting one on the basis of its effect on the association between first use exposure to PPA and risk for hemorrhagic stroke among women. This particular allegation is part of a more global criticism of the investigators' strategy for statistical adjustment of the observed odds ratios.

When the decision to use education as a bivariate variable was made on 1/5/2000, we considered the four analyses listed in Table 1. These were the same four aims we used to examine the effect of all potential confounders.

| Table 1. Effect of various strategies for adjustment for education on odds ratios for four exposure definitions*. | | | | |
|---|---|---|---|---|
| Education | Cough-cold in 3 days (all) | AS in 3 days (women) | Any PPA in 3 days (all) | 1st dose (women) |
| **2 – level (HS)** | **1.30 p=.198** | **16.58 p=.011** | **1.57 p=.062** | **3.13 p=.042** |
| 2 – level (Coll) | 1.38 p=.150 | 15.22 p=.013 | 1.67 p=.044 | 2.68 p=.070 |
| 3 – level | 1.37 p=.156 | 16.41 p=.014 | 1.65 p=.047 | 2.62 p=.076 |
| Continuous | 1.31 p=.200 | 17.74 p=.016 | 1.58 p=.068 | 2.00 p=.160 |
| * All results are also adjusted for the following pre-specified variables: race, hypertension and smoking (a-priori adjustment features). P-values are one-sided. | | | | |

As these results demonstrate, none of the choices for education adjustment affected the findings for appetite suppressant use in women materially and a significant finding (1-sided) for any PPA use would have resulted if the college/no college dichotomy had been chosen. In fact, the choice of high school/no high school dichotomy for the final report to the FDA and for the main publication resulted in the lowest point estimates for the association of cough-cold and any PPA use findings.

The final decision to use a high school/no high school dichotomy as the measure of education was not based on the examined results, but rather on prior published work demonstrating that this categorization was the strongest predictor of health outcomes[2].

The petitioner is not correct in stating that two different definitions for education were used for the final report to the FDA (dated May 9, 2000) and the main publication. In both reports, education was dichotomized as high school/no high school. In both reports, the adjusted odds ratio for the association between first use of PPA and risk for hemorrhagic stroke in women was 3.13. The p-values for this odds ratio differed between the FDA Final Report and the publication because of the choice of one-sided compared with two-sided statistical testing. The choice of one-sided testing compared with two-sided testing is clearly stated in each document.

### 3.4 The 10% rule was not in the HSP protocol and was not applied

In the final report of the HSP findings, the association between PPA and risk for hemorrhagic stroke was adjusted for known risk factors (race, smoking, hypertension) plus pre-specified established or suspected risk factors that were potential confounders and which changed the odds ratio by >10% in any of the 4 exposure definitions (PPA in cough-cold in 3-days, appetite control in 3 days-women, any PPA in 3-days, 1st use-women). The 15 pre-selected variables were specified before the analyses were undertaken and included hypertension, diabetes, congestive heart failure, polycystic kidney disease, sickle cell anemia, hemophilia, clotting disorders, parent or sibling history of hemorrhagic stroke, obesity (body mass index >30), current cigarette smoking, consumption of >14 alcoholic beverages per week, cocaine use within 1 day of focal time, oral contraceptive use within 3 days of the focal time, black race, and education. The HSP investigators discussed other approaches to adjustment of potential confounding features, including estimating different models for each aim. This specific model approach was rejected in favor of the use of a consistent model of adjustment (i.e., one that included race, smoking hypertension) and for all aims because it was less prone to over-tailoring models for the dataset and had the advantage of ease of interpretation. Our

use of the 10% rule in this manner is well established in epidemiology research[3-5], although some researchers have used a lower 5% limit[6] and other simply recommend use of a "clinically important" change[7].

Wyeth is critical of the HSP investigators for not adjusting the main analyses for pharmaceutical caffeine exposure despite the fact that caffeine changed the observed odds ratio for appetite suppressant use by more than 10%. Adjusting for caffeine in drugs did increase the odds ratio from 16.58 to 18.39 for the association between PPA as an appetite suppressant and risk for hemorrhagic stroke, with only marginal changes in the other analyses. We did not adjust for caffeine in drugs because it was not among the 15 potential confounders we identified a-priori for examination. We believe our decision not to adjust for caffeine in pharmacologic agents was correct. The decision was in accordance with our pre-specified analytic plan. Because caffeine is ubiquitous in soft drinks and other beverages, furthermore, adjustment for use of OTC caffeine pills is not adequate to control for this factor. In preparing this response to the Citizen Petition, we realized that the main HSP publication incorrectly describes the potential confounders we examined. In the publication, we indicate that we examined all variables in tables 2 and 3 of that publication, when in fact we only examined the 15 listed above.

The HSP protocol states that data on potential confounding variables will be collected and that "we plan to conduct suitable multivariable analyses to adjust for the effects of potential susceptibility variables." As an operational means to select potential confounding variables for inclusion in adjusted models, the 10% rule was adopted prior to the start of the analysis.

**4. Allegation: Blinding was compromised**

Wyeth speculates that failure to blind interviewers to the study hypothesis may have resulted in bias. The company is presumably referring to bias in exposure ascertainment and case selection bias.

Blinding to hypothesis is just one potential mechanism to prevent ascertainment bias. In the HSP, we put in place procedures and practices to assure a uniform approach to exposure ascertainment in every case and control subject. Specifically, each subject was interviewed using a standard script and interview instrument that reduced or eliminated the influence of interviewer opinion. We believe these policies and practices were effective in reducing the risk of ascertainment bias.

Wyeth points out that subjects in one research site, Rhode Island, were asked to sign an informed consent document that referred to PPA. While it is possible that this advertisement may have prompted case subjects to recall PPA use, it would also prompt control subjects to do the same. It is impossible to know if this advertisement differentially affected case and control subjects or if it affected them at all. We believe this advertisement of PPA as a study focus was not likely to affect exposure ascertainment, given the highly structured interview instrument we used that prompted both case and control subjects to recall exposures to diet pills and cough and cold remedies.

One way to eliminate the potential ascertainment bias in Rhode Island is to re-do the HSP analysis without the 99 cases and 194 matched controls identified at that research site. The results do not affect the overall findings of the HSP. Specifically, the

odds ratio for the association between PPA use as an appetite suppressant and risk for hemorrhagic stroke remains elevated (OR=14.19), as does the odds ratio for the association between first use of PPA and risk for hemorrhagic stroke in women (OR=3.18). The association between any use of a PPA-containing product and risk for hemorrhagic stroke also remains elevated (OR=1.30).

## 5. Allegation: Confounding was not adequately controlled

### 5.1 Major risk factors were more prevalent in the case group

Wyeth argues that unequal distribution of stroke risk factors between the case and control groups is evidence for uncontrolled confounding. Unequal distribution, however, is not unexpected. We would expect that hypertension and cigarette smoking would be unequally distributed among cases and controls because these are known to be associated with increased risk for stroke. Both of these were included in adjusted models. Race was a matching variable that we included in adjusted models because 4% of case-control pairs that were not adequately matched. These pairs were included in the analysis because we believed that exclusion would have a more adverse effect on the study than inclusion.

Wyeth argues that we could not adjust for illicit drug use because five PPA exposed case subjects had a history of illicit drug use, whereas none of the exposed controls were cocaine or heroin users. In fact, we did examine cocaine use as a pre-specified potential confounder in each of the four aims listed in Table 1. Our results indicated that adjustment did not alter the findings for two of the study aims (cough/cold use in 3 days, any PPA use in 3 days). In the analyses restricted to women (appetite control use in 3 days and 1$^{st}$ dose use), adjustment for cocaine was not needed because cocaine use was only reported in 5 female case subjects and none of those subjects reported use of PPA.

### 5.2 No adjustments were made for numerous confounders

Wyeth points out that no adjustment was made for cocaine use, family history of hemorrhagic stroke, or regular use of alcohol. None of these features met the specified criterion for inclusion in adjusted models in which we analyzed the association between consumption of PPA-containing products and risk for hemorrhagic stroke (i.e., did not change any OR by more than 10% when added to the basic model).

Wyeth also criticizes the HSP investigators for not adjusting for risk factors that emerged in subsequent papers based on the HSP data that focused on aneurismal subarachnoid hemorrhage and intracerebral hemorrhage. The company is presumably referring to menopause, consumption of $\geq$5 caffeinated drinks/day, consumption of caffeine in drugs, low body mass index, and nicotine in pharmaceutical products[8, 9]. None of these variables were identified at the time of our original analyses as established or suspected risk factors for hemorrhagic stroke and, therefore, were not examined.

### 5.3 Many of the cases were confounded by multiple risk factors

Wyeth states that "the most profound example of the problem of multiple risk factors and their potential impact on the interpretation of the HSP, however, exists among the six cases of stroke in women who took PPA for appetite suppression. Although these six appetite suppressant cases provide the strongest foundation in the Notice for FDA's

proposal, five of the six can be explained by other, clear causative factors, including pre-existing lesions, trauma, and use of multiple substances (cocaine, tobacco, alcohol, caffeine)." The company argues that these women each had risk factors for stroke and that these risk factors, not PPA or the combination of PPA and other risk factors, were responsible for their stroke events.

The HSP protocol specifically did not exclude case subjects with a subarachnoid or intracerebral hemorrhage related to a previously unrecognized vascular lesion, such as an aneurysm or AVM. There were three main reasons for not excluding patients with newly discovered vascular lesions. First, the investigators and members of our Scientific Advisory Group recognized that patients with unrecognized vascular lesions may be at particular risk for adverse effects of PPA. Second, we had no safe and effective way to screen control subjects for asymptomatic intracranial vascular lesions and, therefore, could not apply such an exclusion criterion equally to case and control subjects. Third, the investigators were aware of case reports of intracranial hemorrhage associated with sympathomimetic agents involving patients with vascular lesions[10, 11]. After the start of the HSP, further reports appeared[12].

Wyeth alleges that patient 62-0094 experienced a SAH as a result of trauma. The HSP investigators who reviewed this patient for eligibility in the field and at the Coordinating Center and who were blinded to PPA exposure status determined that her SAH event was most probably not the result of trauma. We disagree, therefore, with Wyeth's characterization of this patient's stroke event as being related to trauma.

The relevant comparison for the effect of any exposure on stroke risk is between cases and their matched controls. In reply to the complaint of confounding of the association between appetite suppressant use and hemorrhagic stroke by 'multiple substances', we display below the exposure status for the seven exposed female subjects and their matched controls (or case). Among the multiple substances noted by Wyeth in their petition, alcohol use and cocaine use were examined for their effect as confounders and neither met the criteria for inclusion. Caffeine use in pharmaceuticals was not considered in the analysis, but as shown in the table, only one subject (a control subject) was exposed. Smoking and hypertension as a priori risk factors for stroke were included in all adjusted analyses.

Table 2. Exposure to selected risk factors among 7 subjects exposed to PPA in appetite suppressants and their matching subjects.

| Study # | Case/ Ctl | Hypertension Hx. | Current Smoker | Alcohol Use[1] | Cocaine Use[2] | Caffeine in Rx.[3] |
|---|---|---|---|---|---|---|
| 31-0001 | C* | - | - | - | - | - |
| | A | - | - | - | - | - |
| | B | - | - | - | - | - |
| 33-0059 | C* | + | - | + | - | - |
| | A | - | - | - | - | - |
| | B | - | - | - | - | - |
| 35-0043 | C | + | + | - | - | - |
| | A* | - | - | - | - | - |
| | B | - | + | - | - | - |
| 46-0080 | C | - | - | - | - | - |
| | A | - | - | - | - | + |
| | B | - | + | - | - | - |
| 46-0201 | C* | - | + | - | - | - |
| | A | + | + | - | - | - |
| | B | - | - | - | - | - |
| 62-0094 | C* | - | + | - | - | - |
| | A | - | + | - | - | - |
| | B | - | + | - | - | - |
| 71-0398 | C* | + | + | - | - | - |
| | A | - | - | - | - | - |
| | B | - | - | - | - | - |

[1]>14 drinks/week(average). [2]Cocaine used in 1-day window. [3]Caffeine in pharmaceuticals used in 3-day window.
*Exposed subject

**6. Allegation: The HSP failed to obtain an adequate number of controls**

Wyeth argues that the observed exposure rate among control subjects is evidence for a "serious deficiency in the data gathering process for finding matching controls" because it deviated from exposure estimates obtained from consumer marketing surveys that were used in creating the sample size estimate for the HSP. The argument put forth by Wyeth is contrary to the logic that compels investigators to use concurrent control subjects for case-control research. Carried to its logical conclusion, Wyeth's argument suggests that results from marketing surveys could substitute for control group exposure estimates, obviating the need for conventional case-control research. This argument would seem to reflect a misunderstanding of epidemiologic design, analysis, and inference.

Wyeth seems to be suggesting that the observed exposure prevalence among the HSP control group is evidence for selection bias. The estimates in our sample size estimation were based on a nation-wide marketing survey that relied on self-report use. In contrast, the HSP took place in just four geographic areas and employed a rigorous procedure for exposure confirmation that included physical inspection of pill bottles or identification in product books. Exposures that could not be confirmed were excluded.

In addition, if selection bias explained the observed exposure rate, similarly low exposure rates might be expected for other commonly available drugs. To check for propensity to take medications, we performed the analysis shown in Table 3. Exposures for other drugs were quite similar between case and control subjects and do not suggest that control subjects were inherently less likely to take a variety of medications used to treat common ailments. Caffeine and nicotine (in drugs) were more commonly used by case subjects compared with controls. The finding for caffeine may indicate that this is an unrecognized risk factor for stroke. The finding for nicotine is probably explained by the higher rate of use of cigarettes among case subjects compared with controls.

Table 3. Use of common OTC medications among patients and control subjects in the HSP.

| | Patients (N=702) | | Controls (N=1376) | | |
|---|---|---|---|---|---|
| Agent | N | % | N | % | OR |
| Aspirin | 79 | 11.2 | 133 | 9.7 | 1.19 |
| NSAIDS | 115 | 16.4 | 293 | 21.3 | 0.72 |
| Dextromethorphan | 25 | 8.2 | 116 | 8.4 | 0.98 |
| Sympathomimetic* | | | | | |
| Oral Preparations* | 58 | 8.2 | 116 | 8.4 | 0.98 |
| Inhaled Preparations | 11 | 1.6 | 32 | 2.3 | 0.67 |
| Nasal Preparations | 8 | 1.1 | 15 | 1.1 | 1.05 |
| Stimulants/anorexiants | 4 | 0.6 | 12 | 0.9 | 0.65 |
| Oral anticoagulants | 2 | 0.3 | 6 | 0.4 | 0.65 |
| Caffeine (in drugs) | 49 | 7.0 | 40 | 2.9 | 2.51 |
| Nicotine (in drugs) | 9 | 1.3 | 1 | 0.1 | 17.86 |

*pseudoephedrine HCL, phenylephrine, ephedrine, epinephrine

Rather than proving selection bias, the difference between the observed and estimated exposure rates in control subjects may illustrate the importance of standardized data acquisition and use of contemporary controls. In case control research, as in other research on causal associations, use of historical controls may lead to erroneous inferences. In particular, historical data from consumer-based marketing survey research does not qualify as a substitute for rigorous epidemiological research.

Having described the fallacy of the petitioner's inferences based on its comparison of the observed and expected PPA exposure rate in our control group, we wish to point out survey marketing data that suggests our observed value was close to the national average reported in surveys conducted as the HSP was on-going. In 1998, Novartis Pharmaceuticals commissioned Consumer Health Sciences (CHS) to estimate consumer exposure to pharmaceutical agents containing PPA and US adults. CHS obtained data from 16,619 individuals over age 18 years who were asked to recall prescription and over-the-counter drugs taken during the month preceding the survey (May, 1998). Among all respondents, 57.7% were ages 8-49 years, the range included in the HSP. Results indicated that 3.8% of patients took any PPA-containing product during the month and 0.08% took an OTC appetite suppressant containing PPA. Three-day exposure prevalence rates in the survey were not reported, but presumably would have

been somewhat lower for all product categories. The survey results compare with 3-day exposure prevalence rates in the HSP control group of 2.4% and 0.1% for any use and use of an appetite suppressant containing PPA, respectively[13]. In contrast to the allegation of the petitioner, the rates of exposure to appetite suppressants containing PPA among HSP control subjects would appear to be about equal to, and perhaps higher than, the rate expected from consumer research.

Wyeth continues its argument to state that "lack of a sufficient number of controls artificially increases the chance of finding a high odds ratio between cases and controls." Here again, we believe that the company misunderstands the statistical and research issues. In the absence of selection bias, which we believe did not exist, increasing the number of control subjects would not substantially affect the magnitude of the observed odds ratio, only the p-value.

## 7. Allegation: HSP deviated in other ways from the protocol and good epidemiological practices

### 7.1 Interviews often were done improperly

Following the HSP protocol approved by the Scientific Advisory Group, the FDA, and representatives of industry before the start of the research, HSP control subjects were interviewed in mutually convenient locations that afforded adequate privacy. Several control subjects requested to be interviewed outside their homes. Rather than lose these control subjects and following the protocol, we respected their wishes. HSP interviewers were instructed to assure that interviews took place under conditions of adequate privacy. Each interview was timed. The interview instrument was designed to be simple and short. Questions about recent exposure to drugs were purposely placed at the beginning of the interview to assure they were completed if, for any reason, the interview was truncated. Interviewers were instructed to skip many questions depending on responses. For a control subject who reported no drug exposures, alcohol use, tobacco use, or significant medical history, the interview could easily be completed in a short time. For 1368 control interviews with times recorded, the average duration was 36 minutes (minimum of 5 minutes); for 689 case interviews with times recorded, the average duration was 47 minutes (minimum of 7 minutes).

Wyeth alleges that "some interviewers were rushing through the questionnaire, skipping questions, and not probing for critical exposure on medications." This allegation is not consistent with the facts and mischaracterizes the quality of the HSP interviews. Interviews were routinely checked for completeness as soon as they arrived at the HSP Coordinating Center. Any omissions or errors were immediately addressed with the interviewer. There were no interviews in which a subject was not queried about illness symptoms and medication use in the 2 week period. In addition, all subjects were asked about any other medication uses and a similar proportion of control and case subjects reported such uses (62% and 59%, respectively). The finding of comparable or higher rates of exposure to several pharmaceutical agents among control subjects, compared with case subjects, is further evidence against Wyeth's allegation (see Table 3).

In responding to Wyeth's allegations regarding the quality of interviews, we examined the 10 interviews included in Tab 52 of attachment 1 to the Citizen Petition.

All interviews were fully completed with no significant, uncorrected errors. Contrary to the allegation of the petitioner, no questions were skipped. For 9 of the interviews examined, the interviewers commented they were "very confident" or "fairly confident" in the accuracy of the history provided (on a scale of very confident, fairly confident, confident, somewhat confident, little or no confidence). For one interview that was recorded as lasting 7 minutes and took place in the control subject's home, the interviewer was "fairly confident", but the interview form indicates detailed information on family history, illicit drug use, and occupation that suggests the interview was of good quality, within the ability of the interviewee to provide answers. For all 10 of the interviews, interviewers indicated that they were "completed with little or no missing information." For 5 interviews, the recorded times for start and stop were clearly incorrect or, in themselves, misleading. For example, three interviews involved negative or zero times (i.e., 070043B, 620250A, 020023B). This minor administrative error was not considered important enough to include in our standard review of data from the body of the interview that was completed as soon as the instrument was received at the HSP Coordinating Center. Two interviews were started one day and completed the next (710599C, 710003C), a fact that would have been missed without checking dates on the interview forms and which may explain times which appeared negative or short. In summary, the petitioner allegations regarding the quality of HSP interviews is misleading and not consistent with the facts that emerge from a more complete review of the interview forms.

We believe the company has reversed an obvious causal pathway. The company alleges that controls were not probed for "critical exposures on medications" because their interview duration was short. In fact the reverse is true. For control subjects who were healthy and not exposed to many medications, the interview was necessarily shorter than for case subjects who were more often unhealthy or exposed.

**7.2 Matching was conducted incorrectly**

Wyeth argues that the HSP investigators did not properly perform random digit dialing. Specific criticisms were that the investigators did not call back patients who met the study criteria and were willing to participate, identified control subject who were more likely to remain at home during the day, and matched cases and controls from different parts of the country. The company sites an overall participation rate among eligible controls of 18.5% compared with 65%-70% which it cites as the "current standard for acceptance of a case-control study by a medical journal."

Wyeth has mischaracterized the control matching procedure used by the HSP investigators and failed to represent the important design features that helped assure that control subjects were drawn from the same underlying population as case subjects.

Once a case subject was identified, telephone operators in New Haven, CT began making calls from a computer generated list of phone numbers. Because control subjects had to be interviewed within 30 days after the patient's stroke (to control for season), operators often had only a few days to identify a control that was matched to the case subject for telephone exchange, sex, race, and age (within 5 years). The need to identify a control subject within 30 days of the case subject's stroke limited the number of times we could attempt any one number. On average, we called 47 numbers to identify one potential control subject who was matched to a case subject for race, gender, and age.

Several attempts were typically made to reach parties at non-responding numbers, although the mean number of attempts is unknown because only the last outcome of multiple call attempts on one phone number was entered into the database. Among those who were eligible, 573 refused the screening interview and 256 did the screening interview but refused to participate. Thus the overall refusal rate was 39.6% (829/2096). Excluding 12 potential control subjects who reported a history of stroke, the refusal rate was 39.8% (829/2084). The reason for refusal was not entered into the HSP database.

The petitioner alleges that the HSP investigators "regularly matched cases with controls from different parts of the United States." This allegation amounts to a gross misstatement. Standard procedure for the HSP was to match control subjects to case subjects on residential telephone exchange (area code and first three digits of the seven digit local number). In creating its allegation, the petitioner analyzed zip code data for HSP cases and controls (Tab 54 of the Citizen Petition). The resulting document reveals 3 instances in which cases and control appear to list zip codes from different parts of the country.

We examined each of these three instances. We also searched the computerized HSP database for others and found four more. The two cases specifically cited by the petitioner (460058, 620047) did not, according to HSP records, involve a geographic mismatch. Case 460058 reported a residential address in Indiana. One matched control subject was from the same city in Indiana as the case subject and the other matched control was from another city in Indiana. Case 620047 lived and worked in Rhode Island. Two control subjects were matched to this case on area code and telephone exchange. Both control subjects reported residences in Rhode Island. In makings its allegations of mismatching for these two cases, the petitioner relied on zip code data that were miscoded on the original data forms.

When we examined the case report forms for the other 4 instances of zip code mismatch, none proved to represent actual geographic mismatch between case and control subjects; street addresses and telephone numbers provided by the subjects indicated excellent geographic matching. In one instance (710336), zip code mismatch was the result of a single-digit data entry error. In another instance (710154), mismatch was the result of an obvious error by the control subject (the correct zip code was provided with his full address, but reported incorrectly during later questioning). In two others (460122, 710129), the subjects' provided zip codes that were not consistent with their stated street addresses and telephone numbers.

Overall, the petitioner's allegation regarding geographic mismatching amounts to misinformation and does not help inform a careful assessment of the HSP protocol or its implementation.

**7.3 Interviewers were not randomly assigned**

Wyeth is correct that interviewers were not randomly assigned. Instead, they were assigned based on availability and convenience which we felt was adequate to assure data quality. All interviewers were trained in the administration of a highly structured, brief interview that was designed to minimize discretion or potential bias in the ascertainment of clinical information. Contrary to the Wyeth's suggestion, all interviewers interviewed both control and case subjects and had equal understanding of the study design, purpose, and procedures.

**7.4 The study was not population-based**

Wyeth is critical of the HSP for sampling case-subjects from selected hospitals rather than defined geographic areas. Suggesting that the sampling strategy was "nonrandom or selective," Wyeth argues that the HSP findings for an association between PPA and increased risk for hemorrhagic stroke were incorrect. Wyeth specifically suggests that physicians referred case subjects to the HSP based on exposure to PPA.

The sampling strategy in HSP was unbiased. It was comprehensive within participating hospitals and included features to minimize referral or selection bias. Our strategy for sampling cases from regional hospitals is commonly used in case control studies[14-16]. Our methods helped assure the internal validity of our results. While it is possible that our results may not apply beyond the hospitals involved in this study, we believe they are probably generalizeable. A recent case-control study from Korea used methods very similar to the HSP and also found an association between PPA and increased risk for hemorrhagic stroke[1].

At each participating research site, we attempted to reduce referral bias by requiring that investigators identify potentially eligible patients by active surveillance in which the investigators seek and find case subjects. Methods might include review of emergency room admission lists, service rosters, or daily interviews with clinicians who would know of all admissions for hemorrhagic stroke. A total of 14 cases were identified by direct referral before they could be identified by the HSP investigators. We do not know how many of the direct-referral cases would have been subsequently identified by active surveillance, but it is likely that all or most of them would have been. All of the direct referral patients were admitted to institutions at which we had active surveillance systems in place.

If treating clinicians selectively referred case subjects on the basis of exposure to PPA-containing products, it would be reasonable to expect such exposure to be documented in the medical record. In an attempt to understand the potential role of selective referral, we examined the medical records for all 27 PPA-exposed case subjects for documentation by treating clinicians of exposure to any cough and cold or diet remedy. We found 9 case subjects for whom such an exposure was recorded; all of the exposures were to cough and cold products. This finding provides further evidence that our main finding, which related to diet remedies, was not explained by referral bias.

In an analysis that excludes patients who were directly referred (shown in Table 4), the odds ratio for the association between PPA use in appetite suppressants and risk for hemorrhagic stroke in women becomes 16.31 ($p=0.02$), compared with 16.58 ($p=0.02$) in the original publication[13].

| Table 4. Comparison of the original analytic results for the HSP with an analysis excluding patients who were directly referred by clinicians to the HSP investigators and an analysis excluding 99 patients from the Rhode Island research site. | | | | |
|---|---|---|---|---|
| | Cough-cold in 3 days (all) | AS in 3 days (women) | Any PPA in 3 days (all) | 1st dose (women) |
| Original | 1.23 p=0.49 | 16.58 p=0.02 | 1.49 p=0.17 | 3.13 p=0.08 |
| W/o 14 directly referred cases and their 28 controls | 1.15 p=0.66 | 16.31 p=0.02 | 1.41 p=0.25 | 2.85 p=0.12 |
| W/o 99 RI cases and their 194 controls | 1.10 p=0.75 | 14.19 p=0.04 | 1.295 p=0.39 | 3.15 p=0.08 |

## 8. Other Allegations

### 8.1 Choice of one-tailed or two-tailed test of statistical significance

Wyeth is critical of the FDA for not recognizing the distinction between a one-tailed and two-tailed test of significance for the HSP in various communications. The sample size for the HSP was designed with a one-sided test of significance and we had intended to report the results using a one-sided test. One-tailed tests, in fact, were used for sample size estimation. Despite these facts, the petitioner is critical of the HSP investigators for using a one-tailed test of significance in its report to the FDA in May, 2000. In a surprising demonstration of bad faith and lack of regard for the facts, the petitioner accuses the HSP investigators of using one-sided testing in an effort "to establish a statistically significant relationship with first use." We used a one-sided test because it was specified in the protocol and was consistent with good science. The editors at the New England Journal of Medicine, however, asked us to report our results with a two-sided test in keeping with the journal's policy. Wyeth and others[17] do not quarrel with this change which caused the P value finding for the association between PPA and risk for hemorrhagic stroke in women under the first use definition to change from significant (P=0.042) to non significant (P=0.08). It also caused the P value for the association between PPA and risk for hemorrhagic stroke in men and women combined under the first use definition to change from significant (P=0.029) to non-significant (P=0.06).

Wyeth comments on a letter written by the HSP investigators to the New England Journal of Medicine in which we comment that the finding of an association between first use of products containing PPA among women and increased risk for hemorrhagic stroke did not reach conventional criteria for statistical significance (odds ratio 3.13, P=0.08, two-sided). Wyeth failed to fully capture our meaning when it truncated the full text of our comment. We wrote, "Although this odds ratio did not reach conventional criteria for statistical significance (P<0.05), this criterion itself may be too stringent for evaluating potentially harmful associations."

**8.2 Appropriateness of combining subarachnoid and intracerebral hemorrhage**

Wyeth questions the appropriateness of the HSP analyses plan which combines patients with SAH and ICH. The HSP analysis plan was approved at the start of the research by the FDA, industry, and the Scientific Advisory Group based partly on case reports indicating that PPA may be associated with both ICH[18-21] and SAH[22, 23]. By design, therefore, the HSP was not expected to have sufficient statistical power to detect an association at a conventional level of statistical significance (e.g., $P<0.05$) between consumption of PPA containing products and either stroke type separately.

To support its questioning, Wyeth cites a communication from one of the HSP investigators, Dr. Tom Brott, to another, Dr. Walter Kernan, objecting to the omission of an analysis according to stroke type in the main paper[13]. Wyeth failed to represent subsequent discussions among the investigators in which it was agreed to include a footnote in table 4 of the main publication regarding the distribution of stroke types among case subjects and to plan for subsequent publications on type-specific risk factors. The HSP was designed to combine SAH and ICH in the main analysis and this design feature was not seriously questioned at any time during the conduct of the study. To change the analytic plan at the conclusion of the study would have been scientifically inappropriate. The final HSP manuscript was approved by all authors.

Wyeth describes results from the two subsequent publications from the HSP dealing separately with aneurismal SAH and ICH[8, 9]. The company argues that "neither analysis found PPA to be a risk factor for hemorrhagic stroke." This interpretation, while technically correct, is deceptive and fails to account for significant differences in the design, purpose, and statistical power of the two papers on SAH and ICH compared with the main paper. The two type-specific papers only examined exposure to any PPA-containing produce within 3 days in men and women combined. Not surprisingly, one odds ratio (for aneurismal SAH) was lower than the odds ratio reported for the combined cohort (1.15 compared with 1.49) and one (for ICH) was higher (2.17 compared with 1.49). These papers on aneurismal SAH and ICH did not include analyses separately for men and women or by indication. In the combined cohort, a statistically significant association (two-sided P value<0.05) was only found for the use of PPA as an appetite suppressant in women[13].

Sincerely,

Walter N. Kernan, M. D.
Professor of Medicine
Yale University School of Medicine
On Behalf of the HSP Investigators

## REFERENCES


1. Yoon BW, Bae HJ, Hong KS, et al. Phenylpropanolamine contained in cold remedies and risk of hemorrhagic stroke. *Neurology.* 2007;68:146-149.
2. Ickovics JR, viscoli CM, Horwitz RI. Functional recovery after myocardial infarction in men: independent effects of social class. *Ann Intern Med.* 1997;127:518-525.
3. Hernan MA, Hernandez-Diaz S, Werler MM, et al. Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. *Am J Epidemiol.* 2002;155:176-184.
4. Karagas MR, Stukel TA, Dykes J, et al. A pooled analysis of 10 case-control studies of melanoma and oral contraceptive use. *Br J Cancer.* 2002;86:1085-1092.
5. Petitti DB, Sidney S, Bernstein A, et al. Stroke in users of low-dose oral contraceptives [see comments]. *N Engl J Med.* 1996;335(1):8-15.
6. Writing Committee, contraception WHOsocdash. Venous thromboembolic disease and combined oral contraceptives: results of international multicentre case-control study. *Lancet.* 1995;346:1575-1582.
7. Hosmer DW, Lemeshow S. *Applied Logistic Regression.* 2nd ed. New York: Wiley; 2000.
8. Broderick JP, Viscoli CM, Brott T, et al. Major risk factors for aneurysmal subarachnoid hemorrhage in the young are modifiable. *Stroke.* 2003;34:1375-1381.
9. Feldmann E, Broderick JP, Kernan WN, et al. Major risk factors for intracerebral hemorrhage in the young are modifiable. *Stroke.* 2005;36:1881-1885.
10. Sloan MA, Kittner MA, Rigamonti D, et al. Occurrence of stroke associated with use/abuse of drugs. *Neurology.* 1991;41:1358-1364.
11. Humberstone PM. Hypertension and cold remedies. *Br Med J.* 1969;i:846.
12. Chung Y-T, Hung D-Z, Hsu, C-P, Yang, D-Y, Wu, T-C. Intracerebral hemorrhage in a young woman with arteriovenous malformation after taking diet control pills containing phenylpropanolamine: A case report. *Chin Med J (Engl).* 1998;61:432-435.
13. Kernan WN, Viscoli CM, Brass LM, et al. Phenylpropanolamine and the risk of hemorrhagic stroke. *The New England Journal of Medicine.* 2000;343:1826-1832.
14. Heinemann LAJ, Lewis MA, Thorogood M, et al. Case-control study of oral contraceptives and risk of thromboembolic stroke: results from international study on oral contraceptives and health of young women. *BMJ.* 1997;315:1502-1504.
15. Who Collaborative Study of Cardiovascular Disease and Steroid Hormone Contraception. Ischaemic stroke and combined oral contraceptives: results of an international, multicentre, case-control study. *Lancet.* 1996;348:498-505.
16. Forsyth BW, Horwitz RI, Acampora D, et al. New epidemiologic evidence confirming that bias does not explain the aspirin/Reye's Syndrome association. *Journal of the American Medical Association.* 1989;261:2517-2524.
17. Stier BG, Hennekens CH. Phenylpropanolamine and Hemorrhagic Stroke in the Hemorrhagic Stroke Project: A reappraisal in the context of science, the Food and Drug Administration, and the law. *Ann Epidemiol.* 2005;16:49-52.
18. Kase CS, Foster TE, Reed JE, et al. Intracerebral hemorrhage and phenylpropanolamine use. *Neurology.* 1987;37:399-404.
19. Maher LM. Postpartum intracranial hemorrhage and phenylpropanolamine use. *Neurology.* 1987;37:1686.
20. Forman HP, Levin S, Stewart B, et al. Cerebral vasculitis and hemorrhage in an adolescent taking diet pills containing phenylpropanolamine: case report and review of literature. *Pediatrics.* 1989;83:737-741.
21. McDowell JR, LeBlanc HJ. Phenylpropanolamine and cerebral hemorrhage. *West J Med.* 1985;142:688-691.
22. Bernstein E, Diskant BM. Phenylpropanolamine: a potentially hazardous drug. *Ann Emerg Med.* 1982;11:311-315.
23. Kikta DG, Devereaux MW, Chandar K. Intracranial hemorrhages due to phenylpropanolamine. *Stroke.* 1985;16:510-512.