DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE

Food and Drug Administration

[Docket No. 76N-0433]

Warner-Lambert/Parke-Davis & Co.; Benylin; Final Decision

AGENCY: Food and Drug Administration.

ACTION: Final Decision following a formal evidentiary public hearing in an adjudicatory proceeding.

SUMMARY: The Commissioner of Food and Drugs is issuing his Final Decision concerning a supplemental new drug application for Benylin (diphenhydramine hydrochloride) a drug which the sponsor, Warner-Lambert/Parke-Davis & Co., claims is indicated for use in the treatment of cough due to colds or inhaled irritants. The Commissioner has determined that Benylin has not been shown to be effective for this use and is refusing to approve the application. In view of the decision on effectiveness, the Commissioner has not decided whether Benylin is safe for over-the-counter (OTC) distribution. The Decision reverses the Initial Decision of the Administrative Law Judge, which found that Benylin is effective for its recommended use and is safe for OTC distribution.

EFFECTIVE DATE: June 29, 1979.

ADDRESS: The transcript of the hearing, evidence submitted, and all other documents cited in this decision may be seen in the office of the Hearing Clerk (HFA-305), Food and Drug Administration, Rm. 4-65, 5600 Fishers Lane, Rockville, MD 20857, from 9 a.m. to 4 p.m., Monday through Friday.

FOR FURTHER INFORMATION CONTACT: Tenny P. Neprud, Compliance Regulations Policy Staff (HFC-10), Food and Drug Administration, Department of Health, Education, and Welfare, 5600 Fishers Lane, Rockville, MD 20857, 301-443-3480.

SUPPLEMENTARY INFORMATION: The purpose of this proceeding has been to determine whether the Food and Drug Administration (FDA) should approve a supplemental new drug application (NDA 6-514 S-007) for the over-the-counter (OTC) marketing of Benylin Expectorant (Benylin) as an antitussive drug (a drug which specifically inhibits or suppresses coughs).

Benylin is a liquid preparation containing, among other ingredients, diphenhydramine hydrochloride, ammonium chloride, and sodium citrate. Warner-Lambert/Parke-Davis (WL/PD), the sponsor of the supplemental new drug application (NDA), holds an approved NDA for prescription distribution of Benylin as an antitussive drug. WL/PD markets two other well known products, Benadryl and Benylin DM, which should not be confused with the product that is the subject of this proceeding. Benadryl is a prescription antihistamine drug with several other uses; it has diphenhydramine hydrochloride as its active ingredient. The recommended adult antihistiminic dosage of Benadryl provides 50 mg diphenhydramine hydrochloride every 4 hours. The recommended adult antitussive dosage of Benylin provides 25 mg diphenhydramine every 4 hours. Benylin DM is an OTC antitussive drug having dextromethorphan, rather than diphenhydramine hydrochloride, as its active ingredient. Benadryl has been on the market since the 1940's; marketing of Benylin DM began during this proceeding.

The effect of this Decision is to refuse to approve the supplemental NDA for marketing of Benylin as an OTC antitussive drug. This Decision leaves unaffected the approved NDA for marketing of Benylin as a prescription antitussive. However, the finding that Benylin has not been shown to be effective for its recommended use is equally applicable whether the drug is marketed OTC or is subject to prescription requirement. Accordingly, in light of this finding, the Bureau of Drugs (the Bureau) should consider whatever action seems appropriate with respect to the approved NDA for prescription Benylin.

In a notice published in the **Federal Register** of November 30, 1976 (41 FR 52537), the Bureau proposed to deny approval of the supplemental NDA for the OTC marketing of Benylin as an antitussive drug on two grounds:

1. That diphenhydramine hydrochloride, one of the ingredients in Benylin, causes a level of drowsiness in those who take it that is sufficient to render it unsafe for use except under the supervision of a physician or other practitioner licensed to dispense prescription drugs.

2. That the studies submitted to establish the effectiveness of Benylin as an antitussive drug do not provide substantial evidence of its effectiveness for that use as required by section 505(d) of the Federal Food, Drug, and Cosmetic Act (the act) (21 U.S.C. 355(d)) and § 314.111(a)(5) of the regulations (21 CFR 314.111(a)(5)).

I. Background

A more detailed history of the regulation of Benylin and of this proceeding is found in the Bureau's November 30, 1976 proposal to deny approval of the supplemental NDA. The original NDA, approved in 1948, was for use of the drug as an expectorant (a drug that makes it easier for a patient to raise secretions from the chest and throat). Restriction to prescription was a condition of the approval of the drug, then called "Benylin Expectorant."

In 1966, FDA permitted the inclusion of an antitrussive indication in the labeling. During the same year, FDA began a review of the effectiveness of all drugs that had been approved before the Drug Amendments of 1962 added to the act the requirement that the sponsor of an NDA show the new drug to be effective, as well as safe. This review, carried out in cooperation with the National Academy of Sciences/National Research Council (NAS/NRC), resulted in a 1971 NAS/NRC report and an FDA notice in the **Federal Register** of June 18, 1971 (36 FR 11758), classifying diphenhydramine hydrochloride as "lacking substantial evidence of effectiveness" as an antitussive. In a notice published in the **Federal Register** of February 9, 1973 (38 FR 4006), the Bureau announced its conclusion that there is a lack of substantial evidence of the effectiveness of Benylin Expectorant and certain other products as fixed combinations for the indications in their labeling, and offered an opportunity for a hearing on a proposal to withdraw approval of the NDA for Benylin Expectorant.

By letter of March 9, 1973, WL/PD requested a hearing on the proposed withdrawal. In its request, the firm referred to the submission it had filed for review by the Advisory Review Panel on OTC Cold, Cough, Allergy, Bronchodilator, and Antiasthmatic Drugs (CCABA Panel). The CCABA Panel had been established as part of FDA's program for review of the safety and effectiveness of all OTC drugs. The firm also responded to the notice of opportunity for a hearing by filing, on March 22, 1973, a supplemental NDA for Benylin Expectorant. The supplemental NDA provided for changing the name of the product to "Benylin Cough Syrup," for deletion of all ingredients except diphenhydramine hydrochloride from from the list of active ingredients, but not from the product formulation, and for a change in the description of the product's mechanism of action and in the indications for which the product was recommended. By letter of November 28, 1973, the Bureau refused to approve the supplemental NDA.

On February 5, 1974, WL/PD also submitted a supplemental NDA with

two studies relating to the effectiveness of Benylin as an antitussive. On November 25, 1974, the firm submitted a supplemental NDA with revised labeling providing for OTC use of Benylin as an antitussive. The later supplemental NDA is the subject of this Decision. By letter of March 11, 1975, the Bureau informed the firm that no action would be taken on the two supplemental NDA's pending completion of review by the CCABA Panel of the data before it.

In response to a letter from WL/PD, FDA's Division of OTC Drug Evaluation informed the firm on March 18, 1975, that OTC marketing of Benylin would be unlikely to be subject to regulatory action under the enforcement policy then in effect comcerning new OTC products. Thereafter, WL/PD commenced OTC marketing of Benylin with indications for use as an antitussive.

On August 4, 1976 (41 FR 32580), however, FDA published a final regulation, based on a proposal published on December 4, 1975 (40 FR 56675), that changed FDA's enforcement policy concerning OTC marketing of drug ingredients that had previously been limited to prescription use and for which OTC use had not been approved by FDA. This regulation, codified in §§ 310.200 and 330.13 (21 CFR 310.200, 330.13) allows the OTC marketing of products containing such ingredients upon publication of the report of an OTC advisory review panel recommending that the ingredients and indications be classified as generally recognized as safe and effective for OTC use (Category I) unless the Commissioner of Food and Drugs (the Commissioner) disagrees with that decision.

In the Federal Register of September 9, 1976 (41 FR 38312), the Commissioner published a proposal setting forth the report and recommendations of the CCABA panel. The Panel recommended that diphenhydramine hydrochloride be classified in Category I for OTC use as an antitussive. In the preamble to the proposal, the Commissioner stated that his decision concerning the safety and effectiveness of diphenhydramine hydrochloride as an antitussive would be made in the context of his ruling on the supplementyal NDA for OTC marketing of Benylin.

On September 8, 1976, the Bureau informed WL/PD that its supplemental NDA for Benylin as an OTC antitussive was not approvable. On September 17, 1976, the firm requested that the supplemental NDA for OTC use of Benylin be filed over protest pursuant to § 314.110(d) (21 CFR 314.110(d)).

In a notice published in the Federal Register of November 30, 1976 (41 FR 52537), the Bureau proposed to deny approval of the supplemental NDA for OTC marketing of Benylin Expectorant and provided WL/PD with notice of opportunity for a hearing on this proposed action. On the same date, the Commissioner published a notice (41 FR 52536) announcing that he did not, at that time, accept the CCABA Panel's recommendation that diphenhydramine hydrochloride be classified in Category I for OTC antitussive use. Accordingly, any OTC product marketed containing diphenhydramine hydrochloride was subject to immediate regulatory action. The Commissioner had concluded that the recommended antitussive dose of diphenhydramine hydrochloride (25 mg) causes an unacceptable level of drowsiness for an OTC drug. Furthermore, although he agreed with the Panel that some data indicated that this ingredient has some antitussive effect, he found that there was a lack of substantial evidence consisting of adequate and well-controlled studies, as required by § 314.111(a)(5)(ii) (21 CFR 314.111(a)(5)(ii)), on which to base a determination concerning the effectiveness of Benylin for the temporary control of cough.

On November 29, 1976, the firm filed an action seeking a declaratory judgment that Benylin is not a new drug or, in the alternative, an order enjoining FDA enforcement actions involving Benylin pending final determination of the drug's status (Civil Action No. 6–72464, E.D. MI). On November 30 and December 1, 1976, three United States Attorneys for other districts filed complaints resulting in seizures of Benylin. The Michigan case ultimately resulted in a decision by the United States Court of Appeals for the Sixth Circuit holding that the district court lacked jurisdiction to review the agency's decision to initiate enforcement action and that the pending enforcement actions provided an opportunity for full hearing on all issues. *Parke, Davis & Co.* v. *Califano*, 564 F.2d 1200 (6th Cir. 1977), *rev'g Parke, Davis & Co.* v. *Mathews*, Civil Action No. 6–72464 (E.D. MI, Memorandum Opinion issued Jan. 7, 1977).

WL/PD then submitted a request for hearing, which the Commissioner granted in a notice published in the Federal Register of March 29, 1977 (42 FR 16675). The proceeding was referred to the presiding officer, Administrative Law Judge Daniel J. Davidson.

The March 29, 1977, notice of hearing observed that the issue of the effectiveness of Benylin as an OTC product is indistinguishable from the issue of its effectiveness as a prescription product and therefore announced that the hearing would concern the effectiveness of Benylin for prescription use as well as for OTC use. For this reason, the Administrative Law Judge (ALJ) in a pretrial order dated June 2, 1977, broadened the issues at the hearing to include consideration of "[w]hether there is any other evidence relating to the effectiveness of Benylin as an antitussive." Although the effectiveness of Benylin as a prescription antitussive is required to be resolved in a separate withdrawal proceeding, "there would be no need to duplicate the hearing process with respect to other evidence relating prescription use" (ID,[1] p. 4).

As revised by the ALJ, the issues at the hearing were:

1. Whether the NDA for Benylin contains reports of investigations or other information adequate to demonstrate the safety of the drug for OTC distribution, as required by 21 U.S.C. 355(d)(1), (2), and (4);

2. Whether the NDA for Benylin demonstrates that there is substantial evidence of effectiveness of the drug as an antitussive in the form of adequate and well-controlled clinical investigations, as defined by § 314.111(a)(5) (21 CFR 314.111(a)(5)), on the basis of which it could fairly and responsibly be concluded by experts, qualified by scientific training and experience to evaluate drugs, that Benylin is effective as an antitussive for OTC use;

3. Whether there is any other evidence relating to the effectiveness of Benylin as an antitussive;

4. Whether Benylin is generally recognized, among experts qualified by scientific training and experience to evaluate the safety and effectiveness of drugs, as safe and effective for use as an OTC antitussive under the conditions prescribed, recommended, or suggested in its labeling.

The oral portion of the hearing was held from October 11 through 25, 1977. The parties were the Bureau, in support of the proposed denial, and WL/PD, in opposition to the proposed denial. Both parties filed briefs.

[1] The following abbreviations have been used in citing material in the record: Initial Decision: ID; Transcript: Tr; Brief of a participant to the ALJ: Brief; the Bureau's exceptions to Initial Decision: Exceptions; WL/PD's reply to the Bureau's exceptions to Initial Decision: Reply. This decision refers to the exhibits submitted to the record, including written direct testimony, by the symbol for the participant and number assigned to them upon filing by the Hearing Clerk. The Hearing Clerk used the following symbols to refer to the exhibits by the participants: WL/PD: P; Bureau: G.

II. Alleged Deficiencies in Clinical Studies of Benylin's Effectiveness

In an appendix to the Initial Decision, the ALJ summarized the major clinical studies submitted by WL/PD to show that Benylin is an effective antitussive agent. For convenience, excerpts from that summary are set forth as Appendix A to this Decision, and the Bureau's allegations concerning deficiencies in these studies are listed below.

The Bureau contends that studies of Benylin in groups other than the target population may be considered only if diphenhydramine hydrochloride is shown conclusively to act specifically on the cough center of the brain, and not through some other mechanism such as soothing the throat or causing general sedation (G–11, pp. 11–13; G–29, pp. 2–3). The Bureau also asserts that it has not been shown that this drug acts specifically on the cough center (G–1 pp. 8, 14–15; G–11, pp. 12–13; G–29, pp. 2, 7; T–529, T–961). The Bureau contends that Benylin has not been shown to be effective as an antitussive because the two submitted studies of Benylin in the drug's target population contain deficiencies that prevent the studies from being considered adequate and well-controlled investigations meeting the requirements of 21 CFR 314.111(a)(5)(ii).

The Bureau assets that the Tebrock study (Protocol 266–17, P9–1 through P9–6) is deficient in the following respects:

(a) Both a positive control and a placebo control should have been used (Brief, pp. 98–108; G–11, p. 5; G–13, pp. 6–7; G–17, p. 14; G–27, pp. 6–7; G–138, p. 465).

(b) The study lacks an adequate placebo control in that the Benylin vehicle contains ingredients, ammonium chloride and sodium citrate, that may be active (Brief, pp. 24–52; G–11, pp. 8–10, 18, 20; G–13, pp. 5–6; G–15, pp. 7–8; G–27, pp. 18–19).

(c) The subjective measurement of cough improvement used in the study does not permit quantitative evaluation of test results as required by 21 CFR 314.111(a)(5)(ii)(*a*)(*4*); the sponsor should have measured cough improvement by objective means, e.g., having subjects keep detailed diaries throughout the day, rather than rate their cough improvement once daily on a 4-point scale (Brief pp. 82–92; G–11, p. 21; G–17, pp. 12–13; T–564).

(d) Having subjects evaluate their cough improvement by comparing their cough experience on one day to that on the previous day does not permit quantitative evaluation and statistical analysis of the results of treatment and control as required by 21 CFR 314.111(a)(5)(ii)(*a*) (*4*) and (*5*); the protocol should have had subjects compare their coughs after treatment to their coughs on the day immediately preceding the 3-day trial (G–13, pp. 8, 9; G–7, p. 4; G–17, p. 11).

(e) Data from the 5 test centers should not have been pooled because homogeneity of treatment effect among centers is a prerequisite to such pooling, and statistical tests show lack of homogeneity (Brief, pp. 52–72; G–7, pp. 7–8; G–13, p. 9).

(f) The results of the study are not clinically significant; an improvement in cough of only 9 percent in the Benylin group compared to the Benylin vehicle group is not sufficient to satisfy the requirement of substantial evidence of effectiveness (Brief, pp. 73–81; G–11, pp. 20–21; G–13, pp. 10–11; G–15, pp. 9–10; G–27, p. 19; T–560).

The Bureau contends that the Burke study (Protocol 266–9 P7–1 through P7–5) is deficient in the following respects:

(a) Both a positive control and a placebo control should have been used (Brief, pp. 98–107; G–11, p. 5; G–13, p. 7; G–27, pp. 6–7; G–138, p. 465).

(b) Because the effectiveness of a combination of Benylin and codeine has not been established, the use of this combination is not an "effective regimen of therapy," i.e., a positive control, meeting the requirements of 21 CFR 314.111(a)(5)(ii)(*a*)(*4*)(*iii*) (Brief, pp. 137–143; G–11, p. 19; G–13, p. 11; G–15, pp. 10–11; G–17, p. 11; G–27, p. 17).

(c) The study results do not permit quantitative evaluation because the study relied on subjective evaluations by children, most of whom were between the ages of 6 and 12; children of this age group are unreliable for the subjective evaluation of an antitussive drug when the evaluation requires a single judgment of the effects of the drug over 3 to 4 days (Brief at 147–153; G–3, pp. 8–9; G–11, p. 19; G–31, p. 17).

III. Initial Decision

On May 31, 1978, the ALJ issued an Initial Decision in which he found that Benylin had been shown, by adequate investigations, to be safe and effective for use as an OTC antitussive and ordered that the supplemental NDA (6–514/S–007) be approved.

A. Specific Findings

Specifically, the ALJ found (ID, pp. 21–22):

(1) The evidence of record in this proceeding is insufficient to determine whether Benylin, in the opinion of the medical community, is generally recognized, among experts qualified by scientific training and experience to evaluate the safety and effectiveness of drugs, as safe and effective for use as an OTC antitussive under the conditions prescribed, recommended, or suggested in its labeling. Therefore, the exemption from new drug status * * * has not been shown to apply to Benylin.

(2) The supplemental NDA for Benylin (NDA 6/514/S–007) contains reports of investigations or other information adequate to demonstrate the safety of the drug for OTC distribution as required by 21 U.S.C. 355(d) (1), (2) and (4).

(3) The supplemental NDA for Benylin (NDA 6–514/S–007) demonstrates that there is substantial evidence of effectiveness of the drug as an antitussive in the form of adequate and well-controlled clinical investigations, as defined by § 314.111(a)(5) [21 CFR 314.111(a)(5)], on the basis of which it could fairly and responsibly be concluded by experts, qualified by scientific training and experience to evaluate drugs, that Benylin is effective as an antitussive for OTC use.

(4) There is no other evidence relating to the effectiveness of Benylin as an antitussive for prescription use.

Also included in the initial Decision was a detailed discussion of the ALJ's findings with respect to the safety and effectiveness of Benylin. (To conform to the format elsewhere in this Decision, the summary below of the Initial Decision first discusses effectiveness, then safety.)

B. Effectiveness

The ALJ considered only the Tebrock study (Protocol 266–17, P9–1 through P9–6) and the Burke study (Protocol 266–9, P9–1 through P9–6), the two studies in patients with cough due to cold, in support of Benylin's claimed effectiveness in cough due to cold. He rejected all of the Bureau's contentions with respect to alleged deficiencies in these studies. He found with respect to both studies that there is no requirement in the regulations that both a positive and a negative control be used (ID, pp. 9, 20).

With respect to the Tebrock study, the ALJ also found:

(a) The Benylin vehicle was a suitable placebo; the Bureau had failed to show that ammonium chloride or sodium citrate possesses activity (ID, pp. 10–11, 20).

(b) The subjective measurement of cough improvement used in the study was acceptable (ID, pp. 11–12, 20–21);

(c) Having subjects compare cough improvement by comparing experience on one day with that on the previous day is acceptable, given the study results (ID, p. 12);

(d) It was appropriate for WL/PD to pool data from 5 test centers (ID, pp. 12–14);

(e) WL/PD had met the statutory requirement of proof of effectiveness, because the results of this study showed that Benylin offers a statistically

significant benefit over the control and these results were corroborated by the Burke study, which the ALJ interpreted as showing Benylin to be 8 percent more effective than codeine [2] (ID. pp. 14–15, 21).

With respect to the Burke study, the ALJ found that:

(a) The study was adequately controlled (ID, pp. 11, 20, 21), and

(b) The study results were susceptible to quantitative evaluation despite the reliance upon subjective evaluations by young children (ID, p. 12).

The ALJ considered two of the induced cough studies (Protocols 266–18, P6–6 through P6–9; and 794–1, P6–10 through P6–14) as substantial evidence of the effectiveness of Benylin for cough due to inhaled irritants (ID, pp. 16–19). He also accepted a third study (the 1960 Bickerman Protocol, P6–3 through P6–5) as corroborative evidence of the effectiveness of Benylin for this use (ID, p. 18). He rejected the Bureau's contentions that these studies qualify only as Phase I studies of the pharmacological action of Benylin, that there is no target population having cough due to inhaled irritants, that the numbers of subjects were too small, that the studies are not adequate and well-controlled, that cough reduction in Protocol 266–18 may have been due to an anesthetic effect from gargling rather than antitussive activity, and that Protocol 794–1 did not test the finished drug product.

C. Safety

The ALJ found that the supplemental NDA for Benylin contains reports of investigations or other information adequate to demonstrate the safety of the drug for OTC distribution as required by section 505(d) of the act (ID, pp. 9, 22). (He did not, however, mention any particular studies as satisfying this requirement.) The ALJ cited the extensive use of Benylin over a period of 30 years and its approval for prescription use as "a substantial indication that it can be used safely" (ID, p. 5).

The ALJ also stated that the Bureau's argument that limiting Benylin to prescription use would reduce the chance of the drug causing injury "does not take into account the limited scope of agency authority" under section 503(b) of the act to restrict drugs to prescription, and under section 505(d) of the act to consider use of drugs contrary to label provisions (ID, pp. 6–8). He stated his belief that drowsiness due to Benylin while driving a motor vehicle is the only hazard that the Bureau contends is associated with the drug (ID, p. 6). He pointed out that this drowsiness presents a hazard only when a patient ignores the proposed lable warning against use of Benylin while driving a motor vehicle (ID, pp. 6–8). He considered this use to be outside the scope of the provision that the safety of a drug is to be evaluated in terms of its use "under the conditions prescribed, recommended, or suggested in the proposed labeling thereof" (ID, p. 7). The ALJ held that "questions of the safety of a drug when it is used in a manner contrary to the label warnings can be considered in only limited circumstances" (ID, p. 8). Once the manufacturer has shown a drug to be safe under the conditions prescribed, recommended, or suggested in the labeling, the drug is approvable unless the Bureau can show, by "convincing evidence of harm," that "there is a reasonable probability that such non-label indicated uses can be expected to occur" before the Bureau can deny approval of a drug because of hazards when used contrary to label warnings (ID, p. 8). Finding that the Bureau had made no such showing in the record of this case, the ALJ found the safety record of Benylin during 30 years of its use to be "controlling in this proceeding" (ID, pp. 8–9).

The Bureau filed exceptions to the Initial Decision and appealed it to the Commissioner. WL/PD filed a reply disagreeing with each of the Bureau's exceptions and urging that the Initial Decision be affirmed.

IV. Statute and Regulations

A. Effectiveness

The act requires that a sponsor show a new drug to be effective before FDA may approve it for marketing. Section 505(d) of the act requires the Commissioner (by delegation from the Secretary, 21 CFR 5.1) to issue an order refusing to approve an NDA if the Commissioner finds, after notice and opportunity for a hearing, that:

(5) evaluated on the basis of the information submitted to him as part of the application and any other information before him with respect to such drug, there is a lack of substantial evidence that the drug will have the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the proposed labeling thereof * * *[T]he term "substantial evidence" means evidence consisting of adequate and well-controlled investigations, including clinical investigations, by experts qualified by scientific training and experience to evaluate the effectiveness of the drug involved, on the basis of which it could fairly and responsibly be concluded by such experts that the drug will have the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the labeling or proposed labeling thereof.

FDA has promulgated regulations (21 CFR 314.111(a)(5)(ii)) that set forth the principles concerning the conduct of adequate and well-controlled clinical investigations and that provide the basis for determining whether there is "substantial evidence" to support claims of effectiveness for new drugs.

Also relevant to this proceeding are the regulations describing the principles to be used to determine general recognition that an OTC drug is effective (21 CFR 330.10(a)(4)(ii)):

(ii) Effectiveness means a reasonable expectation that, in a significant proportion of the target population, the pharmacological effect of the drug, when used under adequate directions for use and warnings against unsafe use, will provide clinically significant relief of the type claimed. Proof of effectiveness shall consist of controlled clinical investigations as defined in § 314.111 (a)(5)(ii) of this chapter, unless this requirement is waived on the basis of a showing that is not reasonably applicable to the drug or essential to the validity of the investigation and that an alternative method of investigation is adequate to substantiate effectiveness. Investigations may be corroborated by partially controlled or uncontrolled studies, documented clinical studies by qualified experts, and reports of significant human experience during marketing. Isolated case reports, random experience, and reports lacking the details which permit scientific evaluation will not be considered. General recognition of effectiveness shall ordinarily be based upon published studies which may be corroborated by unpublished studies and other data.

B. Safety

The act requires that a sponsor show a new drug to be safe before FDA may approve it for marketing. Section 505(d) of the act (21 U.S.C. 355(d)) requires the Commissioner to issue an order refusing to approve an NDA if the Commissioner finds, after notice and opportunity for a hearing,

That (1) the investigations, reports of which are required to be submitted * * *, do not include adequate tests by all methods reasonably applicable to show whether or not [the new] drug is safe for use under the conditions prescribed, recommended, or suggested in the proposed labeling thereof; (2) the results of such tests show that such drug is unsafe for use under such conditions or do not show that such drug is safe for use under such conditions; * * * (4) upon the basis of any other

[2] As noted correctly in the summary of the Burke study in the appendix to the Initial Decision, Benylin was compared to a combination of Benylin plus 20 mg of codeine. However, elsewhere in the Initial Decision it is stated, incorrectly, that the Burke study compared Benylin to a combination of the Benylin vehicle plus codeine (ID, pp. 10, 18), or to codeine (ID, pp. 15, 20, 21).

information before him with respect to such drug, he has insufficient information to determine whether such drug is safe for use under such conditions * * *.

Because the safety issue in this proceeding concerns whether Benylin is safe for OTC distribution, section 503(b)(1) of the act (21 U.S.C. 353(b)(1)) defining the criteria for restricting drugs to prescription use also is important in this proceeding. Section 503(b)(1) states:

A drug intended for use by man which—
(A) Is a habit-forming drug to which section 502(d) applies; or
(B) Because of its toxicity or other potentiality for harmful effect, or the method of its use, or the collateral measures necessary to its use, is not safe for use except under the supervision of a practitioner licensed by law to administer such drug; or
(C) Is limited by an approved application under section 505 to use under the professional supervision of a practitioner licensed by law to administer such drug; shall be dispensed only (i) upon a prescription of a practitioner licensed by law to administer such drug * * *.

The agency's authority to restrict drugs to prescription did not originate with the prescription drug provisions set forth above, which were added by the Durham-Humphrey Amendments of 1951 (Pub. L. 82–215). Before the enactment of these Amendments, FDA had issued regulations restricting certain drugs to prescription. These regulations were upheld in *United States* v. *Sullivan*, 332 U.S. 689 (1948) and *United States* v. *El-O-Pathic Pharmacy*, 192 F.2d 62 (9th Cir. 1951). The House Report on the 1951 Amendments explained the legal basis for the regulations:

The Present law prohibits the over-the-counter sale of drugs labeled as being for prescription only, but it does this in the following indirect manner.
A drug is required by present law to bear adequate directions for use. This requirement may be relaxed by the Federal Security Administrator [predecessor of the Commissioner] when such directions are not necessary for the protection of the public health. Drugs suitable for use only by or under the order of a licensed practitioner have been exempted from the adequate directions requirement on condition that they be labeled "Caution—To be dispensed only by or on prescription of a physician." If a druggist just sells without a prescription a drug bearing this caution label, the drug is misbranded and the druggist violates the act.

H.R. Rept. No. 82–700, 82d Cong., 1st Sess. 4 (1951). The Durham-Humphrey Amendments' criteria for classifying a drug as a prescription drug were essentially the same as those contained in the FDA regulations. H.R. Rept. No. 82–700 at 11; S. Rept. No. 82–946, 82d Cong., 1st Sess. 4 (1951).

Also relevant to the agency's decision concerning the safety of Benylin as an OTC antitussive are the regulations setting forth the principles to be used to determine general recognition that a category of OTC drugs is safe (21 CFR 330.10(a)(4)(i)):

Safety means a low incidence of adverse reactions or significant side effects under adequate directions for use and warnings against unsafe use as well as low potential for harm which may result from abuse under conditions of widespread availability. Proof of safety shall consist of adequate tests by methods reasonably applicable to show the drug is safe under the prescribed, recommended, or suggested conditions of use. This proof shall include results of significant human experience during marketing. General recognition of safety shall ordinarily be based upon published studies which may be corroborated by unpublished studies and other data.

C. General Recognition of Safety and Effectiveness

Section 201(p)(1) of the act (21 U.S.C. 321(p)(1)) defines "new drug" as:

Any drug * * * the composition of which is such that such drug is not generally recognized, among experts qualified by scientific training and experience to evaluate the safety and effectiveness of drugs, as safe and effective for use under the conditions prescribed, recommended, or suggested in the labeling thereof * * *.

The applicant has the burden of proving that a drug is generally recognized as safe and effective for use under the conditions prescribed, recommended, or suggested in the labeling. See *Weinberger* v. *Hynson, Westcott, and Dunning*, 412 U.S. 609, 620–21, 628–31 (1973); 21 CFR 12.87(e). The regulations governing hearing procedures for proposed denials of new drug applications (21 CFR 314.200(e)) further state:

(1) A contention that a drug product is generally recognized as safe and effective within the meaning of section 201(p) of the act must be supported by submission of the same quantity and quality of scientific evidence as is required to obtain approval of a new drug application for the product * * *. General recognition of safety and effectiveness shall ordinarily be based upon published studies which may be corroborated by unpublished studies and other data and information.

The Supreme Court has held that "general recognition" of a drug's safety and effectiveness requires at least substantial evidence of effectiveness as required for approval of an NDA. *Weinberger* v. *Hynson, Westcott and Dunning, supra*, at 629.

A. Effectiveness

In reviewing the Initial Decision, I have all the powers that I would have in making the Initial Decision (21 CFR 12.130). The Initial Decision and this Decision must meet the requirements of 21 CFR 12.120 and 12.130.

After reviewing the record carefully, I reverse the Initial Decision. I find, based on the evidence in the record, that there is a lack of substantial evidence that Benylin will have the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the proposed labeling. Expert witnesses who testified in this proceeding came to the same conclusion (G–1, pp. 5–6, 18; G–11, pp. 30–31; G–13, p. 4; G–15, pp. 5–6, 19; G–27, pp. 4, 27–28).

1. *Which Studies May be Considered.* An understanding of the principles for study of drugs to suppress coughs requires, first, an understanding of why and how humans cough.

Cough is a protective reflex that occurs in the healthy as well as in the ill (G–30, p. 113; G–40, p. 443; G–138, p. 462; P14–1, p. 38321; P22–18, p. 395). The function of a cough is to clear the respiratory tract of secretions or inhaled irritants (id.). In many cases, it is unwise to suppress coughs (G–10, pp. 7–8; G–35). A variety of stimuli can cause cough, including mechanical, thermal, chemical, allergic, and toxic agents (G–38, p. 492; G–138, p. 471).

Coughing is normally produced by the stimulation of a patient's "cough receptors" (sensory endings of the glossopharyngeal and vagus nerves) in the patient's throat and respiratory passageway (G–40, p. 426; G–138, p. 461). The "cough center," located in the medulla of the brain, is largely responsible for control of the cough reflex in humans (G–1, p. 7; G–40, p. 428; P22–18, p. 395). The central nervous system pathways for coughing are not fully understood (G–40, p. 445; P22–18, p. 395).

The CCABA Panel described an antitussive agent as one that "specifically inhibits or suppresses the act of coughing" (P14–1, p. 38338). Drugs may inhibit cough in several ways: by depressing the cough center or higher centers of the brain by diminishing the sensitivity of the cough receptors; by interrupting the transmission of cough impulses to the brain or to the muscles that are involved in the act of coughing; or by removing irritants and excessive secretions through improved bronchial drainage (P14–1, p. 38338, P19–2, p. 16; G–1, p. 7, G–29, pp. 2–3; G–40, p. 446; G–138, p. 462; P22–18, p. 397).

Among the causes of cough are chronic conditions (such as chronic bronchitis or emphysema), the common cold, and other acute upper respiratory infections. Chronic cough is quite different from acute cough (G–1, p. 13–

15; G–15, pp. 11–12; G–17, pp. 8–9, T–938–39). In chronic lung conditions, chronic inflammatory processes in the patient's lower bronchial pulmonary passages produce thick, tenacious sputum; patients with such conditions need to cough in order to evacuate this sputum. Patients with acute upper respiratory infection usually cough because of minor irritation of the upper portions of the throat or respiratory system by thin secretions from the upper nasal and nasal pharyngeal passages. In such patients, the cough is generally of short duration and generally does not produce the thick, tenacious secretions seen in chronic lung conditions. Aside from these observable clinical differences between acute cough and chronic cough, chronic cough may involve nerve pathways from the throat or lungs to the cough center that are different from those involved in acute cough (T–995; T–942). Also, in chronic cough, patients cough for such a long period that they adapt to the cough by becoming less aware of the cough and sometimes less sensitive to cough stimuli (G–36, p. 422; G–46, p. 427; T–559–60). Adaptation to chronic cough may have physiological as well as pyschological origin; it may result from decreased excitability of certain receptors in the lung (T–995; G–40, p. 445).

The major indication for use of antitussive drugs is suppression of cough in a patient who has an upper respiratory infection that is interfering with rest and sleep (P22–18, p. 397). By suppressing cough, the drug allows the patient to sleep.

Investigation of an antitussive drug involves both animal studies and human (clinical) studies. With respect to new drugs generally and antitussive drugs in particular, there is no absolute requirement that a sponsor show the precise mechanism of action of a drug. In fact, mechanisms of action have not been identified for many drugs, including certain drugs marketed subject to approve NDA's. This situation reflects certain limits, both technical and ethical, on the extent to which biomedical research may establish pharmacological mode of activity in human subjects. Because these limits exist in the case of antitussive drugs, animal studies are used to determine, among other things, the mechanism of action of such drugs, i.e., specific inhibition of the cough center or other action as described earlier in this section. Many drugs have antitussive effects but do not act specifically on the cough center.[3] For example, a drug that depresses the entire central nervous system also will suppress cough. However, it is important to eliminate general depression of the central nervous system as the cause of a drug's antitussive effect: if a drug only suppresses cough due to such general depressant effect, the patient would have to take a dose large enough to produce sedation to a hazardous degree (P19–1, p. 6).

Clinical studies of antitussive drugs have been done in healthy volunteers in whom cough is experimentally induced, in chronic cough patients, and in patients with cough due to cold or other upper respiratory infections (P14–1, p. 38341). It is generally, but not universally, believed that an antitussive drug needs to be studied in the target population, unless the drug is shown to act by specific suppression of the cough center (G–1, pp. 13–18; G–11, pp. 11–12; G–17, pp. 8–9). Some clinical studies of antitussive drugs have employed subjective techniques in which evaluation is based on the opinions of patients regarding the frequency and intensity of cough and other matters (G–46, p. 433; G–61, p. 146; G–64, p. 125; G–138, pp. 463–64). Others have employed objective measures, such as cough counting devices or use of trained observers (id.).

Applying these principles to this proceeding, I first confront the threshold question of which clinical studies of Benylin or diphenhydramine hydrochloride (see Appendix A) may be considered in support of Benylin's claimed effectiveness "[f]or the temporary relief of cough due to minor throat and bronchial irritation as may occur with the common cold or with inhaled irritants" (P14–5). Different answers to this question have been offered by the parties, the CCABA Panel and the ALJ.

The Bureau contends that studies of Benylin in groups other than the target population may be considered only if diphenhydramine hydrochloride is shown conclusively to act specifically on the cough center of the brain, and not through some other mechanism such as soothing the throat or causing general sedation (G–11, pp. 11–13; G–29, pp. 2–3). The Bureau also asserts that it has not been shown that this drug acts specifically on the cough center (G–1, pp. 8, 14–15; G–11, pp. 12–13; G–29, pp. 2, 7; T–529; T–961).

WL/PD argues that all of the studies submitted should be considered, and that the issue of the mechanism and site of action of diphenhydramine hydrochloride as an antitussive has been "distorted out of proportion" and "is of only marginal importance in this case" (Brief, pp. 87–93). The companies claim that, in any event, two of the animal studies (P5–5; study by Rispat, et al., attached to P19–16) show that diphenhydramine hydrochloride's antitussive activity occurs in the cough center of the brain.

The CCABA Panel reached the following conclusion concerning the mechanism of activity of diphenhydramine hydrochloride (41 FR 38312; September 9, 1976):

> The exact mechanism of action of diphenhydramine is not precisely known. However, because of its ability to inhibit the cough reflex resulting from stimulation of the superior laryngeal nerve, the Panel believes a central site of activity of diphenhydramine is a reasonable mode of action. Furthermore, the animal studies are cited as evidence that cough inhibition is not due to a general depression of the central nervous system but to a specific action, similar to codeine, on the "cough center."

Based upon this conclusion, the CCABA Panel apparently considered all studies of diphenhydramine in its determination that the drug is a safe and effective antitussive agent.

In the Initial Decision, the ALJ considered only the two studies in subjects with acute cough in support of Benylin's claim for relief of cough due to cold; he considered only the studies of subjects with experimentally induced cough in support of Benylin's claim for relief of cough due to inhaled irritants (ID, pp. 15–17, 19, 20). The ALJ found that animal studies indicated that diphenhydramine hydrochloride possessses antitussive activity (ID, p. 17). He did not, however, find that the drug acts specifically on the cough center. In addition, he did not accept clinical studies in patients with chronic cough (Protocols 702–3, P7–6 through P7–10; 702–4, P8–1 through P8–5; 184–35, P9–7 through P9–10; 184–36, P9–11 through P9–14; and 184–37, P9–15 through P9–19) as evidence of Benylin's effectiveness. He reasoned that chronic cough is not among the proposed conditions of use in the labeling (ID, p. 16).

Due to the differences among types of coughs and the possibility of underlying neurological differences, as described above, I find that an antitussive drug needs to be studies in its intended target population, unless the sponsor shows that the drug specifically suppresses activity in the cough center of the brain (G–1, pp. 13–15; G–11, pp. 11–13, 18, 30–31; G–17, pp. 8–9). If the sponsor shows that an antitussive drug intended for use

[3] For example, before a patient is bronchoscoped, a local anesthetic may be applied to the bronchial mucosa to prevent the patient from coughing (T–940; G–46, p. 435). Although the anesthetic stops cough, it is not classified as an antitussive drug (id.). Similarly, drugs that paralyze the musculature prevent cough, but are not classified as antitussive drugs (id.).

in cough due to cold specifically suppresses activity in the cough center of the brain, the sponsor may study the drug in patients with other forms of pathological cough, e.g., chronic cough (id.). Nevertheless, studies in the target population are best and are considered by some experts to be indispensable (G–15, pp. 11–12; G–27, pp. 13, 27; G–78, p. 2193). I find that an antitussive drug that acts by inhibiting the cough center is likely to be effective in both chronic and acute cough (G–1, pp. 13–15; G–11, pp. 11–12, 18, 30–31; G–17, pp. 8–9, G–46, p. 437), at least in the absence of certain overwhelming stimuli that may exist in extreme cases of certain diseases (G–1, p. 14). If the sponsor does not show that an antitussive drug's mechanism of action is specific suppression of activity in the cough center, the sponsor must submit at least two studies in the target population to satisfy the statutory requirement of substantial evidence of effectiveness (G–11, p. 31; G–27, pp. 5–6).

Because WL/PD has not shown that diphenhydramine hydrochloride suppresses or inhibits cough by means of specific activity in the cough center, it is scientifically inappropriate for FDA to consider the chronic cough studies in support of Benylin's claimed effectiveness for cough associated with the common cold. If in the future WL/PD proves such activity, the agency will consider these studies.

The requirement of at least two studies is found in § 505(d) of the act "*investigations*" (emphasis added) and in §§ 314.1(c)(2)12.c. and 314.111(a)(5)(i) of the regulations (21 CFR 314.1(c)(2)12.c. and 314.111(a)(5)(i)). These requirements are founded upon a basic proposition of science that an experiment must be reproducible in order for the results to be considered valid (G–11, p.13).

In reaching my finding about diphenhydramine hydrochloride's mechanism of action, I considered those animal studies in the record that were submitted as relevant to this question. Several of the reports submitted (P5–4, P5–6 through P5–11) were not designed to allow determination of the drug's mechanism of action; others lack sufficient detail to permit scientific evaluation or do not even mention diphenhydramine (G–1, pp. 8–12; G–29, pp. 6–7). WL/PD's one expert witness on the drug's mechanism of action does not refer to these reports (P19–16).

The company relies on two studies, the Domenjoz study (P5–5) and the Rispat study (Attachment, P19–16), to show that the antitussive activity of diphenhydramine occurs in the cough center of the brain (Brief, p. 93; P19–16, pp. 4–6).

The 1952 Domenjoz study was conducted in anesthetized cats in which cough was induced by stimulating the superior laryngeal nerve (P5–5; G–1, p. 12; P19–16, pp. 2–3). This is one of the major nerves carrying impulses from the upper respiratory tree, including the throat and larynx, to the medulla, where the central cough center is located (G–1, p. 12). The principal purpose of this study was to describe the mechanical aspects of this type of experiment (P5–5; G–1, p. 12; G–11, p. 12). Although the author states that "we were also able to observe a clear antitussive action of benedryl [diphenhydramine hydrochloride] * * * with the use of our method," he gives no supporting data for this statement that would allow any sound conclusions to be drawn about diphenhydramine's mechanism of action (G–1, p. 12; G–11, p. 12; G–29, p.3).

The 1976 Rispat study (Attachment, P, 19–16) was a French study, also conducted in anesthetized cats. WL/PD learned about this study from its expert witness, during the proceeding. The Rispat study employed the Domenjoz method and another method which, the witness conceded, does not distinguish between central and peripheral action of an agent (P19–16, p. 5). The study compared a drug called zipeprol to several antitussive drugs, including diphenhydramine. From the report of the study, it is evident that the authors did not consider the method used to be capable of establishing conclusively a drug's mechanism of action. With respect to zipeprol, the article states that "[t]he antitussive properties *appear to be* due to a central action" and that "[t]he antitussive properties of zipeprol *may be* attributable to a central action" (Attachment, P19–16, pp. 523, 530, emphasis added). The article also states that "other properties may be implicated in the antitussive action of zipeprol (id., p. 530). Moreover, it is unclear whether "a central action" refers to general action on the central nervous system or to specific action in the cough center (T–961). Thus, the results of this study are consistent with the Bureau's position that the mechanism of action of diphenhydramine has not been conclusively established (T–529). In analyzing the Rispat study, WL/PD's expert witness claimed only that "[t]he antitussive properties of diphenhydramine *may* thus be attributable to a central action" and conceded that the study does not rule out cough suppression due at least in part to diphenhydramine's "potential anticholinergic and antihistaminic properties" (P19–16, p. 5). The Bureau agrees (G–29, pp. 5–7; T–529).

Both of these studies are subject to the fundamental criticism that the superior laryngeal nerve stimulation method is incapable of distinguishing general depression of the central nervous system from specific action in the cough center of the brain (G–29, p. 5; T–523–29; T–993). In recent years, many researchers have begun using more sophisticated methods such as insertion of needle electrodes directly into the cough center of the laboratory animal's brain to determine the mechanism of action of an antitussive drug (G–1, p. 13; T–981, 982, 985, 986, 993; P23–12; P23–9). The superior laryngeal nerve stimulation method still is sometimes used to determine whether a drug under study possesses *any* antitussive activity (P23–10, P23–11).

Another limitation of the Domenjoz method is the need for controls to eliminate the possibility that the drug used to anesthetize the subject cats, rather than the test drug, is responsible, through general depression of the central nervous system, for any observed antitussive effect (G–29, pp. 3–5; G–46, p. 429). Alternatively, or additionally, local anesthetic properties of diphenhydramine may have contributed to suppression of cough in the Domenjoz and Rispat studies of this drug (G–29, p. 5; G–46, p. 436; G–51, p. 606).

WL/PD's expert witness argues that because the Domenjoz method requires that the cats be deeply anesthetized before challenge with the potential antitussive agent, the Rispat study rules out the possibility that diphenhydramine stops cough through general central nervous system depression (P19–16, p. 5). I reject this argument, for two reasons. First, the authors do not state that the cats were deeply anesthetized (Attachment, P19–16; T–961–62). Second, Domenjoz describes this method as requiring light anesthesia (P5–5, p. 3). Pronounced anesthesia would interfere with the elicitation of cough by means of electrical stimulation of the superior laryngeal nerve (G–29, p. 4).

I find, therefore, that the Domenjoz and Rispat studies do not provide any basis for concluding that any antitussive effects of diphenhydramine are due to specific action on the cough center of the brain (G–1, pp. 8, 14–15; G–11, pp. 12–13; G–29, p. 7; T–529, T–961). Due to limitations in the method used in these studies, their results are not inconsistent with the Bureau's belief that diphenhydramine hydrochloride has antitussive activity but that neither the mechanism by which diphenhydramine acts nor the proper dose for cough suppression has been established (G–11,

pp. 11, 30; G–17, p. 15; G–27, pp. 9, 26; G–29, p. 7; T–586–89). This activity may be general central nervous system depression rather than specific depression of the cough center (G–1, pp. 7–8, 14–15; G–5, p. 5; G–11, pp. 11–30; G–29, pp. 6–7; G–40, p. 446; T–529).

To determine the mechanism of action of diphenhydramine hydrochloride, the CCABA Panel considered several of the animal studies referred to above (P5–4; P5–6; P5–7) and a review article, Loew, E. R., "Pharmacology of Antihistamine Compounds," *Physiological Reviews*, 27:542, 1947) (P14–1, p. 38341). Among the literature the Panel considered was a discussion in another paper (P5–9, p. 425) that mentions the Domenjoz study (P5–5). As mentioned above, the Panel acknowledged that "[t]he exact mechanism of action of diphenhydramine is not precisely known," but concluded that "a central site of activity of diphenhydramine is a reasonable mode of action" (P14–1, p. 38341). The Panel went on to say that "the animal studies are cited as evidence that cough inhibition is not due to a general depression of the central nervous system but to a specific action, similar to codeine, on the 'cough center'" (id.).

It is evident from the Panel's own words that it had reservations about its findings concerning diphenhydramine's mechanism of action. The Panel found merely that a central site of action is a "reasonable" (meaning "plausible") mechanism of action and apparently treated the animal studies as persuasive although not conclusive evidence in support of this finding. It is on the question of how certain one must be that an antitussive drug acts specifically on the cough center, before FDA may consider clinical studies in other than the target population, that I part company with the Panel. I believe that the Panel applied too loose a standard. It is not enough that specific action on the cough center be a plausible inference from the various animal studies. Plausibility is not proof. Under section 505 of the act, approval of an NDA must be based, among other things, on scientifically rigorous and valid evidence of effectiveness. The requirement for such evidence applies to each step in the chain of reasoning intended to show effectiveness. It is necessary, therefore, that a sponsor show with scientific rigor and validity that specific action on the cough center is an antitussive drug's mechanism of action, before FDA may accept studies done in other than the target population as substantial evidence of the drug's effectiveness. All of the evidence in the record concerning the mechanism of action of diphenhydramine hydrochloride is merely suggestive; more carefully designed studies are needed to establish precisely what this mechanism is.

I am also concerned that the Panel's protocols for studies required to move a Category III drug (one for which available data are now insufficient to permit final classification of the drug) into Category I (consisting of drugs that are generally recognized as safe and effective) (P14–1, pp. 38354–55) do not recognize the principle that an antitussive drug needs to be studied in the target population unless scientifically rigorous and valid evidence exists that the drug acts specifically on the cough center. The Panel recommended that the two required studies consist of either one study with experimentally induced cough plus a study in cough due to respiratory disease, or two studies by different investigators in patients with respiratory disease (P14–1, p. 38355). FDA will respond to this recommendation, and to my concerns, in its tentative final regulation on the CCABA Panel report and the September 9, 1976 proposal.

The induced cough studies considered in this proceeding (see Appendix B, paragraphs 3, 4, and 5) show that diphenhydramine hydrochloride possesses some pharmacologic activity in subjects with experimentally induced cough (G–17, pp. 6–8; G–27, pp. 10–11), but do not clearly demonstrate antitussive effects even in this artificial setting (G–11, pp. 13–15). Moreover, there is strong evidence that studies of an antitussive drug in subjects with experimentally induced cough are only of limited value in demonstrating that the drug has antitussive properties for treatment of pathological cough (G–1, pp. 17–18; G–11, pp. 14–15, comparing G–136, G–46, G–137, and G–64; G–17, pp. 7–8; G–36, p. 421; G–46, pp. 431–33; G–51, pp. 278–79; G–61; G–62; G–64; P19–11, p. 10; T–745–46).[4] Accordingly, I find that, in determining the effectiveness of a drug whose precise mechanism of antitussive action has not been established and that is offered for treatment of cough associated with the common cold, induced cough studies are no substitute for adequate and well-controlled studies in the target population (id.). (If the sponsor shows specific action on the cough center, the sponsor may study the drug in patients with chronic cough.) Induced cough studies do, however, have value as an early screening technique to determine whether an agent has any form of antitussive activity (G–1, pp. 17–18; G–11, p. 14; G–17, p. 7; G–138, p. 465), but cannot be offered as substantial evidence of the drug's effectiveness in pathological cough. Induced cough studies simply do not show that the drug will have the effect it purports or is represented to have under the recommended conditions of use (section 505(d) of the act).

Because I do not believe that diphenhydramine has been shown to act by specific suppression of the brain's cough center, it is necessary that Benylin has been the subject of two or more adequate and well-controlled studies producing evidence of a significant reduction in coughs in the target population.

What is the target population for Benylin? The sponsor-proposed labeling for Benylin states that the drug is "[f]or the temporary relief of cough due to minor throat and bronchial irritation as may occur with the common cold or inhaled irritants." In its brief, WL/PD identifies two studies (Protocols 266–9 and 266–17) as having been done in the "target population of patients with coughs associated with the common cold" (Brief p. 48). Similarly, the Bureau identifies these two studies of Benylin as the only ones done in the target population (Brief, p. 159). The Bureau also asserts that no target population exists, outside industrial settings, that has cough due to inhaled irritants (id.). In the Initial Decision, the ALJ interpreted the labeling as creating two distinct conditions for Benylin usage: cough due to the common cold, and cough due to inhaled irritants (ID, pp. 15–16).

I find that the target population for Benylin is cough associated with the common cold. In the first place, even if there exists a target population with cough due to inhaled irritants, it is not clear that it would be desirable for these patients to suppress coughs that would help eliminate the irritants (G–46, p. 427; T–801). Second, to the extent that such population consists of individuals who work in industrial environments in which they are exposed to inhaled irritants, many of these individuals would be performing tasks, e.g., operation of machinery, that contraindicate use of a drug that produces drowsiness (G–5, p. 5). Third, even if there exists a target population with cough due to inhaled irritants, I find that the record does not support the

[4] The cited sources raise additional and legitimate doubts about the wisdom of the CCABA Panel's recommendation that FDA treat significant results from one induced cough study and one study in cough due to respiratory disease as acceptable evidence of effectiveness.

ALJ's finding (ID, p. 17) that there is sufficient similarity between experimentally induced cough and cough due to inhaled irritants to extrapolate test results from one situation to another. Early studies of the induced cough technique explored the suitability of various irritants for inducing coughs (G–36, pp. 402–407; G–46, pp. 428–29). These studies suggest that different irritants (or even different concentrations of the same irritant) do not uniformly have the same effects on subjects in terms of production of coughs (id.). Nor are the coughs produced by different irritants consistently affected by standard antitussives (id.). These studies show that it cannot be assumed that an agent effective in suppressing cough induced by citric acid aerosol or another test substance will suppress cough due to any other inhaled irritant.

Finally, cough due to inhaled irritants to which a patient is exposed over a long period of time may differ from experimentally induced cough in which the subject is exposed suddenly and briefly to an irritant; this possible difference also renders invalid any prediction of antitussive effect on cough due to inhaled irritants that is based on induced cough studies (G–36, p. 422).

In light of the above findings, FDA should consider disagreeing with the recommendation of the CCABA panel that the emphasized phrase in the following claim be regarded as acceptable indication for a safe and effective OTC antitussive drug: "For the temporary relief of cough due to minor throat and bronchial irritation as may occur with the common cold (cold) *or with inhaled irritants*" (P14–1, p. 38342, emphasis added).

2. *Adequacy of studies in patients with cough due to cold.* Two of the studies of Benylin were done in patients with cough due to the common cold (the Tebrock study, Protocol 266–17, P9–1 through P9–6), and the Burke Study, Protocol 266–9, P7–1 through P7–5. Neither is an adequate and well-controlled study that offers substantial evidence of Benylin's effectiveness as an antitussive drug (G–3, p. 8; G–7, pp. 3, 9; G–11, p. 31; G–13, p. 5; G–15, pp. 7, 10–11).

Before I discuss the particular deficiencies in these studies, I believe I should express my general disappointment that the sponsor, a major pharmaceutical firm, submitted so many studies that fall so short, in so many different ways, of meeting scientific standards for clinical evaluations. It seems almost tragic that the Tebrock study, a large study in the target population, was not designed to include adequate controls and to use sufficiently objective methods so that Benylin's effectiveness as an antitussive could be evaluated.

Even if the design and conduct of the Tebrock study satisfied fully the requirements of the act and 21 CFR 314.111(a)(5)(ii) for an adequate and well-controlled investigation (which they do not, as explained elsewhere in this Decision), the results of the study do not provide substantial evidence of Benylin's effectiveness under the conditions of use recommended in its labeling. In this study, Benylin relieved cough in 32.4 percent of patients receiving it, while the "placebo" (Benylin vehicle) relieved cough in 23.2 percent of patients receiving it. Thus Benylin offered only a 9 percent advantage over the control. As discussed below, the statute's requirement that a drug "have the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the proposed labeling thereof" (section 505(d) of the act) is not met unless a drug offers a clinically significant benefit.

Both the Tebrock study and the Burke study employed unsuitable controls. The "placebo" in the Tebrock study contained ingredients that may be active. Also, no positive control was used. The Burke study is deficient in that it did not include a placebo drug. Furthermore, the positive control used (Benylin plus codeine) is not a suitable positive drug control.

The issues of clinical significance and the need for adequate controls are of overriding importance and deserve detailed treatment here. There are other important deficiencies in the Tebrock and Burke studies that prevent their consideration as substantial evidence of Benylin's effectiveness. These deficiencies are addressed below in section VI, of this Decision, below.

3. *Clinically Significant Effectiveness.* The evaluation of any drug starts with a determination of its effectiveness. FDA may approve a new drug only if the sponsor shows, by substantial evidence, "that the drug will have the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the proposed labeling thereof" (section 505(d) of the act). Even a drug that is not a new drug is misbranded and subject to regulatory action under section 502 (a) and (f) of the act (21 U.S.C. 352 (a) and (f)) if it does not perform as claimed.

Although the act provides an exact standard for evidence of new drug effectiveness, it does not define "effectiveness" itself. Congress has left to FDA the task of deciding exactly how effective a new drug must be to merit approval. As mentioned above in section IV.A. of this Decision, FDA has by regulation established a definition of "effectiveness" for OTC drugs (21 CFR 330.10(a)(4)(ii)).

In administering the act and regulations thereunder, the agency has not quantified the required degree of effectiveness for new drugs, and cannot do so. FDA recognizes that a drug may not be effective for all patients with a disease, or may not be equally effective in all patients. See S. Rept. No. 1703, 87th Cong., 2d Sess. 16 (1962). Determination of a drug's effectiveness is a part of the risk-benefit analysis that lies at the center of the drug approval process and that takes into account such factors as the seriousness of the disease for which the drug is intended, the availability of other therapies, and the public health implications of the product's availability (Ref. 1, pp. 3, 61). Thus, a drug that is moderately or even highly effective in the treatment of such minor, self limiting conditions as the common cold would not be approved if it were characterized by even moderately serious side effects. On the other hand, relatively toxic drugs are approved if they have value in the treatment of malignant tumors.

Drug effectivenes is determined principally through comparative clinical experiments (see 21 CFR 314.111(a)(5)(ii)). The objective of these experiments is to detect a difference between the test drug and a control. A statistical test can determine the probability that an observed difference between the test drug and the control occurred by chance. As a matter of scientific custom, a statistically significant difference has sometimes been considered one that is likely to occur by chance 1 in 20 times or less (P19–9, p. 2; T–557). The probability (P) of 1 in 20 or less is expressed as a decimal $P<0.05$. With increased use of computers, it has become more common for statisticians to supply the probability that a difference is likely to occur by chance and leave to others judgments about the significance of the result in the particular settings in which it is to be used (e.g., medicine, engineering, sociology, or industry). One expert on design of clinical trials explains that, "[i]t is well to remember that statistics prove nothing—they are merely a device for establishing the betting odds on the reproducibility of the results by mere chance" (G–78, p. 2196).

Clinical significance concerns the degree of benefit a patient will receive from a drug (G–13, pp. 10–11). In its

exceptions to the Initial Decision, the Bureau essentially equates clinically significant effectiveness with substantial evidence of effectiveness as required by section 505 of the act (Exceptions, p. 56). Statistically significant evidence is not the same as clinically significant evidence that a drug will have the effect it is represented to have (G–13, p. 10).

Achievement of a statistically significant difference between groups depends heavily on sample size (T–278, 946–47). Using large enough samples of subjects, an investigator testing a particular drug could obtain results that are highly significant statistically even if the clinical advantage of the drug is trivial (Ref. 2):

> [In a therapeutic trial of 200 patients, improvement occurs] in 49 of 100 patients treated with Drug W, and in 50 of 100 patients treated with Drug V. Noting the small increment of 1 percent between 49 percent and 50 percent, we would conclude that the difference was clinically unimportant. Suppose we now conducted the same therapeutic trial with a sample size of 50,000 rather than 200 patients. If we encountered the same difference of 49 percent improvement for Drug W and 50 percent for Drug V, we might still be unimpressed by the difference of 1 percent. Nevertheless, in the second trial, with 50,000 patients, the 1 percent difference is "statisticaly significant" ($X^2=5.0$; $P<0.05$), whereas in the first trial of 200 patients, the identical increment of 1 percent is not "significant" ($X^2=0.02$, $P<0.75$).

In this example, the 1 percent difference would not be judged clinically significant if the improvement represented 2 pounds lost after treatment with an anorectic agent (Ref. 1, pp. 18–19). But if the treatment were life-saving, the 1 percent difference, representing 250 individuals out of 25,000 whose lives were saved by Drug V, would be judged highly important. Thus, the seriousness of the disease for which a drug is indicated is a critical element in a judgment whether a statistically significant difference is also clinically significant.

The fallacy of equating statistical significance with clinical significance can be shown by considering that in a study of 100,000 subjects in which 50,000 subjects receive the test drug and 50,000 subjects receive the control, an improvement in just 5 subjects who received the test drug would have statistical significance ($P=0.03$). Yet the percentage of subjects in the test drug group who improved would be only 0.01 percent. Although only one in 10,000 subjects responded, the results would be statistically significant because of the large sample size. This result is a *reductio ad absurdum* of the argument that statistical significance is the same as, or is sufficient for, clinical significance. (Obviously, with such a result, no risks associated with the drug could be considered acceptable.)

These examples show why, in determining whether results are clinically significant, a decisionmaker must take into account factors in addition to the 95 percent confidence level. It should also be noted that, in other cases, a result that is only at the 90 percent confidence level could have great clinical significance.

Applying these principles to the Tebrock study of Benylin, I find that the 9 percent difference on the first day of the study between patients receiving Benylin and those receiving the vehicle may be statistically significant because of the large number of patients involved and the analysis used, but is not clinically significant (G–11, p. 21; G–13, p. 11; G–15, pp. 9–10; G–27, p. 19, T–557–560).[5] Aside from the smallness of this difference, the percentages of patients reporting improvement in both groups—32.4 percent in the Benylin group and 23.2 percent in the Benylin vehicle group—seem surprisingly low considering the placebo effects that have been observed with the administration of antitussive drugs (G–36, p. 400; G–46, pp. 434–35) and considering the self-limiting nature of upper respiratory infections. Of every 3 patients who received Benylin, 2 were not helped; of every 5 patients who received the vehicle, 4 were not helped. Thus, calculations of the statistical significance of differences between the groups are based on figures that already seem low, in terms of clinical significance. With these considerations in mind, a demonstration of only 9 percent difference between the groups is particularly unimpressive. Based on the results of the Tebrock study, of every 9 patients who receive Benylin, 6 will report no benefit. Of the 3 who report benefit, only 1 will report benefit from the drug itself, while the other 2 who report benefit might have been equally benefitted by a placebo. (For purposes of this analysis, I am treating the Benylin vehicle as a true placebo.) A demonstration of no more than a 1 in 9 [6] chance of obtaining relief from a cough preparatation does not show that the drug will have the effect it purports to have under the conditions of use suggested in Benylin's labeling.

The discussion so far considers the Tebrack study results from the viewpoint of a physician or other scientist trying to decide whether statistically significant results are also clinically significant. Also important is the viewpoint of the patient, especially when one considers that the sponsor has proposed to make Benylin available to the public without a prescription. A patient who buys an OTC drug for relief of symptoms associated with the common cold expects that it will, in fact, treat the indicated condition effectively. The patient does not expect to be taking a 1 in 9 chance that the drug will help. If the drug were for a more serious condition (which generally, but not always, would be a prescription drug), the patient may be more willing to take such a chance when that course of action appears medically advisable. In the case of a prescription drug, the labeling sometimes makes available to physicians summaries of the findings of clinical studies, including differences between results with the drug and those with the controls. Some physicians may even explain to patients that studies of a prescription drug have not shown effectiveness in all who take it. In the case of OTC drugs, however, such information is not provided in the labeling or otherwise available to patients. Certainly, the proposed labeling of Benylin does not provide information about the clinical studies of the drug's effectiveness.

It should also be noted that the results of the Tebrock study were statistically significant only on the first day of the study, and that the results even on that day were just on the borderline of statistical significance (G–15, pp. 9–10). A change in the reports of just a few patients would have eliminated this significance (id.).

In addition, results of the study's other measures of drug-attributable results, which do not support Benylin's effectiveness, should not be ignored (G–7, p. 6). The second such measure was a question directed to patients as to whether they would take this medication the next time they got a cough. More patients were satisfied with the Benylin vehicle (90.3 percent) than with Benylin (84.3 percent).

The other such measure was a question directed to investigators, which called for an overall rating as to beneficial drug-attributable results from medication. The results did not reveal any statistically significant differences

---

[5] In another context, witnesses for WL/PD who testified concerning the Tebrock study recognized the distinction between statistical significance and clinical significance by arguing that the incidence of drowsiness in the Benylin group, although statistically significant, was not clinically significant (P19–5, p. 3; P19–20, p. 3).

[6] The Bureau's argument that the likelihood of benefit from Benylin is 1 in 11 (G–11, p. 21; G–13, p. 10; T–560) is based on the difference between 32.4 percent in the Benylin group and 23.2 percent in the placebo group. The argument does not take into account the large numbers of subjects unaided by either medication.

between Benylin and the Benylin vehicle: positive ratings were 91.6 percent for Benylin and 89.9 percent for the vehicle. Thus, the slight but statistically significant improvement in cough response reported by the Benylin group on day 1 does not carry over to other responses that also are relevant to overall evaluation of the drug's treatment effect.

Witnesses for WL/PD have speculated that the relatively small difference between results with Benylin and results with placebo may have been due to demulcent (soothing) activity of ingredients in the Benylin vehicle, which was used as a control (P19-3, p. 6). The company has also speculated that the percentage of patients who reported subjective observation of reduction in cough is lower than the percentage whose coughs actually were reduced (P19-3, p. 6, T-558-559). (This speculation is based on studies of chronic cough patients; these studies which are not pertinent here because chronic cough patients become acclimated to their coughs in a way that patients with acute cough associated with the common cold do not (T-560, T-995). Adaptation by chronic cough patients is discussed above in section V.A.1. of this Decision.) Moreover, it is at least possible that the subjective methodology resulted in more differences between the two groups than actually existed (T-232; G-46, p. 434; G-64). These theories would not need to be offered if the Tebrock study had employed a true placebo, an active control, and more objective methods of measuring improvement, as discussed in sections V.A.4., VI.A.5, and VI.A.6. of this Decision. In any event, I cannot base my decision on speculations about what the Tebrock study might have shown had it been designed differently.

I also reject the assertion that "[t]he issue of whether 32 percent versus 23 percent (a difference of 9 percent) is clinically significant can only be answered yes when subjective estimates of benefits are used in a disease with a highly variable course" (P19-7, p. 16; T-76; T-108; T-111-113; P19-19, p. 2). As discussed above, I cannot base a decision on speculation that, because of the subjective methodology employed, the test results may show less difference than actually existed. In addition, the bare fact that a drug is for use in a disease with a highly variable course does not have much bearing on the acceptability of a 9 percent difference between test drug and control without consideration of the seriousness of the disease and other relevant factors.

4. *Need for Adequate Controls.* A clinical investigation to show new drug effectiveness must be "adequate and well-controlled" (section 505(d) of the act). FDA's regulations on adequate and well-controlled clinical investigations (21 CFR 314.111(a)(5)(ii)) spell out the principles recognized by the scientific community as the essentials of such investigations (G-1, p. 5, G-3, p. 4; G-5, p. 4; G-11, p. 4; G-13, p. 13; G-15, p. 5; G-17, p. 6; G-31, p. 6). The regulations' "criteria for an adequate and well-controlled clinical investigation * * * are minimal requirements for any valid objective study." *Pharmaceutical Manufacturers Ass'n* v. *Richardson*, 318 F. Supp. 301, 310 (D. Del. 1970). Compliance with the regulations does not guarantee, however, that the investigation will be scientifically valid in every respect.

The regulations require, among other things, that the protocol and the report of study results provide "a comparison of the results of treatment or diagnosis with a control in such a fashion as to permit quantitative evaluation" and state that, "[g]enerally," four types of comparison are recognized: no treatment, placebo control, active treatment control, and historical control (21 CFR 314.111(a)(5)(ii)(*a*)(*4*)).

The Bureau argues that this provision, in some cases, requires more than one of the four types of "generally recognized" controls (G-11, p. 5; G-27, pp. 6-7; T-492-96, T-690-92; Brief, pp. 98-99).

I agree with the Bureau and find that § 314.111(a)(5)(ii)(*a*)(*4*) establishes general principles and essential requirements of adequate and well-controlled studies generally, but does not prescribe detailed directions for every type of study. Scientific principles determine whether, in a particular study on series or studies, all or some of the types of comparison listed in the regulation are properly used (G-11, p. 5; G-27, pp. 6-7).[7] In the case of some drugs or conditions, use of more than one type of comparison may be necessary. The types of comparison listed thus are illustrative, rather than alternative and mutually exclusive. It is no more correct to interpret the regulation as requiring use of but one of the methods of comparison listed than to interpret the regulation as requiring use of all four of these methods. Moreover, the regulation requires that the controls be used "in such a fashion as to permit quantitative evaluation." As discussed below, quantitative evaluation of drug effects is impossible in some cases without use of more than one type of control. Interpreting the regulation to mean that selection of any of the four named types of controls is acceptable in all cases would compel FDA to accept as substantial evidence of effectiveness studies that, due to the nature of certain drugs and the conditions in which they are used, are not in fact well-controlled, susceptible to quantitative evaluation, or otherwise scientifically valid. This result obviously would be at odds with the statutory intent.

I find, based upon what appears to me to be the preponderance of expert opinion, that use of a placebo is necessary to assure the validity of studies of antitussive drugs; in addition, use of a positive control is essential in some such studies, e.g., those that rely on patients' subjective responses, insensitive experimental methods, or experimental methods of unknown sensitivity (G-7, p. 5; G-11, pp. 5, 20-21; G-13, pp. 6-7; G-15, pp. 7, 10-11; G-17, p. 14; G-39, p. 118; G-78, pp. 2195, 2198; G-79, p. 120; G-138, p. 465; T492-96).

A placebo control is needed to distinguish between a true pharmacological effect of the test drug and fortuitous matters such as psychological effects of taking medication or spontaneous improvement of the disease (G-13, p. 5; G-15, p. 7; G-39, p. 126; G-46, p. 431; G-78, p. 2195). A positive control (e.g., a drug known to be effective) is needed to measure the sensitivity of the experimental procedure, i.e., its ability to measure or detect a drug-related effect when one occurs (G-13, pp. 6-7; G-17, p. 14; G-27, p. 7; G-78, p. 2198; G-79, p. 120). Both controls are essential when the sensitivity of an experimental procedure is in doubt (id.). Both controls are at least highly desirable, and may sometimes be essential, when a high degree of sensitivity is needed to detect or measure an effect, such as suppression of cough due to colds, or where the disease or condition is of short duration or is self-limiting (G-7, p. 5; G-11, p. 5; G-17, p. 14; G-27, pp. 6-7).

Applying these principles to Benylin, I find that neither the Tebrock study (Protocol 266-17) nor the Burke study (Protocol 266-9) is well-controlled. The Tebrock study is deficient because it

[7] There are many ways in which FDA informs sponsors of what FDA expects in investigations. FDA encourages sponsors to consult with FDA to determine what controls are needed in studies of particular drugs (T-754-55). FDA provides sponsors with the opportunity to submit protocols for FDA comments before studies begin and to confer with the agency's reviewing staffs (T-754). FDA also makes available guidelines and other information concerning the conduct of drug investigations (G-11, pp. 3-4). Although WL/PD did consult with the CCABA Panel concerning the additional clinical studies needed to show safety and effectiveness, it apparently did not consult the agency's new drug evaluation staff, as would be appropriate for a new drug subject to a prescription requirement.

failed to include a positive control (G–11, pp. 20, 21; G–13, pp. 6–7; G–27, p. 19), which was essential given the study's reliance upon patients' subjective evaluations. Inclusion of a positive control woud have helped determine whether the lack of any statistically significant difference on days 2 and 3 of the study was due to the normal course of the cold, by showing whether a recognized antitussive such as codeine demonstrated a statistically significant difference on these days (G–17, p. 14). Without this information, we do not know whether the lack of difference between Benylin and the vehicle shows that Benylin is ineffective or that most patients' colds had improved rapidly (G–17, pp. 14–15). In addition, the Tebrock study did not employ a true placebo (see section VI.A.5. of this Decision, below).

The Burke study is deficient because it failed to include a placebo control (G–11, p. 19; G–13, p. 11; G–15, p. 10–11). It therefore is not known whether either of the drugs tested, Benylin or Benylin plus codeine, actually had a significant effect of suppressing cough. Furthermore, the Burke study did not include an adequate positive control; a combination of Benylin and codeine is not adequate because the combination has not itself been established as an effective antitussive drug product (G–11, p. 19; G–13, p. 11; G–15, pp. 10–11, 13; G–17, p. 11; G–27, pp. 14, 17). With a combination that has not been established as an effective drug product, there may be an interaction among the ingredients that results in a different pharmacological effect, e.g., decreased effectiveness of one or more ingredients due to lack of adequate absorption (G–15, p. 10; T–535). The CCABA Panel characterized the Burke study as "uncontrolled" (P14–1, p. 38341). Due to the lack of controls, no conclusions may be drawn from the Burke study at to Benylin's effectiveness.

Witnesses for WL/PD contended that the Burke study should not be considered deficient due to lack of placebo control because it would be unethical to use a placebo with sick children (P19–4, p. 4; P19–11, pp. 9–10). Although I agree that there are many cases in which it would be unethical to use a placebo control with sick children (P2°–3, pp. 17–18), I do not believe that this practice is unethical in tests of drugs for mild, self-limiting conditions such as cough due to the common cold (T–782–88, T–849). Indeed, the principal investigator in the Burke study testified that none of the subjects was critically ill or had a disease of a serious nature (P19–4, p. 2).

B. *Safety*

In view of my finding that Benylin has not been shown to be effective, I need not reach a final decision on whether Benylin is safe for OTC distribution. When the effectiveness of a drug is unknown, the agency cannot decide whether its benefits outweigh its risks. These alternative products are discussed in section V.B.7. of this Decision, below. As discussed earlier, it is appropriate for the agency to scrutinize carefully the risks associated with use of a drug for a mild condition, especially when effective alternative products are available that present only acceptable risks. Moreover, any additional clinical studies of Benylin intended to show its effectiveness as an antitussive may shed new light on its safety for OTC use. Accordingly, I cannot now find that WL/PD has satisfied the requirements for establishing the safety of a new drug.

In fairness to the parties, who have long been at odds on the question of Benylin's safety as an OTC drug, I will give my opinion on this question based on studies in the present record and FDA policies. My opinion is that the risks associated with the use of Benylin in the recommended adult dosage of 25 mg might not be so severe as to require a prescription, if Benylin were shown to be effective and were adequately labeled and packaged for OTC use.

1. *Bureau's Safety Concerns.* The Bureau concedes the safety of Benylin if restricted to prescription (Bureau's Motion to Strike, p. 27). It contends, however, that WL/PD has not met its burden of proof that Benylin is safe for OTC distribution and, further, that the available evidence shows conclusively that Benylin is unsafe for use as an OTC antitussive (Bureau's Brief, p. 232).

The Bureau's belief that Benylin is unsafe for OTC distribution is based principally on its soporific (sleep-inducing) effects, which the Bureau believes are associated with impairment of psychomotor functions required in the performance of such tasks as driving a motor vehicle and operating machinery (41 FR 52537 (November 30, 1976); Bureau's Brief, p. 232). The Bureau's position is that a drug that causes drowsiness in the degree that Benylin does is safer when restricted to prescription because patients believe that a prescription drug is potent and are more likely to handle the product carefully and heed any oral warnings given by a physician concerning a prescription drug than the written warnings accompanying an OTC drug (see section V.B.5. below).

The Bureau asserts that Benylin presents additional hazards besides drowsiness: undesirable drying of secretions, which may interfere with the ability of individuals with chronic coughs to raise secretions and to breathe (G–19, pp. 4–5; G–15, p. 18); the possibility of injury to children due to accidental ingestion of Benylin (G–3, pp. 4–6; G–9, pp. 5–6; G–31, pp. 6–13); the possibility that Benylin will be used to treat children, due to its sedative effects, and its toxicity in children (G–3, pp. 5–6; G–9, pp. 5–6, 9; G–31, pp. 6–13, 15); evidence that diphenhydramine's soporific effects are magnified by interaction with barbiturates (G–29, p. 4; G–63; G–66; G–67) or alcohol (G–3, p. 8; G–21, pp. 6–7; G–23, pp. 12–13; G–53, G–55, G–83, G–146); and concern about the interaction between Benylin and other drugs (G–21, p. 7).

I find, however, that the critical safety issue is whether Benylin causes such drowsiness that it should continue to be restricted to prescription. The other safety concerns expressed by the Bureau are also raised by many products that are now distributed OTC and can be addressed adequately by labeling and child resistant packaging.

2. *FDA's Authority to Restrict Benylin to Prescription.* Like the Bureau, I believe that the Initial Decision erred in finding that FDA has limited authority under section 503(b) of the act (21 U.S.C. 353(b)) to restrict a drug such as Benylin to prescription use (Exception A.2, pp. 10–17, citing ID, pp. 6–9). The Initial Decision adopts an overly restrictive interpretation of section 503(b) of the act. That interpretation is unsupported by the statute's plain language, its legislative history, and judicial interpretation. As discussed below in numbered paragraph VI.B.2., Rulings on Exceptions, FDA has ample authority to restrict a drug such as Benylin to prescription use if the agency finds that this restriction is necessary for the protection of the public health.

Although FDA has broad authority to restrict a drug such as Benylin to prescription use, the agency is not compelled to use this authority if it determines that there is sufficient evidence that the drug is safe for OTC distribution.

3. *Evidence Concerning Benylin's Safety.* The record includes voluminous evidence relevant to Benylin's safety. Benylin has been the subject of a number of studies aimed at measuring either effects of diphenhydramine on psychomotor function (Protocol 266–15, P2–6; the Moskowitz study, G–23A, G–146) or drowsiness (Protocol 184–15, P2–10; Protocol 184–18, P3–3; Protocol 184–40, P16–1; Protocol 184–41, P17–1).

Certain of the clinical studies intended to show Benylin's effectiveness also produced data on drowsiness and other effects (Protocols 286–17, 184–35, 184–36, 184–37, P9–3, P9–9, P9–13, P9–17).

I find that two of the studies of drowsiness (Protocol 184–15, P2–10; Protocol 184–18, P3–3) are not valid evidence concerning the safety of Benylin because they were conducted on prisoners, who are an unsuitable population for the study of drowsiness due to medication because prisoners often are drowsy and bored anyway due to inactivity) G–11, pp. 26–27; G–13, pp. 15–17; G–19, pp. 10–11; G–21, pp. 20–21, P10–5, p. 8; P19–9, p. 2; T–248–49; T–583). Prisoners may also be unusually suggestible, may exchange information or the test drugs themselves, and may not be motivated to cooperate fully in the study (id.). In both studies, there was a high frequency of drowsiness reports by subjects in the placebo groups as well as by subjects in the diphenhydramine groups. See Appendix B, paragraphs 3 and 4. I disagree with WL/PD's argument (P19–9, p. 2) that the double-blind, randomized nature of Protocol 184–15 allows comparison of various treatment groups despite the normal drowsiness among prisoners. Subjects who would have reported drowsiness with a placebo can also actually experience a drug effect; it cannot be determined how many of the treated subjects were reporting a drug effect and how many, if any, were reporting a placebo effect (G–13, p. 16; T–281, T–571–72).

Based on the remainder of the studies, I find that there is substantial evidence in the record to support either a finding that Benylin is safe for OTC distribution, or that Benylin is not safe for OTC distribution, if the required information establishing Benylin's effectiveness were available for use in essential therapeutic risk-benefit judgments.

When the record supports either finding, it is important to identify the factors that are criticial in making a choise, to consider these factors from the perspective of public protection, and to reach a decision based on the preponderence of evidence on these factors. Although I cannot now make a final decision on safety, I will, nevertheless, identify the critical factors in such a decision in the hope that this effort will assist the agency should Benylin be shown in the future to be an effective antitussive. First, FDA should pinpoint the extent and degree of risks to the public due to the drowsiness that Benylin causes. Second, FDA should consider whether, and to what extent, prescription status will reduce these risks. Third, FDA should consider the benefits of OTC distribution. Fourth, FDA should consider the benefits and risks of alternative OTC antitussive drugs.

4. *Risks Due to Drowsiness From Benylin.* Diphenhydramine hydrochloride has a pronounced tendency to produce drowsiness in a relatively high proportion of those who take it (G–51, p. 608; G–9, p. 6; G–11, p. 22; G–21, p. 4; G–49, p. 197; G–115, p. 57304). Drowsiness is dose-related: more patients report drowsiness after receiving the antihistaminic dose (50 mg) than after receiving the antitussive dose (25 mg) (G–51, p. 608; P9–10). Objective studies using brain wave measurements have demonstrated that both 25 mg and 50 mg of diphenhydramine hydrochloride have a sedative effect (G–50; G–11, p. 25).

Based upon Protocols 184–40 (P16–1) and 184–41 (P17–1), I find that approximately one-fourth to one-third [8] of patients receiving diphenydramine 25 mg can expect to feel drowsy as a result of the medication. Most of the drowsiness experienced would be moderate or mild, but some patients would feel extremely drowsy (id.; G–17, pp. 16–18; G–21, pp. 14–15).

I reject WL/PD's argument that the evidence of reports of drowsiness in these studies should be disregarded because these reports were inflated by the informed consent procedure and the choice of a relatively sophisticated group of subjects (P16–2, p. 3, P17–2, p. 3; P19–16, pp. 6, 8; T–129). First, I am not convinced that the informed consent procedure employed was unduly suggestive) G–21, pp. 15–16; G–118, G–119). Furthermore, even if this procedure, or the nature of the subject population, or both, tended to inflate the number of reports of drowsiness, these factors would have had an equal impact on reports of drowsiness in the placebo, dextromethorphan, and codeine groups in this double-blind study. The study results show, however, that the incidence of drowsiness in these other groups was much lower than the incidence of drowsiness in the diphenhydramine group (P16–3, p. 4, P17–3, p. 4; G–21, p. 16). See Appendix B, paragraph 7. Thus, the results show that most of the reported drowsiness is due to the pharmacological activity of diphenhydramine (T–130–31).

Drowsiness is a common side effect of OTC antihistamine products (P14–1, pp. 38379–92). However, Benylin probably causes drowsiness in more cases than do other products (id.). The degree of drowsiness that Benylin causes may also be more pronounced than that caused by current OTC products (id.). Moreover, it is inappropriate to do a comparative risk/benefit analysis of drugs used for different indications or under different conditions of use (G–21, pp. 17–18; G–31, p. 18). The benefits of many OTC antihistamines under their recommended conditions of use are established; the benefit of Benylin as an antitussive is not.

5. *Significance of Prescription Status in Assuring Drug Safety.* I am not persuaded by the record that a drug that causes drowsiness is safer when restricted to prescription. The Bureau argues that patients believe that a prescription drug is potent and are more likely to handle the product carefully and heed any oral warnings given by a physician than the written warnings accompanying an OTC drug (G–3, pp. 5–6; G–5, p. 6; G–8, p. 9; G–11, pp. 24–25; G–15, pp. 17–18; G–19, p. 5; G–21, p. 22; G–31, p. 16; G–49, pp. 213–216; T–592, 636, 646–47, 668–70, 842). Yet virtually all of the reports of injuries from Benylin found in the record (e.g., G–3, pp. 4–5; G–9, pp. 5–6; G–31, pp. 7–11; G–41; G–44; G–45; G–49, pp. 203–10; G–60; G–65; G–85; G–88; G–91; G–92; G–117; G–135) occurred with Benylin or Benadryl that had been dispensed upon prescription. Obviously, prescription status is no guarantee that injuries from a drug will not occur (G–49, p. 214).

I believe that, if Benylin were shown to be an effective antitussive drug, it might be possible to devise labeling that would provide adequate warnings of the risk of drowsiness and other ill effects and that, coupled with child-resistant packaging, would enable the product to be safely used as an OTC drug. In devising any such labeling, WL/PD and the Bureau would have to consider inclusion of some or all of the information in the approved labeling for prescription Benylin as well as that

[8] In Protocol 184–40, the incidence of drowsiness in the diphenhydramine group was 38 percent and in the placebo group, 10 percent, a difference of 28 percent. In Protocol 184–41, the incidence of drowsiness in the diphenhydramine group was 34.3 percent and in the placebo group, 10.1 percent, a difference of 24.2 percent. The true difference between the two groups may be greater than the 28 percent found in Protocol 184–40 or the 24.2 percent found in Protocol 184–41. Some of the subjects who reacted to placebo might also have reacted to diphenhydramine (G–13, p. 16; T–281; T–571–72). (In the alternative, some of the subjects who reacted to diphenhydramine might also have reacted to placebo; if this were the case, the true incidence of drowsiness due to diphenhydramine would be less than the difference between the reported figures.) The Bureau argues, in addition, that individuals afflicted with cough due to cold are weak and therefore more susceptible to drowsiness than were the healthy volunteers who were the subjects in these studies, so that drowsiness in the target population may be actually higher (G–15, p. 15).

recommended by the CCABA Panel.[9] The risk to patients from a drug that causes drowsiness is indirect. The drowsiness itself does not cause harm. It is only when the patient tries to undertake a task that requires alertness, such as driving a car, that the drug's sedative qualities pose a risk to the patient and to other members of the public. Suitable labeling of an OTC drug may provide sufficient safeguards for a drug that presents such indirect risks. When a drug presents serious direct risks (e.g., of cancer or other serious disease), adequate labeling for lay use without medical supervision generally cannot be written.

With respect to indirect risks associated with drowsiness, I do not believe that the risk that consumers will not read, or will read and disregard, the OTC drug labeling is greater than the risk that the prescribing physician will not give or that the patient will not remember and heed equivalent oral warnings.

Physicians should provide patients with information about prescription drugs (G–48, p. 33), but often do not (G–80, Refs. 3, 4). Patients often cannot remember much of what their physicians do tell them, especially about treatments (G–80, Refs. 5–10). Oral communications by health professionals are useful when they occur, but they simply are not a dependable source of information that patients need to have in order to use prescription drugs properly. For this reason, FDA has required patient labeling for certain prescription drugs (see, e.g., §§ 310.501, 310.501a, and 310.502 of the regulations (21 CFR 310.501, 310.501a, and 310.502)). Moreover, the agency has underway a program to require most prescription drug products dispensed for human use to be dispensed with labeling directed to the patient to serve as an adjunct to oral communications between the physician and the patient. In the Federal Register of November 7, 1975 (40 FR 52075) FDA published a notice requesting comments to help formulate this program. In the future, FDA will publish a proposed regulation concerning patient labeling for most prescription drugs.

Available information on the inadequacy of oral communications undermines the Bureau's argument that a soporific drug such as Benylin is safer when restricted to prescription because a consumer is more likely to heed a physician's oral warnings than OTC written warnings. A consumer cannot heed a warning he or she does not receive or does not remember. Although it is undoubtedly true that some consumers would not read and heed written information accompanying an OTC drug[10] because that information accompanies each package, it will at least be available to reach more consumers than are reached by physicians' oral warnings (P19–15, pp. 1–2). Moreover, the written information that would appear on the OTC drug will remain with the drug for its shelf life (P19–15, p. 2), long after many consumers would have forgotten any information provided orally. (Eventually, this advantage of OTC status will diminish as prescription drugs become subject to patient labeling requirements, when FDA implements its comprehensive prescription drug patient information program. Meanwhile, more and more pharmacists use stick-on labels to provide certain information vital to safe and effective use of prescription drugs.)

The Bureau's expert witnesses testified that consumers are more likely to assume that a prescription drug is potent and, therefore, are more likely to adhere to the recommended dosage and heed any warnings that are given (G–31, p. 14; T–842). Studies show, however, that many consumers (30 to 80 percent) fail to follow the recommended regimen for prescription drugs (G–80, Refs. 11–14). It is clear that consumers often do not take prescription drugs and what physicians say about them seriously (T–301). To the extent that patient noncompliance stems from a patient's lack of information about the drug, written labeling may improve compliance (Ref. 15). Moreover, even if it is true that many consumers believe that prescription drugs are more potent than OTC drugs, I question whether this belief alone would justify restricting to prescribing a drug that could be adequately labeled and packaged for OTC distribution.

With respect to some drugs, including Benylin, the principal reason for the prescription restriction is the desire to assure patient compliance with instructions and warnings. (Other reasons for such a restriction are discussed in section VI.B.1. of this Decision, below.) It is tempting to reason that all drugs restricted to prescription solely for this reason can be transferred from prescription to OTC status, once adequate patient labeling is required. In some cases, however, the agency may properly decide that the risks a drug poses when patients do not comply are so serious that it is necessary to provide for the additional protection of physician warnings (when they are given) as well as the required written information. For such drugs, the ideal situation would be one in which the physician provides essential drug information orally when writing the prescription, the patient receives a patient information leaflet containing clear, concise, and complete drug information, and the vial of the dispensed drug bears on its label the most important features of this information.

In the case of Benylin, however, the risks presented by the drug do not seem sufficient to warrant continued restriction to prescription as a way of gaining the added protection of physician communications, when they occur, provided that the manufacturer provides essential information in the labeling and packages the drug in child-resistant containers.

6. *Benefits of OTC Distribution.* The preceding section discusses one present advantage of OTC status; the guarantee that important drug information is at least available to consumers. As mentioned, this advantage will eventually be eliminated through implementation of FDA's prescription drug patient information program.

There are other benefits of OTC status where a drug is for a condition that can be self-diagnosed and self-treated, including cough due to cold. OTC status makes a product for such a condition available to consumers without wasting physicians' time and consumers' time and money in office visits. OTC drugs are cheaper than prescription drugs. In addition, because OTC drugs may be purchased in other retail outlets in addition to pharmacies, they also are more readily available and can be obtained with less delay.

Of course, a drug can be made safer by restricting it to prescription, particularly if this restriction is accompanied by a requirement of patient labeling with the dispensed drug. But these restrictions carry costs that should not be imposed unnecessarily.

Self-medication is important, both to consumers as individuals and to a society that is concerned about how best to use its scarce health care resources. FDA should, therefore, strive to allow drugs to be available without need of a prescription whenever this can

[9] In advising WL/PD on any new labeling for Benylin, and in preparing the tentative final regulation based on the CCABA Panel report, FDA needs to examine carefully the Panel recommendation that the labeling of diphenhydramine 25 mg warn of the possibility of "marked drowsiness." (P14–1, p. 38341). The word "marked" may not be understood by the average consumer.

[10] In Protocol 184–41 (P17–1), at least 2 of the 100 subjects failed to heed both written and oral warnings not to drive. We know this because two subjects, one in the diphenhydramine group and one in the dextromethorphan group, reported impaired driving ability as an adverse effect (P17–3, p. 4).

be done without exposing the public to unacceptable risks.

7. Risks and Benefits of OTC Antitussives. When evaluating a drug, it is important to carry out a therapeutic risk/benefit analysis that includes consideration of alternative therapies. There are two drugs, codeine and dextromethorphan, that may now be used as active ingredients in OTC antitussive products. The CCABA Panel found that both drugs have been shown to be effective antitussive drugs, with specific activity on the cough center (G–46, p. 437; P22–18, pp. 400–01), and that they present acceptable levels of risk (P14–1, pp. 38339–40). Neither the Bureau nor the agency has disagreed with this finding (G–11, p. 17). The risks associated with the use of these drugs are described in the CCABA Panel report (P14–1, pp. 38339–40). The principal adverse effect of these drugs is respiratory depression, generally due to doses well above antitussive doses (id.). Neither of these drugs causes much drowsiness at antitussive doses (id.; P16–3, pp. 4; P17–3, pp. 4–5; G–3, p. 7; G–9,pp. 7, 9–10; G–11, p. 23; G–13, p. 14; G–15, p. 14–15; G–17, p. 20; G–21, p. 16; G–31, p. 18; G–46; G–47).

Considering diphenhydramine's unproven effectiveness as an antitussive, its pronounced tendency to produce drowsiness, and the relatively low toxicity of codeine and dextromethorphan in OTC antitussive doses, I find that Benylin exposes the patient to greater risks, with little or no proven benefit, compared to codeine and dextromethorphan. This view is supported by substantial evidence in the record (G–3, pp. 9–10; G–5, p. 7; G–11, p. 27; G–15, p. 19; G–17, p. 20; G–19, pp. 5–7; G–21, p. 16; G–27, p. 22; G–31, pp. 18–19).

* * * * *

As noted at the outset of this section, V.B., on safety, because I have found that Benylin has not been shown to be effective, I do not decide whether it has been shown to be safe.

C. *General Recognition of Safety and Effectiveness*

Because WL/PD has failed to submit substantial evidence of effectiveness (see section V.A. of this Decision above), Benylin is a "new drug" within the meaning of section 201(p)(1) of the act. In addition, there is a lack of published medical and scientific data on the safety and effectiveness of the drug, including literature describing adequate and well-controlled studies demonstrating the effectiveness of the drug (G–1, p. 6; G–15, p. 5; G–17, p. 5; G–31, pp. 4–5). I find that Benylin is not generally recognized by qualified experts as safe and effective under the conditions of use prescribed, recommended, or suggested in the proposed labeling (G–1, p. 6; G–3, p. 3; G–5, pp. 6–7; G–7, p. 9; G–9, pp. 4–5; G–11, p. 6; G–15, p. 5; G–17, pp. 4–6; G–19, p. 11; G–21, pp. 13–14; G–27, pp. 8, 28; G–31, pp. 4–5).

VI. Rulings on Exceptions

A. *Effectiveness*

The Bureau expressed general concern that by relying upon such poorly controlled, poorly designed, and poorly executed studies as the Tebrock and Burke studies, protocols 266–17 and 266–9, the Initial Division has set a new low standard for the approval of new drugs and that this standard will adversely affect the Bureau's evaluation of effectiveness studies generally (Exceptions, p. 32). The Bureau requests that I consider the impact of the Initial Decision on the Bureau's evaluation of all NDA's (Exceptions, p. 6).

1. *Burden of Proof Concerning Alleged Deficiencies in Studies.* The Bureau contends that the Initial Decision improperly places on the Bureau the burden of substantiating certain alleged deficiencies in effectiveness studies of Benylin (Exception B.1, pp. 32–35, citing ID, pp. 11, 12, 21). The Bureau argues that because the Initial Decision has misplaced the burden of proof, the Initial Decision's ultimate finding that Protocols 266–17, 266–9, 256–16 and 794–1 are adequate and well-controlled studies is invalid. WL/PD maintains that the Initial Decision does not shift the burden of proof from the manufacturer to the agency and that WL/PD has met its burden in the proceeding (Reply, pp. 27–31).

It may be useful to explain who has the burden of showing what under the new drug provisions of the act. There appears to be no dispute that a manufacturer has the burden of establishing the safety and effectiveness of a new drug. See *Weinberger* v. *Hynson, Westcott and Dunning, Inc., supra, North American Pharmacal, Inc.* v. *HEW*, 491 F.2d 546 (8th Cir. 1973); *Ubiotica Corp.* v. *FDA*, 427 F.2d 376, 378 (6th Cir. 1970). In fact, WL/PD concedes that it has the burden of proof in this proceeding (Reply, pp. 8–10, 27, 48). It is also clear that the manufacturer has the initial burden of coming forward with some evidence of safety and effectiveness to which the Bureau may respond. If there were absolutely no evidence on either point, the NDA plainly could not be approved because the Commissioner would be required to make the negative findings enumerated in section 505(d).

When the Bureau proposes to withdraw approval of an NDA, it bears the initial burden of adducing new information that, when evaluated together with the information available when the NDA was approved, shows that the drug is not shown to be safe or effective for use under the conditions of use upon the basis of which the NDA was approved. *Hess & Clark* v. *FDA*, 495 F.2d 975 (D.C. Cir. 1974). To meet its burden, the Bureau need only raise significant doubts as to the prior showing of safety or effectiveness. Once this threshold burden is met, the manufacturer is required to prove the drug's safety, effectiveness, or both.

In a proceeding such as this one involving a proposed denial of approval of an NDA (in contrast to a withdrawal of a previous approval), it is less clear whether, and to what extent, the Bureau has a burden of coming forward with evidence or arguments raising an issue. Less guidance is available from the legislative history and case law. Clearly, the Bureau's burden is less than in a withdrawal proceeding, because "new information" obviously is not a prerequisite to initiation of a denial proceeding. In any event, when, in a denial proceeding, the manufacturer has presented its initial evidence of safety and effectiveness and the Bureau has come forward with evidence or arguments that raise a significant question about the adequacy of that evidence to support approval, it seems reasonable that the manufacturer then be required to persuade the decisionmaker that, notwithstanding the question, the drug is safe and effective; if the manufacturer fails to carry its burden of persuasion, the NDA must be denied approval. This allocation carries forward the de facto allocation that prevailed during the period when the NDA was under review in the Bureau. It is the obligation of the applicant to present sufficient evidence to resolve significant questions raised by the Bureau concerning effectiveness or safety. Any different allocation of burden at the denial hearing would distort the pre-hearing review process.

The Bureau takes exception to the following findings: (1) "* * * in the absence of quantitative evidence of activity, ammonium chloride and sodium citrate must be considered inactive substances" (ID, p. 11); (2) "* * * in the absence of a showing of unreliability, the subjective responses of children from the ages of 6 to 12 for Protocol 266–9 will be considered reliable" (ID, p. 12); (3) "[e]ven though the applicant has the burden of showing that a study is adequate and well-controlled, subjective

responses will be considered reliable in the absence of a showing that such responses are inaccurate" (ID, p. 21); and any other findings that similarly place a burden on the Bureau to make a showing as to the inadequacy of a particular aspect of a study. I have identified two such findings in addition to those quoted by the Bureau: (1) "[m]ere allegation by counsel for the Bureau without substantiation that these hypothesized differences between test groups are fatal to the validity of a test cannot be held to be controlling in the present determination" (ID, p. 13); and (2) "[n]o evidence in support of [the Bureau's assertion that the reduction of cough by Benylin in Protocol 266-18 may have been due to gargling the drug for 15 seconds before swallowing, resulting in topical anesthesia of the contact surface and not in antitussive activity] is offered by the Bureau" (ID, p. 18).

I agree with the Bureau that these findings place upon it an inappropriate burden to substantiate possible deficiencies in effectiveness studies. I believe that the Bureau has met its burden of coming forward with evidence or arguments concerning these alleged deficiencies. The Bureau's evidence on these issues, and my resolution of them, are discussed below in numbered paragraphs 5, 6, 7, 9, and 14 of the Decision. The questions raised by the Bureau are not frivolous or trivial; they are significant and material. Therefore, the manufacturer bears the burden of persuading the decisionmaker that the studies are adequate and well-controlled despite the Bureau's allegations.

2. *Definition of "Antitussive".* The Bureau argues that the Initial Decision improperly takes official notice of the definition of "antitussive" in Dorland's Illustrated Medical Dictionary; the Bureau contends that the Initial Decision should, instead, have relied on the definition of "antitussive agent" in the CCABA Panel report (41 FR 38312, September 10, 976; P14-1) (Exception B.2, pp. 35-39, citing ID, pp. 9, 20-21). WL/PD replies that the Bureau's objection to reference to Dorland's definition is "just preposterous" (Reply, p. 31).

I do not share the Bureau's belief that the ALJ erred in taking official notice of Dorland's definition of "antitussive" ("an agent that relieves or prevents cough"). A definition in such a work properly may be the subject of official notice under § 12.95 (21 CFR 12.95).

I do, however, agree with the Bureau that it was improper to treat this definition as controlling in this proceeding, particularly when both parties apparently agree with the definition of "antitussive agent" (an agent that "specifically inhibits or suppresses the act of coughing") in the CCABA Panel report (P14-1, p. 38338). The difference between the two definitions is that Dorland's treats as an antitussive a drug that "relieves" cough without necessarily preventing, inhibiting, or suppressing it. The ALJ then interpreted "relief" of cough as follows (ID, pp. 9, 20-21):

> Relief is the removal or lightening of something oppressive, painful, or distressing. This, in turn, means that as long as an antitussive lessens the intensity or discomfort of a cough, then it is effective regardless of whether or not a reduction in cough frequency has occurred.
>
> * * * * *
>
> Antitussive action is not limited to a mere reduction in number of coughs. A reduction in the severity of the cough or its discomfort can constitute an antitussive action.

The Bureau contends that that interpretation of "antitussive" holds Benylin to too low a standard and that, as a result, the Initial Decision's ultimate finding on effectiveness is "meaningless" (Exceptions, p. 37).

I agree with the Bureau that mere relief of the discomfort of a cough is an inadequate criterion against which to judge an antitussive drug. Many products, including alcoholic beverages, may reduce the discomfort of a cough without affecting its frequency or severity. An antitussive drug should reduce the frequency of cough, or stop it altogether.

It is unclear how the definition of "antitussive" adopted in the Initial Decision affected its ultimate finding that Benylin has been shown to be effective. The definition played but a minor part (ID, pp. 20-21) in the discussion of the clinical studies on which the ALJ based his finding of effectiveness (ID, pp. 9-21).

3. *Use of Both a Positive Control and a Placebo.* The Bureau disagrees with the Initial Decision's conclusion that FDA's regulations do not require that a study has employed both a positive control and a placebo (Exception B.3, pp. 39-40, with respect to ID, pp. 9-20). WL/PD replies that no FDA regulation requires use of both a positive control and a placebo (Reply, pp. 33-34).

As explained above in section V.A.4, I agree with the Bureau's view that, in certain cases, a study does not meet the requirements of the act and 21 CFR 314.111(a)(5)(ii) unless it included both a positive control and a placebo.

4. *Adequacy of Control in Burke Study.* The Bureau points out that the Initial Decision states that the control in the Burke study (Protocol 266-9) was the Benylin vehicle plus codeine, when actually the control was Benylin plus codeine (Exception B.4, pp. 41-42, citing ID, p. 11). The Bureau contends that this factual error invalidates the Initial Decision's conclusion that the study employed an adequate positive control because a combination of a known active drug such as Benylin and a known effective drug such as codeine is not an adequate positive control. WL/PD contends that the Bureau has distorted the ALJ's comments (Reply, pp. 34-37).

I agree with the Bureau that the finding in the Initial Decision that the Burke study is adequately controlled may have been based on a misunderstanding concerning the control used (see footnote 2, above, of this Decision). In any event, as explained above in section V.A.4. of this Decision, I have found that a combination of Benylin plus codeine is not an acceptable positive control.

5. *Use of Benylin Vehicle as Placebo in Tebrock Study.* The Bureau maintains that the Initial Decision concludes erroneously that the Benylin vehicle is an adequate placebo in the Tebrock Study (Protocol 266-17), the Bureau argues that ammonium chloride and sodium citrate, ingredients in the vehicle, may be active substances (Exceptions B.5, 3.6, and B.7, pp. 42-46, citing ID, 10-11, 23, 20). (Five ml of Benylin contain 12.5 mg diphenhydramine hydrochloride, 125 mg ammonium chloride, 50 mg sodium citrate, 1 mg menthol and 5 alcohol. 5 ml of the Benylin vehicle contain all of these ingredients except the diphenhydramine). WL/PD replies that the Bureau has mischaracterized both the expert testimony by a witness for the company and the Initial Decision (Reply, pp. 37-39). The company believes that the presence of ammonium chloride and sodium citrate in the Benylin vehicle does not prevent its use as a placebo.

I find that ammonium chloride and sodium citrate in the amounts used in the Benylin vehicle may have expectorant, demulcent, or other pharmacological activity (G-11, pp. 9, 10, 31; G-13, pp. 5-6, 12; G-15, pp. 7-9; G-27, p. 18; P14-1, p. 38359; T38-41, 45, 520-23). Although both the NAS/NRC and the CCABA Panel (P14-1, p. 38312) found that the effectiveness of these ingredients as expectorants has not been proven, these findings do not establish that these ingredients lack pharmacological activity. Inclusion of ammonium chloride and sodium citrate in the "placebo" prevents Protocol 266-17 from being considered an adequate and well-controlled study under

§ 314.111(a)(5)(ii)(a)(4)(iii), (21 CFR 314.111(a)(5)(ii) (a)(4)(iii)). This regulation requires that a placebo be "an inactive prepartation designed to resemble the test drug as far as possible" (id., emphasis added). "[A] test of two possibly effective agents constitutes two uncontrolled tests, not a test with two controls, "*Cooper Laboratories* v. *Commissioner, FDA*, 501 F.2d 772, 784 (D.C. Cir. 1974).

As explained above in section V.A.4. of this Decision, the purpose of a placebo control is to distinguish between a true pharmacological effect of the test drug and fortuitous matters such as psychological effects of taking medication or spontaneous improvement of the disease.

If the purpose of the Tebrock study was to determine the effectiveness of Benylin, the sponsor should have selected an inactive drug for use as a placebo. Selection of a control that has, or may have, pharmacological activity makes it impossible to determine whether any difference between the test drug and the control is due to pharmacological activity of the test drug.

Inclusion of active ingredients in a placebo could make the test drug look worse than it actually is, if these ingredients have activity that is similar to that of the test drug. For example, if ammonium chloride or sodium citrate has demulcent activity that lessens cough by soothing the throat, patients receiving the Benylin vehicle may cough less. Subtracting the percentage of patients reporting improvement with the Benylin vehicle from the percentage of patients reporting improvement with Benylin would, then, underestimate the true difference between the two drugs in effect on cough. In other words, if these ingredients are demulcents, Benylin may be a more effective antitussive drug than the calculated difference suggests. WL/PD argued that this is the case (P19–3,p. 6). (If this theory is correct, selection of the Benylin vehicle as a control against which to test Benylin shows surprisingly poor judgment on the sponsor's part.)

On the other hand, inclusion of active ingredients in a placebo could make the test drug look better than it actually is, if the ingredients in the placebo have activity that worsens the patient's underlying condition. For example, if ammonium chloride or sodium citrate has expectorant activity that increases the frequency, intensity, or duration of a patient's cough, or that makes the cough more efficient in raising phlegm, patients receiving the Benylin vehicle may cough more, at least initially. (Later, as phlegm is removed, the cough may lessen in frequency, intensity, duration, and efficiency.) Subtracting the percentage of patients reporting improvement with the Benylin vehicle from the percentage of patients reporting improvement with Benylin would, then overestimate the true difference between the two drugs. In other words, Benylin may be a less effective antitussive drug than the calculated difference suggests. The Bureau argued that this was the case (G–11, p. 21; G–15, pp. 8–9; T–806). The plausibility of this theory is supported by the fact that 13 of the 14 patients in the Tebrock study who volunteered the information that the drug they received made their coughs more productive were in the Benylin vehicle group; only one was in the Benylin group (G–04 through G–107). If the Bureau's theory is correct, the marginally statistically significant difference between the two groups on the first day of the study could have been produced, wholly or in part, by the temporary "worsening" of cough in patients in the Benylin vehicle group. Interaction between diphenhydramine and one or both of the two ingredients in question could have prevented a similar effect in the Benylin group. Or any antitussive or sedative effect of diphenhydramine could have counteracted any expectorant effect of ammonium chloride, sodium citrate, or both.

It simply is not known whether ammonium chloride or sodium citrate have demulcent, expectorant, or other pharmacological activity. Without this knowledge, it cannot be determined whether the reported difference between the Benylin group and the vehicle group was underestimated or exaggerated due to activity of these ingredients. Nor can we, with any confidence, attribute the reported difference between the groups to antitussive activity of diphenhydramine. Accordingly, inclusion of ammonium chloride and sodium citrate in the "placebo" makes it impossible to perform the quantitative comparison of the test drug to the control that is central to any clinical investigation of a new drug.

I disagree with WL/PD's argument (P17–3, p. 5; P19–6, p. 8; P19–7, p. 16; P19–9, p. 11) that it would be impossible to formulate a placebo other than the Benylin vehicle without unblinding the study. The regulations do not require that the placebo be identical to the test drug in color, taste, and consistency. The regulations require, rather, that the placebo be "designed to resemble the test drug as far as possible," 21 CFR 314.111(a)(5)(ii)(*a*)(*4*)(*iii*). On cross-examination, expert witnesses for the Bureau suggested ways in which a true placebo could be used without impairing the double blind character of the study (T553, Bureau's Brief, pp. 31–32; T–787, 806).

6. *Subjective Evaluation Generally.* The Bureau takes exception to the finding that the subjective methods of evaluation in the Tebrock and Burke studies (Protocols 266–17 and 266–9) are adequate to permit quantitative evaluation (Exception B.6, pp. 46–48, citing ID, pp. 11–12, 21). WL/PD replies that use of subjective responses is unavoidable in a large scale study of cough due to cold, conducted in a target population in an industrial setting (Reply, pp. 40, 28–30).

I agree with the Bureau and find that the subjective methods of evaluation used in these studies are not adequate to permit quantitative evaluation.

The preponderance of expert opinion is that clinical trials of antitussive drugs that base their findings on patients' subjective impressions do not produce reliable evidence of effectiveness (G–17, pp. 12–13; G–36; G–39, p. 122; G–61, p. 146; G–62, p. 384; G–64, p. 130; G–68; G–72;). Patients' subjective responses concerning the effects of antitussive drugs may be "grossly misleading" (G–61, p. 146). The drug may simply make a patient less aware of cough without reducing its frequency or severity (G–17, p. 12; aG–68, p. 1137). A patient's response may be affected by memory, concentration, mood, sense of well-being, pain or discomfort from other causes, degree of distraction, and passage of time (G–17, pp. 12–13). Most individuals do not sense diurnal and daily changes in their coughs and misjudge treatment effects (G–36, p. 422; G–43, p. 9; G–64, pp. 129–30). Studies using both subjective and objective techniques have repeatedly shown a lack of agreement between a patient's subjective assessment of cough and the actual cough count (G–39, p. 122; G–43, p. 9; G–46, p. 434; G–51, p. 279; G–61, pp. 383–84; G–64, pp. 129–30; G–68; pp. 1239–41).

Countering this evidence is testimony by several expert witnesses for WL/PD that objective cough counting techniques could not practically be used in ambulatory subjects having cough due to cold (P19–3, p. 3; P19–6, pp. 9–11; P19–7, pp. 15–16; P19–9, p. 12; P10–11, p. 8). It was argued that such techniques would diminish the value of the study populations in Protocols 266–17 and 266–9 as reflecting the "real world" (P19–6, p. 9; P19–11, pp. 8, 9).

Although I agree that it is more difficult to use objective techniques to evaluate cough improvement in such patients, the record shows that there are several ways in which this evaluation can be done. Use of these methods

would increase, rather than diminish, the value of these studies in the real world by enabling valid conclusions to be drawn concerning Benylin's effectiveness in the target population.

The Bureau is willing to accept studies of antitussive drugs employing subjective techniques that are designed to avoid the well-documented pitfalls in subjective evaluation of antitussives and that permit quantitative evaluation (G–27, p. 18). I agree with the Bureau's contention that use of a daily diary with specific questions on such matters as number and severity of coughs, pain, and the extent to which the patient's coughs raise secretions from the throat and respiratory system, is one method of obtaining detailed and objective information that would permit quantitative evaluation, including comparison of each day's experience to the baseline (G–11, p. 21; G–17, p. 12; G–27, p. 18; G–77; T–564–65). In addition, it may well be possible for WL/PD to devise an objective study of Benylin's effectiveness (T–949). The CCABA Panel recommended use of objective studies employing the actual recording of cough to document a decrease in cough frequency, intensity, or both (P14–1, pp. 38354–55). The Panel did not find objective techniques unacceptable for evaluating drugs for use in cough due to cold. There are populations, e.g., university students or officer workers, that can be used to study antitussives in cough due to cold by either detailed subjective methods or objective techniques. WL/PD was not obligated to test Benylin in an industrial setting. Alternatively, if the sponsor establishes that diphenhydramine suppresses cough by inhibiting the brain's cough center, the Bureau would accept objective studies in chronic cough patients as evidence of Benylin's effectiveness in cough due to cold (see section V.A. of this Decision; T–565). Accordingly, I do not accept the argument that it would not be possible to develop a more objective study of Benylin's effectiveness in acute cough than was used in the Tebrock study (G–39, p. 126; T–413).

The Tebrock study used a patient record instructing patients to check a box showing whether their coughs were "gone," "better," "same," or "worse" compared to the previous day (P9–3). This was the only question asked concerning the effect of the test drug or control upon the patient's cough. In evaluating test results, WL/PD combined the "gone" and "better" responses to identify patients whose coughs had "improved", and the "same" and "worse" responses to identify patients whose coughs had "not improved" (P9–5, p. 1).

Given the known drawbacks of subjective studies of antitussive agents (as discussed in the third paragraph of this section), it is inappropriate to base a decision concerning a drug's antitussive qualities upon a patient's once-daily subjective perception of the progress of the cough and the patient's interpretation of such undefined terms as "gone," [11] "better," [12] "same," or "worse." [13] This rating system is inadequate because it does not ask the patient to describe changes in a cough's frequency or severity, the extent to which sputum is raised, or pain (G–17, pp. 12–13). The rating used may only have measured patients' awareness of cough (G–17, p. 13).

The fact that a large number of patients participated in the study (P19–3, p. 3) cannot compensate for the inherent unreliability of subjective responses obtained in that study. As the Bureau points out, enrolling many patients merely increases the amount of uninterpretable data without improving its quality (Bureau's Brief, p. 90), especially in the absence of adequate controls.

The Burke study involved 100 children aged 6 through 12. To determine effects of the medication given (Benylin, or Benylin plus codeine), the protocol required that one day after receiving medication the subjects be contacted either by having the patients return to the clinic or sending visiting nurses to the patients' homes (P7–3, p. 8). The protocol stated: "Questions should at first be broadly phrased, such as: 1. How did the medicine work? 2. How did it taste? 3. Was there anything wrong with it?" (P7–3, p. 8). If the initial response was unclear, further questions were to be asked to enable the investigator or nurse to complete a form for each patient that, among other things, described whether the medication stopped cough, reduced severity, reduced frequency, or had no effect (P7–3, pp. 8–10 (attachment)). The procedure was to be repeated on the third day (P7–3, p. 8). WL/PD's summary of the study notes, however, that each patient was in fact observed only once for the recording of these opinions and that the reporting forms did not state on which day of medication the recording occurred (P7–5, p. 1).

The protocol did not mention that individuals other than the child were to provide information on cough improvement. According to Dr. Burke, responses obtained from children were supplemented by responses from someone else, often a parent or other relative (T–768). No record was kept of who actually furnished a response and whether this individual actually had been able to observe the child's cough during the treatment period (T–768).

This subjective method used in the Burke study also is unacceptable because it required the patient (or relative) to make a single judgment of the effects of the medication as many as 3 or 4 days after treatment began (G–31, p. 17; G–3, p. 9; G–27, p. 17). The Burke study used non-standardized questions to obtain responses that are classified by the health professional performing the contact into one of four ill-defined categories. Because no record was kept of the day of medication on which the contact occurred, there is no way of determining the extent to which responses may have been affected by the patient's (or relative's) inability, due to passage of time, to remember the nature of the cough. Similarily, the failure to keep records as to who furnished a response concerning a child's cough, and as to that individual's ability to observe the cough, makes it impossible to perform a quantitative analysis of data from the study because the records of individual patient responses are too variable to be reliable.

It is possible to design objective studies of antitussives in children (T–555). Asking parents to record observations of children's coughs would be a valid way to study antitussive drugs in children, as would using a tape recorder or other apparatus (T–641). However, careful records of parents' observations would have to be kept (T–641).

7. Subjective Evaluation By Children. The Bureau takes exception to the Initial Decision's conclusion that the Burke study (Protocol 266–9) is not deficient by reason of its reliance upon the subjective responses of children aged 6 to 12 (Exception B.9, pp. 48–49; citing ID,

[11] "Gone": Evidence from other studies (e.g., G–68, pp. 1139, 1141) suggests that some patients who are still coughing would report a cough as "gone," while others who have stopped coughing or who are coughing but little would report that a cough still is present. A patient's report that a cough is "gone" or "better" may reflect a patient's awareness of the frequency or severity of the cough rather than any actual effect on these parameters (G–17, p. 12). This is especially true when a drug has sedative or hypnotic properties as Benylin does (id.).

[12] "Better": As explained in the text, when a patient reported that his cough is "better," there was no way of knowing what was meant. Was the cough less frequent? Less severe? More productive of sputum? Less painful? One cannot say, based on the study design. The sponsor concedes that patients may have had difficultry in discriminating between "better" and "same" (P9–5, p. 1).

[13] "Same" or "worse": Here, too, it is impossible to say whether a patient's report that a cough is the "same" or "worse" concerns its frequency, severity, or productivity, the presence of pain, the patient's awareness of the cough, or some unknown factor.

p. 12). WL/PD replies that subjective evaluation of cough by children is an acceptable method (Reply pp. 40, 28).

I reject the exception. The preceding section of this discussion explains why subjective responses generally are unreliable. Based on the evidence in the record, I find that children's subjective responses are not inherently less reliable than those of adults (T-554-55; T-642). Neither group is capable of providing valid subjective evaluations over a period of several days. Whether an antitussive is studied in adults or in children, the study should use objective techniques or careful, detailed subjective techniques, as discussed in the preceding section.

8. *Analysis of Tebrock Study Results.* The Bureau disputes the Initial Decision's finding that the results of days 2 and 3 of the Tebrock study (Protocol 266-17) must be disregarded because they do not permit quantitative evaluation without invalidating the study (Exception B.10, p. 50, with respect to ID, p. 12). The Bureau argues that instead of disregarding the study results on these days, the Initial Decision should have found the study invalid because "the average cough due to cold lasts three days" and Benylin must be effective for at least that long (id.). WL/PD replies that the exception misinterprets the Initial Decision, and the company disputes the contention that the average cold last three days (Reply, pp. 40-41).

The statement in the Initial Decision to which the Bureau excepts follows (ID, p. 12):

> The Bureau additionally asserts that Protocol No. 266-17 does not allow quantitative evaluation due to its use of the previous day as a comparison for cough improvement instead of to an established baseline. This is a valid criticism only with respect to the statistically insignificant results obtained for days two and three. Therefore, the results for these days must be disregarded. However, this criticism is not valid for day one because the use of the drug immediately preceding the study is generally the same point at which a baseline would be established. The subjects of Protocol 266-17 recorded their responses concerning cough on day one at bedtime. Therefore, the use of the previous day baseline is valid for day one and those results are appropriate for quantitative evaluation and statistical analysis.

I agree with the Bureau that it was improper to disregard the statistically insignificant results obtained for days 2 and 3. These results, as well as the clinically insignificant results on day 1 and the failure to detect drug effects using other measures, show that this study does not demonstrate Benylin's effectiveness in cough due to cold of any duration.

9. *Pooling of Data in Tebrock Study.* The Bureau also disagrees with the finding that it is appropriate to pool data from the five test centers for statistical analysis of the test results (Exception B.11, pp. 50-54, citing ID, pp. 12-14). The Bureau argues that such pooling is inappropriate because of the lack of homogeneity of treatment effect among the five centers. WL/PD replies that the Initial Decision decided this issue correctly (Reply p. 41).

I reject the exception. In my opinion, both parties have placed too much emphasis on the appropriateness of pooling results of the five centers participating in this study. In making a scientific judgment about the results of the Tebrock study, one should analyze both the results at each center alone and the pooled results, noting the differences in sample size, the differences in results at various centers, and the closeness of results with the two medications both at the individual centers and overall. As an aid to this analysis, it is wise to employ a statistical technique to determine whether differences between study centers are greater than would be expected by chance alone. A finding of significance, i.e., a difference greater than expected by chance, provides information useful in making a scientific judgment about the results of the study but does not dictate that pooled data be ignored.

I disagree with the Bureau's argument (G-7, pp. 7-8; T-435-36, 446, 464-65) that it is in appropriate to use the Mantel-Haenszel test to determine the appropriateness of pooling data from the five centers because of lack of homogeneity of test results. The developers of the method contemplated that it might be used in situations where there are substantial variations among subgroups with respect to the parameter being measured (Ref. 17, p. 740; T-435). There is no evidence in the record, other than the Bureau's expert testimony, that the Mantel-Haenszel technique requires homogeneous treatment effects as a prerequisite to its use. In a study such as this, results from the Mantel-Haenszel test and other statistical techniques provide data that are useful in making a scientific judgment about the results of a study and how they should be intepreted. For example, in this case, application of the Mantel-Haenszel test resulted in a finding that differences among results at the five centers exceed those that would be due to chance alone (P19-9, p. 9; G-25, p. 4). This result is undoubtedly influenced by the differences in group size at different centers, the small size of some groups, the closeness of the reported results in the Benylin and the Benylin vehicle groups, and possible geographical differences in such things as pollen count (P19-8, pp. 8-9; Bureau's Brief, p. 54).

With the above modifications and one exception, I agree with the Initial Decision's discussion (ID, pp. 12-14) of the appropriateness of the statistical analysis of the Tebrock study in which the data from the five centers were pooled. The exception is that I do not adopt the statement that, "Mere allegations by counsel for the Bureau without substantiation that these hypothesized differences between test groups are fatal to the validity of a test cannot be held to be controlling in the present determination" (ID, p. 13). As discussed in section VI.A.1. of this Decision, above, this statement places on the Bureau an inappropriate burden to substantiate its concerns about a study.

10. *Clinically Significant Effectiveness.* The Bureau takes exception to the finding that the Tebrock study (Protocol 266-17) shows Benylin's effectiveness. The Bureau points out that the study results demonstrate only a 9 percent advantage over placebo (Exception B.12, pp. 54-57, citing ID, p. 15). The Bureau also disputes the ALJ's belief that section 505(d) of the act (21 U.S.C. 355(d)) compels this finding. The Bureau argues that, to be found effective, a drug must offer not just a statistically significant benefit over placebo, but a clinically significant benefit. WL/PD replies, in effect, that Benylin is probably more effective than these figures show because a patient must experience a major reduction in coughs per day before recognizing an improvement (Reply pp. 41-42, Brief pp. 60-61).

I agree with the exception, for the reasons given above in section V.A.3. of this Decision.

11. *Execution of Tebrock Study.* The Bureau takes exception to the statement that the Tebrock study (Protocol 266-17) was carefully executed (Exception B.13, p. 57, citing ID p. 21). WL/PD contends that the study is an adequate and well-controlled scientific investigation that meets all criteria of the regulations (Reply pp. 42-43).

I agree with the exception and find that the Tebrock study was not carefully executed (G-7, pp. 3, 5; T-448). I have already discussed inadequacies in the design of the protocol with respect to controls and measurement of drug effects (see sections V.A.4., VI.A.5., and VI.A.6. of this Decision), as well as in the significance of the study's results

(see section V.A.3., above). In addition, the protocol was not closely adhered to. First, the protocol called for each of the five test centers to enroll enough patients so that at least 120 patients would complete the study (P9–3, p. 2). In fact, only 556 subjects completed the study and 325, representing 58 percent, were in New York City (P9–4, p. 3, Table 4). Also, the efficacy analysis used only 472 subjects (id.). A drop of 84 subjects, or 15 percent, is a substantial number for a 3 day study (G–7, p. 5). Second, although the protocol required investigators to control the number of doses given subjects on the first day of the study (P9–3, p. 3), this was not done (P9–4, p. 4). Rather, the distributions of subjects receiving 1, 2, 3, or 4 doses were established by post-stratification, which was not planned in the design (P19–9, p. 7). Third, a few patients included in the study did not meet the entrance criteria of having cough due to cold. These patients had chronic cough of up to 2 months in duration. Inclusion of them could have affected the study's overall results, in view of the marginally significant difference between the Benylin group and the vehicle group (G–11, pp. 19–20). Their inclusion suggests inattention to detail that could be reflected in other aspects of the study as well (id.).

Although these deficiencies in the execution of the Tebrock study would not, by themselves, warrant rejection of the study under § 314.111, 314.111(a)(5)(ii), they show that the study was not carefully executed.

12. *Use of Pediatric Study to Replicate Adult Study.* The Bureau contends that the Burke study (Protocol 266–9), involving a pediatric population, cannot be relied upon as a replication of the results of the Tebrock study (Protocol 266–17), involving an adult population; and that the sponsor thus has failed to satisfy the requirement that a drug be shown effective by at least two studies (Exception B.14, pp. 57–58, citing ID, p. 15). WL/PD replies that there is no regulation that prevents a study in children from being considered a replication of a study in adults (Reply, pp. 43–44).

Where children comprise part of the target population for a drug, it is important that the drug be tested separately in children (P19–11, pp. 9–10; P23–3; T–554; T–846–47). It cannot be assumed that a drug that acts a certain way in adults will act the same way in children (P23–3, p. 20). As individuals grow from conception to adulthood, they undergo profound and complex changes in anatomy, physiology, biochemistry, and behavior (id.). These changes cause variation in the absorption, distribution, metabolism, and excretion of pharmacological agents and in receptor sensitivity (id). For this reason, FDA has sometimes required sponsors of new drugs to conduct separate studies in children, and the agency is considering a broad program to encourage manufacturers of many drugs commonly used in children to conduct tests in pediatric populations.

As explained above in section V.A.1. of this Decision, the requirements that a sponsor submit at least two studies demonstrating effectiveness is founded on the scientific principle that an experiment must be reproducible in order for the results to be considered valid (G–11, p. 13). In this case, the Burke study cannot be treated as replicating the Tebrock study. Due to the possibile differences in drug activity in adults and children, the agency generally cannot rely upon one study of a drug in adults and one in children, as were submitted in this case, as satisfying the requirement of substantial evidence of effectiveness. (There are exceptions to this rule, e.g., with respect to studies of a topical, relatively nonabsorbable, antifungal agent when there is a sound scientific basis for expecting the drug's safety and effectiveness in adults and in children to be similar.) For example, there are observable differences in how children and adults react to toxic dosages of diphenhydramine (G–31, pp. 7–9; G–44, pp. 107–08; G–49, p. 194; G–51, p. 608; G–88; T–91). These differences may reflect variant pharmacological activity. Accordingly, I agree with the Bureau's exception and find that submission of one study of Benylin in adults (Protocol 266–17) and one in children (Protocol 266–9) does not satisfy the substantial evidence requirement of the act and FDA's regulations.

13. *Reliance on Induced Cough Studies.* Several exceptions relate to the Initial Decision's reliance on Protocol 266–18 and 794–1 as adequate and well-controlled studies, and on the 1960 Bickerman protocol as a corroborative study, in support of Benylin's effectiveness in cough due to inhaled irritants.

The Bureau disputes the ALJ's finding that these investigations are not Phase I studies as defined in § 312.1(a)(2) 10a. and b. of the regulations (21 CFR 312.1(a)(2) 10a. and b.) (Exception B.15, pp. 58–59, citing ID, pp. 16–17). The Bureau also takes exception to the findings that there exists a target population with cough due to inhaled irritants (Exception B.16, p. 59, citing ID, p. 17), and that the induced cough studies show that Benylin will suppress cough in this population (Exception B.17, pp. 59–60, citing ID, p. 17). In addition, the Bureau disputes the finding that Protocols 794–1 and 266–18 and the 1960 Bickerman Protocol are adequate and well-controlled studies (Exception B.18, pp. 60–61, citing ID, p. 18). WL/PD replies that the Initial Decision's findings on these matters are correct (Reply, pp. 44–46).

The issues raised by these exceptions have been addressed in the discussion by my finding that induced cough studies are not adequate and well-controlled studies of antitussive drugs for use in cough due to cold or inhaled irritants (see section V.A.1. of this Decision). For the reasons stated there, I agree with these exceptions.

14. *Effect of Gargling Benylin.* The Bureau takes exception to the finding that cough reduction in Protocol 266–18 was due to an antitussive effect rather than to the possible anesthetic effect of gargling the Drug for 15 seconds (Exception B. 19, p. 61, citing ID, p. 18). WL/PD replies that the Initial Decision is correct.

I agree with the Bureau that there is inadequate support in the record for the finding that cough suppression in Protocol 266–18 is "consistent with the type of activity an antitussive would exhibit and not a anesthetic" (ID, p. 18). It is at least possible that diphenhydramine may have local anesthetic effects (G–29, p. 5; G–46, p. 436; G–51, p. 606), and that, in Protocol 266–18, any cough suppression was due to anesthetic effects. In any event, the proposed labeling of Benylin does not call for gargling before swalling. Therefore, a study in which subjects were told to gargle before swallowing may not be offered in support of Benylin's effectiveness "under the conditions of use prescribed, recommended, or suggested in the proposed labeling thereof" (section 505(d) of the act). Because that instruction was given under Protocol 266–18, this study is not acceptable evidence. (Changing the labeling to recommend gargling would not solve the problem because induced cough studies are not acceptable predictors of a drug's effectiveness in cough due to cold or inhaled irritants (see section V.A.1. of this Decision). The Tebrock study is similarly deficient because the record given to each patient instructed the patient to gargle each spoonful of the drug 15 seconds before swallowing (P9–3). (WL/PD's medical interpretation states, however, that each dose was to be gargled for 10–15 seconds before swallowing (P9–4, p.2).)

15. *Effectivenes of Benylin Compared to Codeine.* The Bureau takes exception to the statement that the Burke study (Protocol 286–9) "demonstrated Benylin to be 8 percent more effective than the known antitussive codeine." The Bureau points out that the positive control in the study was Benylin plus codeine, and that no expert witness testified that Benylin is more effective than codeine (Exception B.20, pp. 61–62, citing ID, p. 15). WL/PD replies that the finding is justified by the results of the study.

I agree with the Bureau that a study comparing Benylin to Benylin plus codeine does not support a finding concerning the effectiveness of codeine. Addition of Benylin to codeine creates a distinct combination drug whose effects may be different from those of codeine alone. It is possible that Benylin combined with codeine reduced the codeine's effectiveness (see section V.A.4. of this Decision).

B. *Safety*

1. *Finding of Safety Based on Evidence Other Than Studies.* The Bureau contends that the Initial Decision's factual finding with respect to safety cannot sustain the ultimate determination to approve the supplemental NDA (Exception A.1., pp. 7–9, citing ID, p. 5). The Initial Decision found that the extensive use of Benylin for over 30 years and FDA's approval of prescription Benylin "is a substantial indication that it can be used safely" (ID, p. 5). The Bureau contends that, under the act, it is not enough to have a "substantial indication of safety" and that, accordingly, this factual finding compels disapproval of the supplemental NDA. The Bureau also argues that the Initial Decision relies improperly upon marketing experience, poison control reports, DAWN data, and the agency's approval of the drug for prescription use as the basis for concluding that Benylin is safe for OTC use (Exception A.4., pp. 21–23, citing ID, pp. 5, 9, 19–20). WL/PD contends that the Initial Decision's factual finding as to Benylin's safety is adequate to support an approval decision and that marketing experience with Benylin demonstrates its safety for OTC use (Reply, pp. 8–10, 13–17).

I believe that the ALJ statement that there is "substantial indication of a drug's safety" is merely his way of saying that he believed the statute's criteria with respect to new drug safety are satisfied. However, I do agree with the Bureau's position that neither extensive use of Benylin for over 30 years nor FDA's aproval of prescription Benylin is sufficient evidence on the basis of which to find a prescription new drug to be safe for OTC use.

The Initial Decision does not apply correctly the principles governing the type of data that may be considered to determine the safety of a new drug, including the safety in OTC use of an approved prescription new drug. Since 1938 the law has required that decisions concerning the safety of new drugs be based upon tests. Specifically, the primary evidence demonstrating the safety of a new drug is required to be "investigations" * * * [that] include adequate tests by all methods reasonably applicable to show whether or not [a new] drug is safe for use under the conditions prescribed, recommended, or suggested in the proposed labeling thereof" (section 505(d)(1) of the act). If OTC use is to be a condition prescribed, recommended, or suggested in the labeling, then safety for OTC use must be demonstrated (G–5, p. 6).

FDA may appropriately consider reports of marketing experience, poison control data, and DAWN data as additional safety evidence, especially when a drug has a long marketing history (see, e.g., FDA's OTC drug regulations, 21 CFR 330.10(a)(4)(i)). However, evidence from these secondary sources is merely corroborative and cannot substitute for adequate safety tests. I believe that the Initial Decision does not sufficiently consider safety tests (certainly, none are cited) and relies excessively upon evidence from other sources.

The Bureau argues that the phrase "results of significant human experience during marketing" in § 330.10(a)(4)(i)) (21 CFR 330.10(a)(4)(i) means "adequate tests" (Brief, p. 377). I do not agree with this argument. Nevertheless, the reports in the record on Benylin's marketing history are not of sufficient quality to corroborate a finding, based on the studies, either that Benylin is safe, or that it is unsafe, for OTC use.

Reports of marketing experience, poison control data, and DAWN data are too fragmentary, incomplete, and nonrepresentative to serve as the basis for a conclusion that a drug is safe (G–31, pp. 10–13; G–82, G–90; T–501–07, 612–13, 651, 653–54). For example, poison control data and DAWN data would be expected to include reports of at least a fraction of the overdoses and abuse that may occur, but not reports of drug-induced drowsiness (G–21, pp. 21–22; T–501, 649–651). Similarly, physicians may not report to the manufacturer or to FDA a type of reaction that is already well-known, e.g., drowsiness from Benadryl or from Benylin (G–15, pp. 18–19; G–21, pp. 21–22; G–31, p. 13; T–505, 649, 655). Patients may not associate their involvement in an automobile accident with their having taken a drug (G–15, p. 19; T–501). Adverse reaction reporting systems in other countries such as Canada, where Benylin is sold OTC, may be designed to gather information from physicians concerning prescription drugs, rather than information from consumers on OTC drugs (T–503). Accordingly, a paucity of reports about Benylin in adverse reaction reports, poision control statistics, and DAWN data cannot be treated as proof that Benylin is safe for OTC use.

The Bureau's concern about patient compliance with warnings about drowsiness had led it, since 1948, to limit Benylin to prescription. As stated above in section V.B. of this Decision, the risks presented by the drug do not seem sufficient to warrant continued restriction to prescription as a way of gaining the added protection of physician communications. In contrast to the Initial Decision (pp. 5, 9, 19–20), this view is not based on the agency's approval of Benylin as safe for prescription use, or on lack of legal authority, but on my assessment of the safety studies and other evidence in the record as well as policy considerations about the comparative benefits of prescription and OTC status (see sections V.B. and VI.B.2. of this Decision).

2. *FDA's Authority to Restrict Drugs to Prescription Use.* The Bureau maintains that the Initial Decision adopts an an impermissibly narrow interpretation of the agency's authority to restrict drugs to prescription use under section 503(b) of the act. (Exception A.2., pp. 10–17 citing ID, p. 6–9). The Bureau argues that the Initial Decision's interpretation of section 503(b) is inconsistent with its plain language, its legislative history, and judicial interpretation. WL/PD replies that the Initial Decision interprets section 503(b) correctly (Reply, pp. 10–12).

The Initial Decision discussed FDA's authority under 503(b) of the act as follows (ID at 6–8):

> The Bureau's line of argument [that restricting Benylin to prescription use will enhance its safety because consumers show extra care in following directions for prescription drugs] does not take into account the limited scope of agency authority under the Federal Food, Drug, and Cosmetic Act. The lengthy House and Senate debates which culminated in the expansion of § 503(b) reveal a strong distrust of administrative power in this area [97 Cong. Rec. 2312 (1951), 97 Cong. Rec. 9235 (1951), 97 Cong. Rec. 13128 (1951)].
>
> * * * * *
>
> It is clear from these debates that the main purpose of the legislation was to put the pharmacists on notice as to the legal status of a drug * * *. Even a cursory review of the legislative history of these portions of the

statute reveals that this legislation was not directed toward expanding the powers of the Commissioner to include his dictating prescription status of drugs.

* * * * *

The use of prescription requirements to insure patient compliance with non-technical label restrictions which do not require any medical expertise for their comprehension is not what Congress intended when it revised § 503(b) [21 U.S.C. 353(b)].

I agree with the Bureau's contention that the Initial Decision has adopted an overly restrictive view of the agency's authority to limit a drug to prescription use. The plain language of section 503(b)(1)(B) of the act sets forth four separate grounds for limiting a drug to prescription dispensing. It provides for a prescription limitation whenever (1) the method of use, (2) collateral measures necessary for use, (3) toxicity or, (4) other potentiality for harmful effect of a drug justifies a conclusion that it is not safe for unrestricted self medication. The Initial Decision effectively reads out of the statute any assessment of a drug's "other potentiality for harmful effect" by holding that FDA's inquiry is limited to determining whether "non-technical label restrictions" can be comprehended by laymen.

The obligation of the manufacturer and the agency to the public demands an inquiry more extensive and searching than mere label reading to determine whether a drug may be safely used by all groups in the general population. The public has a justifiable expectation that, within the limits of scientific knowledge, significant uncertainties about a drug will be resolved before general marketing commences. In this respect, it is entirely reasonable for the agency to satisfy itself that label restrictions are not only understandable, but likely to be followed. Unlike an umpire passively calling balls and strikes, FDA has the responsibility to require an affirmative showing that patients will comply with label instructions and that failure to observe those instructions does not place consumers in significant jeopardy. Certainly, in the case of a soporific drug, hazards that ensue from failure to observe labeling directions may threaten the patient and others. The potentiality for harm to the operator of a motor vehicle or heavy machinery, as well as to others in his vicinity, is a serious concern and cannot be ignored when assessing the question of prescription status under section 503(b)(1)(B).[14]

The Bureau's argument that the Initial Decision incorrectly relied upon selected portions of legislative history is persuasive. When the purpose of a congressional enactment "has been effected by plain and unambiguous language, and the act is within the power of Congress, the only duty of the courts is to give it effect according to its terms." *United States* v. *Lexington Mill Co.*, 232 U.S. 399, 409 (1914); see *TVA* v. *Hill*, 437 U.S. 153, 173, 187, 193–195 (1978). The only circumstance in which a statute may properly be construed to mean something other than what it plainly says is where a literal reading " 'would lead to absurd results * * * or would thwart the obvious purpose of the statute.' " *Trans Alaska Pipeline Rate Case*, 436 U.S. 631, 643 (1978), quoting *Commissioner* v. *Brown*, 380 U.S. 563, 571 (1965). Implied exceptions to clearly delineated statutory coverage are disfavored and will not be found unless they are essential to avoiding an obvious inconsistency within the statutory scheme. *United States* v. *Key*, 397 U.S. 322, 324–325 (1970); see *TVA* v. *Hill, supra*, 437 U.S. at 188; 2A *Sutherland, Statutes and Statutory Construction*/47.11 at 90 (C. Sands 4th ed. 1973).

When the proper construction of a statute is in doubt, there is a special principle applicable to public health legislation. It is "the well-accepted principle that remedial legislation such as the Food, Drug, and Cosmetic Act is to be given a liberal construction consistent with the act's overriding purpose to protect the public health * * *." *United States* v. *An article of Drug* * * * *Bacto-Unidisk*, 394 U.S. 784, 798 (1969).

Whether one accepts the plain meaning of the words "other potentiality for harmful effect" in section 503(b)(1)(B) or applies them in the manner that best promotes the remedial purposes of the act, there is no justification for relying upon selected portions of congressional debates to reach a conclusion that section 503(b)(1) should be read less expansively than other sections of the same statute. A policy of liberality in the construction of the act certainly does not justify carving out exemptions to its coverage; on the contrary, a liberal approach to interpreting the statute should serve to carry out its central purpose. *United States* v. *Dotterweich*, 320 U.S. 277, 284 (1943).

Congressional debates of the type relied upon in the Initial Decision are generally accorded little, if any, weight. *National Welfare Rights Organization* v. *Mathews*, 533 F.2d 637, 642–643 (D.C. Cir. 1976); *Castaneda-Gonzalez* v. *Immigration and Naturalization Service*, 564 F.2d 417 (D.C. Cir. 1977); *Warner* v. *Dworsky*, 194 F.2d 277, 279 (8th Cir. 1952). See 2A Sutherland, *Statutes and Statutory Construction* /48.13 (C. Sands 4th ed. 1973). The reason for this reluctance is apparent from a reading of the debates on the Durham-Humphrey Amendments.

Although the Initial Decision states that the debates evidenced "a strong distrust of administrative power in this area," examination of the debates reflects primarily a conventional legislative concern that the statute contain reasonable substantive standards and procedural safeguards. Certain speakers fearing socialized medicine and over-regulation spoke out against the bill. Their comments are for the most part stated in the minority report attached to H.R. Rept. 700, 82d Cong., 1st Sess., pp. 28–37 (1951). Their rhetoric in floor debate focused on Oscar Ewing, then Federal Security

[14] In addition to the authority provided in section 503(b)(1)(B), there appears to be a wholly independent legal basis for requiring a new drug—in this case Benylin—to be dispensed only upon the prescription of a physician. Section 503(b)(1)(C) authorizes prescription status for a drug that "is limited by an approved application under section 505 to use under the professional supervision of a practitioner licensed by law to administer such drug." A drug that satisfies the criteria of either paragraph (A) "or" paragraph (B) "or" paragraph (C) of section 503(b)(1) may be limited to prescription status. The committee reports accompanying this legislation show that a prescription requirement imposed under the authority conferred by section 503(b)(1)(C) on a drug subject to section 505 is independent of the prescription drug classification effected under section 503(b)(1)(B). See S. Rept. 946, 82d. Cong., 1st Sess., pp. 4, 81; H.R. Rept. 700, 82d. Cong., 1st Sess., pp. 6–7, 10 (1951). Indeed, in enacting section 503(b)(1)(C), Congress was codifying what had been, and what it understood to be, existing practice of the FDA in regard to new drugs. As the Senate report noted, "There is no controversy whatever about * * * new drugs restricted to prescription sale by effective new drug applications under section 505 of the *present statute*." S. Rept. 946, 82d. Cong., 1st Sess., p. 4 (1951) (emphasis added). Section 503(b)(3) provides that the Secretary may by regulation remove drugs subject to section 505 from the prescription dispensing requirements "when such requirements are not necessary for the protection of the public health." Because, in practice, FDA has not established different criteria for specifying when prescription requirements for new drugs are "necessary for the protection of the public health," the general prescription drug criteria in section 503(b)(1)(B) also apply to decisions restricting a new drug to prescription use under section 503(b)(1)(C) (see 21 CFR § 310.200(b)). It appears that the principal purpose of section 503(b)(1)(C) and (3) is evidentiary, i.e., to create a presumption that a new drug is restricted to prescription, except where exempted by regulation, and thus to simplify the burden upon the agency in an enforcement action against a new drug that is being sold without a required prescription. These provisions also reinforce my position that WL/PD, as a new drug aspplicant, has the burden of establishing Benylin's safety for OTC use (see section VI.B.4. of this Decision).

Administrator,[15] as the object of their stated antipathy. Mr. Ewing was a controversial and, in some circles, unpopular public figure; the debates presented an opportunity to attack him [97 Cong. Rec. 9533 col. 1 (1951); 97 Cong. Rec. 9547 col. 1 (1951); 97 Cong. Rec. 9549 col. 2 (1951)]. Reliance on statements made in debate are equally available to support a broad, liberal interpretation of section 503(b)(1), e.g., the statement that section 503(b) was "plain and broad and [gave] the Administrator all the power he needs * * *" 97 Cong. Rec. 9547 col. 2, (1951). The problem with statements made in debate is that they generally do not reflect the collective will of Congress.

Legislative consensus is reflected in legislation as finally passed and in authoritative committee reports on that legislation. The committee reports pertinent here show that Congress was concerned with increasing public protection. In explaining the objectives of the bill, the House Report stated that it was intended "To *strengthen* the protection of the public health against dangerous abuses in the sale of potent prescription drugs * * *" H.R. Rept. No. 700, 82d Cong., 1st Sess. p. 2 (1951) (emphasis added). As explained elsewhere in the report, the legislation was considered (Id. at 7):

> * * * important to the enforcement agency because it permits more effective enforcement through appropriate control over drugs that are too dangerous, or otherwise unsuitable, to be used by a layman without medical diagnosis or supervision. Under this proposed legislation it will be possible to prevent injury to the public, as contrasted with the present system which is largely concerned with puniahing past violations.

The Senate Report also stated that the legislation was designed "to protect the public," S. Rept. No. 946, 82d Cong., 1st Sess., p. 1 (1955), and explained:

> The word "safe", as used in the definition, is intended to have its ordinary meaning. * * * The language of the definition clearly shows that toxicity is only one factor to be considered by the courts in determining whether a particular drug is safe for use without medical supervision. *The definition requires the Court to consider other potentialities for harmful effect * * * When this language is given judicial interpretation consistent with the over-all purpose of the Federal Food, Drug, and Cosmetic Act to protect the public health it will effectively restrict to prescription sale all drugs that require professional supervision for their use.* (Emphasis added.)

[15] At the time, FDA was a component of the Federal Security Administration, the predecessor of the Department of Health, Education, and Welfare.

My conclusion is also guided by judicial opinions construing section 503(b)(1) of the act. These opinions arose from challenges by the National Nutritional Foods Association (NNFA) to FDA's regulations classifying high dosages of vitamins A and D as prescription drugs. See *National Nutritional Foods Ass'n* v. *Weinberger*, 366 F. Supp. 1314 (S.D. N.Y. 1973), *aff'd*, 491 F.2d 845 (2d Cir. 1973); 376 F. Supp. 142 (S.D. N.Y. 1974), *remanded*, 512 F.2d 688 (2d Cir. 1975), *cert. denied sub nom., National Nutritional Foods Ass'n* v. *Mathews*, 418 F. Supp. 394 (S.D. N.Y. 1976), *rev'd* 557 F.2d 325 (2d Cir. 1977). For reasons that are not apparent, these *NNFA* cases were not discussed in the Initial Decision. Their relevance here, however, is inescapable.

The position adopted in the Initial Decision is similar to the argument that was advanced by the National Nutritional Foods Association and rejected in the *NNFA* cases.[16] In the three initial *NNFA* cases, the reviewing courts held that section 503(b)(1) affords the agency broad authority to restrict drugs to prescription. In denying a motion to enjoin enforcement of the vitamin A and D regulations, the district court held that the statutory term "potentiality for harmful effect" plainly "imposes broad responsibility on the Commissioner to safeguard human health." 366 F. Supp. at 1346, *aff'd*, 491 F.2d 845. In later opinions both the district court and appellate court held that, although the legislative history of the prescription drug provisions is not entirely clear, 376 F. Supp. at 143, 512 F.2d at 698, it does not detract from the conclusion that the agency's power to limit drugs to prescription is quite broad, 376 F. Supp. at 144–46, 512 F.2d at 699.

In sum, Congress lef it to FDA to decide whether medical supervision will help ensure patient compliance with labeling restrictions needed for safe use of a drug. FDA has authority under section 503(b)(1) to continue to restrict Benylin to prescription use, if it concludes that such medical supervision is warranted. But the agency is not compelled to continue this restriction if it determines that there is sufficient evidence to show that the drug is safe for OTC distribution. See section V.B. of this Decision, above.

[16] The Court of Appeals for the Second Circuit held in its final decision that there was not adequate support in the record for FDA's classification of high dosages of vitamins A and D *as drugs at all*, rather than as foods. On the basis of that court's earlier opinions there is little doubt that if the court had held that FDA could regulate the articles *as drugs at all*, it would have held that there was ample evidence to support restricting them *to prescription use*.

3. *FDA's Authority to Consider Use of a New Drug Contrary to Labeling.* The Bureau maintains that the Initial Decision's strict reliance on the term "under the conditions of use prescribed, recommended, or suggested in the proposed labeling thereof" in section 505(d) of the act is erroneous (Exception A.3., pp. 17–20, citing ID, p. 7). The Bureau contends that the use of a drug without the supervision of a physician is a condition of use within the meaning of section 505(d). WL/PD replies that the Initial decision interprets the act correctly (Reply, pp. 12–13).

The Initial Decision discusses FDA's authority under section 505(d) of the act as follows (ID, pp. 7–8):

> In the section of the statute which sets forth the factors to be considered in the evaluation of new drug applications [§ 505(d)(1)–(6)], each subparagraph includes the provision that the safety of the drug is to be evaluated in terms of its use "under the conditions prescribed, recommended, or suggested in the proposed labeling thereof." For the evaluation of a drug to go beyond this scope, there must be a strong showing that label requirements will be disregarded or cannot practicably be complied with and that the safety implications of the resulting use outside the labeling requirements would require the [Commissioner] to broaden his investigation. * * * In the case of Benylin, however, it has not been demonstrated that the side effects are so dangerous as to require additional administrative controls.
>
> * * * * *
>
> Therefore, questions of the safety of a drug when it is used in a manner contrary to the label warnings can be considered in only limited circumstances. Once the manufacturers have shown the drug to be safe under the conditions prescribed, recommended or suggested in the labeling, they have fulfilled their statutory requirement and the drug is eligible for certification. Consideration of safety evidence relating to uses of a drug outside of the label requirements necessitates a showing that there is a reasonable probability that such non-label indicated uses can be expected to occur. The mere allegation of potential injury resulting from noncompliance with label requirements does not result in the manufacturers being required to submit evidence to disprove such a claim. To support its claim, the Bureau would have to show convincing evidence of harm rather than presenting what is a mere supposition of possible harm. Such a showing has not been made on the record of this case.

I agree with the Bureau. The Initial Decision misconceives the intent of section 505 to require those who would market drugs to demonstrate that the drugs they distribute will be safe for use under circumstances that may reasonably be expected. FDA may require of new drug applicants who propose labeling a convincing showing that proposed labeling will be followed

by users and that the consequences of a failure to follow the labeling will not represent a significant risk to the user and the public. Consumers of OTC drugs would not be well served if the Bureau were required to shoulder the burden of showing that labeling requirements will be disregarded or cannot practicably be followed. It is sufficient for the agency to raise reasonable questions about labeling compliance. At that juncture, the burden of resolving those questions must be carried by the new drug applicants.

The provisions of section 505 do not dictate or support the position adopted in the Initial Decision. Section 505(b)(1) requires the new drug applicant to submit "full reports of investigations which have been made."

The provisions of section 505 do not dictate or support the position adopted in the Initial Decision. Section 505(b)(1) requires the new drug applicant to submit "full reports of investigations which have been made to show whether or not such drug is safe for use * * *" Such reports are not limited by this section to investigations under the conditions of use prescribed, recommended, or suggested in the proposed labeling. Section 505(b)(6) requires the submission of "the labeling proposed to be used for such drug." Approval of the application must be denied under section 505(d)(1) when the reports of investigations "do not include adequate tests by all methods reasonably applicable to show whether or not such drug is safe for use under the conditions prescribed, recommended, or suggested in the labeling * * *." Given the laboratory and clinical data derived from appropriate tests, other questions must also be answered: (1) Can the proposed labeling be comprehended; (2) if so, will it be followed; and (3) if not, what are the consequences for patient safety of failure to follow the labeling? These questions are implicit in any thorough evaluation of an application under section 505(d)(1).

The same questions arise from the scrutiny of a drug under section 505(d)(4), which provides, in pertinent part, for the disapproval of a drug when the FDA has "insufficient information to determine whether such drug is safe for use under such conditions * * *" Even if the FDA were required to limit its inquiry to the labeling a manufacturer proposed under section 505(d)(1), it is apparent that Congress sought through section 505(d)(4) to provide the agency with some authority not granted elsewhere in section 505(d). Otherwise, one of the grounds for disapproval would be a statutory redundancy. Applying the cardinal rule that "significance and effect shall, if possible, be accorded to every word" in a statute and "if it can be prevented, no clause, sentence, or word be superfluous, void, or insignificant," *Ex Parte Public Bank*, 278 U.S. 101, 104 (1928), it is only reasonable to conclude that an inquiry into labeling compliance is authorized under section 505(d)(2) or (4). See also: *Jaracki* v. *G. D. Searle & Co.*, 367 U.S. 303, 307, (1961); *Ginsberg & Sons, Inc.* v. *Popkin*, 285 U.S. 204, 208 (1932).

The conclusion that, in determining the safety of a drug, FDA must consider the totality of patient exposure to the drug, and not just use within the boundaries of the proposed labeling, was also reached in *National Nutritional Foods Ass'n* v. *Weinberger*, 366 F. Supp. 1341, 1346 (S.D. N.Y. 1973), *aff'd* 491 F.2d 845 (2d Cir. 1973). There, Judge Frankel explained that the FDA has the authority and responsibility to protect those consumers who do not heed label instructions, as well as those who do:

> * * * The Commissioner is not only entitled but required to be at least as prudent as lay understanding of human behavior would dictate * * *. He knows, because we all do, of the prevalent wisdom among us pill swillers that "if one is good, two are better." * * * Knowing such obvious things and other matters less patent, the Commissioner is not required to circumscribe his responsibilities in terms solely of the completely "rational" consumer. He may and he must consider the "potentiality for harmful effect" through excessive use to the merely "average" man and even to the substantial numbers of us who help create the average by falling below it.
>
> * * * * *
>
> The Commissioner is not required to set the over-the-counter limit at a maximum which the consumer (be he old, young, weak or strong) might withstand. See *United States* v. *Bodine Products*, 206 F. Supp. 201, 207 (D. Ariz. 1962). A substantial margin of safety probably should—certainly may—be used, at least in a case like this, where it can have only beneficial effects for the Commissioner's paramount subject of health.
>
> "One making a rule for the future which in practical effect will determine whether millions of people shall eat something every day may reasonably refuse to subject the general public to even slight risks and small deceptions." *Atlas Power* v. *Ewing*, 201 F.2d 347, 355 (3d Cir. 1962).

The Second Circuit, although remanding the case on different grounds, adopted the district court's reasoning (512 F.2d at 704):

> We reject petitioners' contention that because the higher dosage levels are not "inherently" unsafe but become unsafe only if used in violation of cautionary labelling (sic) they do not qualify as "prescription" drugs within the meaning of § 503(b) of the Act, 21 U.S.C. § 353(b). The broad language of § 503(b)(1)(B) [footnote omitted] permits consideration of the various factors surrounding the use of a particular drug in determining that it "is not safe for use except under the supervision of" a physician. There was ample evidence before the FDA that Vitamins A and D, when consumed in large quantities over a period of time, can be acutely toxic. It was reasonable for the Commissioner to recognize that the risks of toxicity are increased by over-the-counter availability of readily ingestible, high dosage forms, and therefore he could rationally conclude that these forms have a "potential harmful effect."

The court suggested that FDA might properly decide not to restrict a drug to prescription use in "a case in which there was the mere possibility that a drug would be occasionally misused," in contrast to a case "in which the actual way in which the product is apparently used on a normal basis by many persons presents serious risks of toxicity." ID.

These opinions recognize FDA's broad authority under section 503(b)(1) of the act to protect consumers from "potentiality for harmful effect" due to uses contrary to label warnings. Whether FDA acts to restrict a drug to prescription status under section 503(b)(1)(B) alone or through the new drug procedures set forth in section 505 of the act, FDA believes that it is not limited to a consideration of the safety of a drug under an assumption that the proposed labeling will always be followed; the principle that FDA is authorized to consider possibilities and consequences of use outside the labeling remains operative.

In requiring the Bureau to "make a strong showing" that harm from uses outside the labeling will occur, the Initial Decision places on the Bureau a burden that is administratively infeasible and contrary to the agency's mandate under the new drug provisions of the act. In the usual case in which FDA is considering whether to approve an NDA, factual evidence concerning a drug's potential for causing harm due to uses outside the labeling simply would be unavailable because the drug would not yet be on the market. FDA's decisions that the potential effects of drug use may warrant restricting a drug to prescription use—or may require, in some cases, disapproval of the NDA on grounds that the drug is unsafe—are necessarily based upon FDA officials' scientific judgment and experience. Even in the case of previously marketed drugs that are the subject of supplemental applications, it would not be a wise use of public funds for FDA to have to conduct the tests or surveys necessary to make "a factual showing of

significant noncompliance with * * * labile warnings" (ID, p. 8). Furthermore, a general requirement for such tests or surveys conducted by FDA would be contrary to the agency's mandate under the new drug provisions of the act to provide an "authoritative review of the manufacturers' tests," and not to duplicate tests that manufacturers are supposed to carry out. H.R. Rept. No. 75-2139, 75th Cong. 3d. Sess. (1938). Individual manufacturers are better able than FDA to conduct tests that address concerns about possible dangers from their drugs.

4. *Burden of Proof Concerning Justification for Prescription Status and Harm From Use Outside Labeling.* The Bureau contends that the Initial Decision improperly places on the Bureau the burden of demonstrating that prescription status is justified because it will enhance safety (Exception at A.5., pp. 23–25, citing ID at 5, 20). Similarly, the Bureau argues that the Initial Decision improperly places on the Bureau the burden of showing that "label requirements will be disregarded or cannot practicably be complied with and that the safety implications of the resulting use outside the labeling requirements would require the [Commissioner] to broaden his investigation" (Exception A.1., pp. 6–9, Exception A.6., pp. 25–27, citing ID, p. 7). The Bureau contends that, even if it has some burden to come forward with evidence on these matters, the Bureau has met this burden by evidence in the record. WL/PD claims that the Bureau has misinterpreted the Initial Decision (Reply, pp. 1820). The company argues that it has successfully carried its burden of proof in showing the safety of Benylin in OTC distribution and believes that the Bureau properly has the burden of coming forward with evidence to contradict the company's evidence (Reply, p. 18).

For the reasons given in section VI.A.1. and B.3. of this Decision, I agree with the contention that the Initial Decision improperly places on the Bureau the burden of demonstrating that prescription status is justified and of showing that the proposed label warnings on Benylin will be disregarded. It would have been better for the Initial Decision's conclusion concerning Benylin's safety to have been based on the evidence of record, rather than on a technical, legal approach of resolving the safety issue against the party to whom the Initial Decision, under a novel theory, assigns the burden of proof.

In this case, I believe that the Bureau has satisfied its burden of coming forward with evidence or arguments concerning the difference between the use of a drug under the supervision of a physician and the unsupervised use of an OTC drug (G–3, pp. 5–7; G–5, p. 6; G–9, p. 9; G–11, pp. 24–25; G–15, pp. 17–18; G–17, p. 18; G–19, p. 5; G–21, p. 22; G–31, pp. 13–14, 16–17; G–49, pp. 213–16; T–636, 647, 668–70, 842). Similarly, I believe that the Bureau has satisfied its burden of coming forward with evidence on the harm that has occurred, or may occur, due to disregard of warnings of the type proposed for inclusion in Benylin's labeling (G–3, pp. 4–7; G–5, pp. 5–6; G–9, pp. 5–6, 9–10; G–11, pp. 24–25; G–15, p. 18; G–19, pp. 4–5; G–21, pp. 12–13; G–31, pp. 7–11, 14; G–49; G–65; T–668–70).

As discussed above in section V.B. of this Decision, I also believe that there is substantial evidence to support a finding either that WL/PD has nevertheless—or that it has not—overcome the Bureau's evidence and arguments and met the company's burden of persuasion.

5. *Hazards Associated With Use of Benylin.* The Bureau takes exception to the statement that the "Bureau asserts that accidents which occur due to drowsiness are the only hazard associated with the drug" (Exception A.7. pp. 27–29, citing ID, p. 6). The Bureau argues that the Initial Decision has misinterpreted a statement in the Bureau's motion to strike of September 6, 1977 (p. 26) and ignores evidence in the proceeding of other hazards associated with Benylin. WL/PD replies that the Initial Decision interprets the Bureau's position correctly (Reply, pp. 20–22). The company argues that the other alleged safety hazards from Benylin are speculative and, in any event, are adequately dealt with in the proposed labeling.

I agree that the Initial Decision appears to take out of context a statement in the Bureau's motion to strike. It is clear from the record that the Bureau's principal safety concern with Benylin is accidents that may occur due to drowsiness. See sections V.B.1., 3., and 4. of this Decision. However, it also is clear that the Bureau has other concerns with respect to the safety of Benylin. See section V.B.1. of this Decision.

I believe, however, as explained in section V.B.1. of this Decision, that consideration of these other hazards does not affect the outcome of this proceeding. These other hazards, although not trivial, do not justify a decision to continue restricting Benylin to prescription use. Many current OTC drugs present risks similar to these, and such risks have been addressed adequately by labeling and child resistant packaging.

6. *Initial Decision's Lack of Findings and Conclusions.* The Bureau contends that the Initial Decision is defective because it fails to include "findings and conclusions, and the reasons or basis therefor," 5 U.S.C. 577(a)(3)(A), to support the ultimate conclusion that "[t]he report of investigations which have been made to show whether or not Benylin is safe for use demonstrates the safety of the drug as required under § 505(d)(1)–(6) of the Act" (Exception A.8., pp. 30–31, citing ID, p. 9). The Bureau contends that, in the absence of specific findings and reasoning that the drug has been shown safe "by adequate tests by all methods reasonably applicable" (§ 505(d)(1)), there is no basis for the ultimate conclusion that the supplemental NDA should be approved. WL/PD replies that the Initial Decision's conclusion that Benylin is safe is a sufficient finding, based upon the manufacturer's submitted studies (Reply, pp. 23–24).

Although it would have been desirable for the Initial Decision to include findings respecting the various safety studies of Benylin, I reject the exception because I believe that the Initial Decision satifies the requirements of the Administrative Procedure Act. A narrative presentation of findings of fact and conclusions of law, as incorporated in the Initial Decision, is permissible and sufficient under the Administrative Procedure Act (5 U.S.C. 577(d)). *Gilbertsville Trucking Co.* v. *United States*, 196 F. Supp. 351 (D. Mass. 1961); *State Corporation Commission* v. *United States*, 184 F. Supp. 691 (D. Kan. 1959). FDA's regulations governing this proceeding, 21 CFR 12.120, do not establish a mandatory form for the required findings of fact and conclusions of law. Moreover, in this Decision I am making my own findings and conclusions (incorporated in the text of this Decision) with respect to the safety of Benylin.

7. *Benefit-to-Risk Ratio of Benylin.* The Bureau criticizes the Initial Decision's failure to analyze the benefit-to-risk ratio of Benylin as required by 21 CFR 330.10(a)(4)(iii) (Exceptions, pp. 31–32). In its reply, WL/PD disagrees with the exception (Reply, pp. 24–26).

I disagree with the exception. It appears from the Initial Decision that the ALJ did consider the benefit-to-risk ratio of Benylin in reaching his conclusion that the drug had been shown safe and effective (ID, pp. 19–21). He simply came to a conclusion different from that proposed by the Bureau.

C. General Recognition of Safety and Effectiveness

1. *Sufficiency of Evidence to Determine General Recognition.* Among other things, the Bureau excepts to the finding that the evidence of record in this proceeding is insufficient to determine whether Benylin is generally recognized as safe and effective (Exception C.2, pp. 62–63, citing ID, p. 21). WL/PD disagrees (Reply, p. 48).

I agree with this exception. As stated in section V.C. of this Decision, above, the record contains substantial evidence that Benylin is not generally recognized as safe and effective.

2. *Standard for General Recognition.* The Bureau also disagrees with the reference to such recognition "in the opinion of the medical community," rather than "among [qualified] experts" as specified in section 201(p) of the act (Exception C.3, pp. 62–63, citing ID, p. 21). WL/PD disagrees (Reply, p. 48).

I agree with the exception. The Initial Decision does not state correctly the statutory standard for general recognition of a drug's safety and effectiveness.

Other exceptions concerning the Initial Decision's findings with respect to general recognition of safety and effectiveness have been addressed by this Decision.

VII. Conclusion

Based on the foregoing findings, conclusions, and discussion, I reverse the Initial Decision and conclude that:

1. WL/PD has not shown that diphenhydramine hydrochloride acts to inhibit activity in the brain's cough center. See section V.A.1.

2. In the absence of a showing that diphenhydramine hydrochloride suppresses activity in the brain's cough center, Benylin's effectiveness as an antitussive drug for use in cough due to colds may be established only by two or more studies in the target population. See section V.A.1.

3. The two studies of Benylin in patients with coughs due to cold (Tebrock study, Protocol 266–17, and Burke study, Protocol 266–9) are not adequate and well-controlled investigations, as defined in section 505(d) of the act (21 U.S.C. 355(d)) and § 314.111(a)(5)(ii) of the regulations (21 CFR 314.111(a)(5)(ii)). Accordingly, there is a lack of "substantial evidence" as that term is defined in section 505(d) of the act that Benylin will have the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the proposed labeling thereof. See sections V.A.2. through 4 and VI.

4. Because it has not been shown that Benylin is effective, I cannot now find that WL/PD has satisfied the requirements for establishing its safety. See section V.B.

5. Benylin is not generally recognized, among experts qualified by scientific training and experience to evaluate the safety and effectiveness of drugs, as safe and effective for use under the conditions of use prescribed, recommended, or suggested in the proposed labeling thereof. Accordingly, Benylin is a new drug within the meaning of section 201(p)(1) of the act (21 U.S.C. 321(p)(1)). See section V.C.

The foregoing decision in its entirety constitutes my findings of fact and conclusions of law.

VIII. References

Pursuant to § 12.95 (21 CFR 12.95), I take official notice of the following references, which pertain to matters peculiarly within the general knowledge of FDA as an expert agency, to the extent they are relied upon in factual findings in this Decision that cite these references. If I am taking official notice of a material fact not appearing in the evidence of record, a participant, on submission of a timely petition for reconsideration under § 12.139 (21 CFR 12.139), will be afforded an opportunity to show the contrary.

The following information has been placed in the office of the Hearing Clerk (HFA–305), Food and Drug Administration, Rm. 4–65, 5600 Fishers Lane, Rockville, MD 20857, and may be seen by interested persons from 9 a.m. to 5 p.m., Monday through Friday.

1. Review Panel on New Drug Regulation, DHEW, "Risk, Safety, Efficacy and Benefit," Interim Reports, Vol. II.

2. Feinstein, A., "How Do We Measure Safety and Efficacy?" *Clinical Pharmacology and Therapeutics*, 12:544–556, 1971.

3. Hoff, L. C., "How Often Do Consumers Seek Your Advice on Rx ? & OTC Products?" *Pharmacy Times*, 163:52–55, 1975.

4. Svarstad, B. L., "Physician-Patient Communication and Patient Conformity with Medical Advice," *in* "The Growth of Bureaucratic Medicine," John Wiley and Sons, New York, 220–238, 1976.

5. Ley, P. and M. S. Spelman, "Communications in an Out-patient Setting," *British Journal of Social Clinical Psychology*, 4:114–116, 1965.

6. Joyce, C. R. B., et. al., "Quantitative Study of Doctor-Patient Communication," *Quarterly Journal of Medicine*, 150:183–194, 1960.

7. Bain, D. J. G., "Patient Knowledge and the Content of the Consultation in General Practice," *Medical Education*, 11:347–350, 1977.

8. Ley, P., "Psychological Studies of Doctor Patient Communication" in Rachman, S. (ed.), *Contributions to Medical Psychology*, 9–42, 1977.

9. Hulka, B. S., et al., "Doctor-Patient Communication and Outcomes Among Diabetes Patients," *Journal of Community Health*, 1:15–27, 1975.

10. Decastro, F., "Doctor-Patient Communication," *Clinical Pediatrics*, 11:86–87, 197[illegible].

11. Marston, M. V., "Compliance with Medical Regimens: A Review of the Literature," *Nursing Research*, 19(4):312–323, 1970.

12. Boyd, J. R., et al., "Drug defaulting, part 1: determinants of compliance," *American Journal of Hospital Pharmacy*, 31:362–367, 1974.

13. Sackett, D. L., "The Magnitude of Compliance and Noncompliance" *in* "Compliance with Therapeutic Regimens," 9–25, 1976.

14. Barofsky, I., "Chronic Psychiatric Patient in the Community: Principles of Treatment," Spectrum, Wiley Press: in press.

15. Morris, L. A. and J. A. Halperin, "Effects of Written Drug Information on Patient Knowledge and Compliance: A Literature Review," in press, *American Journal of Public Health.*

16. Mantel, N., and W. Haenszel, "Statistical Aspects of the Analysis of Data From Retrospective Studies of Disease," *Journal of the National Cancer Institute*, 27:719–748, 1959.

Appendix A.—Effectiveness Studies of Benylin

The ALJ's Initial Decision included a summary of certain of the clinical studies intended to show Benylin's effectiveness. That summary is excerpted below:

1. Protocol 266–17 (P9–1 through P9–6) is a large, controlled, double-blind, randomized, multicentered study conducted by Dr. Tebrock involving 566 patients with cough due to cold. Patients were administered either Benylin or the Benylin vehicle (same formula as Benylin but lacks diphenhydramine) in dose of two teaspoons every four hours. Each morning during the study, patients were questioned as to the degree of improvement in their cough. Results were recorded on a questionnaire by the patient on each of the 3 days of treatment. They rated the degree of improvement on a four-point scale of "gone," "better," "same," or "worse." The results demonstrated an improvement of cough for the first day of 32.4 percent for Benylin and 23.1 percent for the Benylin vehicle. Results for day 1 showed a statistical significance of nine percent while the results for days 2 and 3 were not statistically significant.

2. Protocol 266–9 (P7–1 through P7–5) is a double-blind, positive-controlled study conducted by Dr. Burke involving 100 children with coughs due to the common cold. Patients were randomly assigned to receive either Benylin or Benylin combined with 20 mg of codeine for a total of 4 ounces of each preparation to be taken in four daily doses for 3 or 6 days depending upon whether 1 or 2 teaspoon doses were taken. Each child was questioned on symptom improvement, medication taste, and any adverse experiences at least once during the study. An adult closely associated with the child was present during the interview but the

study did not indicate whether the adult and/or child responded to the questions. Severity and frequency of cough were reduced in 58 percent of the Benylin group and 50 percent of the Benylin plus codeine group.

3. Protocol 266–18 (P6–6 through P6–9) is a randomized, double-blind, crossover study which used the inhalation of citric acid aerosol to induce artificial cough. The subjects were nine healthy, trained volunteers. The antitussive effects of Benylin over a 4-hour period were compared to the Benylin vehicle and to the Benylin vehicle plus 15 mg of codeine phosphate. The frequency and intensity of coughs were recorded by a pneumotachograph-strain gauge system and each hourly test period was compared against the control levels. * * *

4. Protocol 794–1 (P6–10 through P6–14) is a randomized, double-blind, crossover study which used the inhalation of citric acid aerosol to induce artificial cough. Ten healthy, trained volunteers compared four preparations in an inactive vehicle (1) diphenhydramine, (2) codeine phosphate, (3) diphenhydramine plus codeine phosphate 20 mg and (4) the inactive vehicle alone. The same methodology used in Protocol 266–18 was employed in this study. * * *

5. The 1960 Protocol of Bickerman (P6–3 through P6–5) is a double-blind, randomized, crossover study which used the inhalation of citric acid aerosol to induce artificial cough in nine healthy, trained volunteers. The antitussive activity of diphenhydramine 50 mg (twice the recommended dosage for Benylin) was compared with codeine 15 mg, with Ambodryl Hydrochloride (bromodiphenhydramine) and placebo. The same methodology used in Protocol 266–18 was employed in this study. * * *

Appendix B.—Safety Studies of Benylin

1. Protocol 266–15 (P2–6) was a double-blind, crossover study of 20 prisoner volunteers using 4 psychomotor function tests in which Benylin, containing 12.5 mg and 25 mg of diphenhydramine hydrochloride (diphenhydramine), was compared to the Benylin vehicle. In its narrative statement, WL/PD contends that this study is "of primary significance" among the company's 5 clinical pharmacology studies (others are described in paragraphs 3, 4, and 7 below) because "it was designed to measure objectively the effects of Benylin cough syrup on the psychomotor performance of patients receiving the drug" (P20, pp. 9–10). Subjects received medication four times daily for two days. The four tests used were (1) the digitspan test, a test of short-term memory used to evaluate a subject's ability to comprehend and carry out oral instructions; (2) the choice reaction time test, a measure of the time required to respond to a given stimulus; (3) the pursuit rotor test, a measure of time required to adjust a position of a vehicle to changing conditions on the road; and (4) the tracking time test, a variation of the pursuit rotor test. The results of this study showed no demonstrable impairment of psychomotor function produced by either dose of Benylin, compared to placebo.

The Bureau contends that the tests used in this study are not of sufficient sensitivity to demonstrate performance impairment, that these tests are not adequate to predict Benylin's effects in typical real life situations, that no drug plasma level determinations were made to measure accumulations of repetitive doses, and that the schedule of testing may have been biased against finding drug related effects (G–11, p. 25, G–19, p. 9; G–23, pp. 8–11; G–27, p. 20).

2. The Moskowitz study (G–23A, G–146) was submitted by the Bureau and involved 12 volunteers who were subjected to psychomotor function tests to examine the following aspects of behavior, which the investigator believed to be relevant to safe driving: tracking, perception, division of attention, and information processing rate (G–23, pp. 5–6). The study compared the effect on psychomotor function of diphenhydramine 50 mg (double the recommended antitussive dosage), alcohol, diphenhydramine 50 mg plus alcohol, and placebo. WL/PD disagrees with the statement in the results (G–146, p. 9); that the treatment administration was double-blind (G–146, p. 9); WL/PD disagrees (Brief, p. 12). The three tracking tests used were (1) a compensatory tracking task, in which the subject must maintain a fluctuating bar in a fixed position, (2) a divided attention task, in which the subject must do two tasks at once: locate and respond to the number "2" in a constantly changing, random array of 24 numerals by using a four-way lever to show the quadrant in which the number "2" appears, and operate with the other hand a lever that controls a compensatory tracking test as described in (1); (3) a critical tracking task, in which the subject must keep a horizontal line centered on a display by using a one dimensional control stick, a task that becomes more difficult with each success; and (4) the information processing task, in which the subject views a card containing four letters and must, after a longer interval in which random bits of letters are displayed, record the letters (G–23A, pp. 2–5).

In the compensatory tracking task, subjects taking diphenhydramine alone showed an impairment in performance compared to subjects taking placebo (G–23A, pp. 6–7). Those who consumed alcohol alone showed a greater impairment, and those taking alcohol plus diphenhydramine showed the greatest impairment. All three of these differenced from placebo were statistically significant. However, the mean error in subjects' ability to center the bar-graph display was 17 milliliters (ml) in the placebo group and 18 ml in the diphenhydramine alone group; WL/PD contends that this 1 ml difference is too minor to treat as significant (Brief, p. 13).

In the divided attention task, diphenhydramine alone and alcohol plus diphenhydramine resulted in statistically significant impairments in performance, compared to placebo (G–23A, p. 7). Alcohol alone caused an impairment that was not statistically significant. WL/PD contends that the test device is "devilishly complicated" and inapplicable to the operation of an automobile (Brief, p. 13). The report of the study states, on the other hand, that a divided attention task "is an experimental analogue of the attention-sharing demand[s] of actual driving, where attention must be shared between control of tracking of the vehicle upon the highway and surveying the environment for potential sources of danger" (G–146, p. 5).

In the critical tracking task, subjects taking alcohol alone or alcohol plus diphenhydramine showed statistically significant impairment in performance, compared to the placebo group (G–23A, p. 7). Subjects taking diphenhydramine alone showed a small impairment that was not statistically significant.

In the information processing rate task, the alcohol alone group and the alcohol plus diphenhydramine group showed statistically significant impairment in performance, compared to the placebo group (G–23A, p. 7). Diphenhydramine alone caused a small impairment that was not statistically significant, compared to placebo.

With respect to the latter two tasks, the investigator stated that it is possible that the tests were of too short duration to detect impairment due to diphenhydramine; subjects may have been able to force themselves to overcome drowsiness for such a short time (G–23A, p. 9; G–146, p. 16).

With respect to all tasks, WL/PD points out that no statistically significant difference between alcohol alone and diphenhydramine plus alcohol was observed (Brief, p. 14).

3. Protocol 184–15 (P2–10) was a double-blind, randomized, parallel group design study of 240 prison volunteers who received diphenhydramine hydrochloride 50 mg, diphenhydramine hydrochloride 25 mg, or placebo three times daily for seven days. The primary measurement used in the study was questioning of the subjects regarding any adverse reactions. Subjects were questioned daily during the study about occurrence of drowsiness. Over the seven day period, drowsiness was reported at least once by 41.3 percent of the placebo group, 50 percent of the diphenhydramine 25 mg group, and 53.8 percent of the diphenhydramine 50 mg group. There were no statistically significant differences among the groups.

4. Protocol 184–18 (P3–3) was a double-blind, randomized, placebo-controlled parallel group design study of 200 prison volunteers. The study compared the incidence of side effects of diphenhydramine in doses of 25 mg, 50 mg, or 100 mg with placebo. The drug or control was administered once daily for 28 days. Side effects were elicited at the time of dosing by the use of a general question about how the subject felt. Drowsiness was reported at least once by 34 percent of the placebo group, by 34 percent of the diphenhydramine 25 mg group, by 38 percent of the diphenhydramine 50 mg group, and by 60 percent of the diphenhydramine 100 mg group.

5. Protocol 266–17 (P9–3) was a double-blind, randomized, multicenter trial conducted by Dr. Tebrock involving 566 patients with cough due to cold, who received either Benylin or the Benylin vehicle without diphenydramine. See Appendix A, paragraph 1 above. Patients were asked questions about, among other things, the occurrence of any "unpleasant effects" while taking the medication. Witnesses for each party stated that this question may not have elicited as many reports of drowsiness as would a more

probing question; the Bureau's witnesses contended, moreover, that the subjects may not have perceived drowsiness as "unpleasant" (P19–6, p. 9; P19–8, p. 8; G–7, pp. 4–5; G–9, p. 7; G–9, pp. 7–8; g–11, pp. 23–24; G–13, p. 18; G–15, pp. 16–17; G–17, pp. 18–19; G–19, pp. 8–9; G–21, p. 19; G–27, p. 19). Witnesses for WL/PD argued, however, that subjects at work would, indeed, have considered drowsiness "unpleasant" and would have reported it (P19–6, p. 9; P19–9, p. 6). Drowsiness was reported at least once by 4.5 percent of the patients receiving Benylin and by 1.4 percent of the patients receiving the vehicle. There was a statistically significant difference between these percentages.

6. Protocols 184–35 (P9–9), 184–36 (P9–13), and 184–37 (P9–17) were studies conducted under identical protocols in a total of 43 patients with stabilized chronic cough associated with chronic pulmonary disease. After two days of monitoring without medication, patients received either placebo, diphenhydramine 25 mg, or diphenhydramine 50 mg by random assignment. In all of these studies, information on side effects was not elicited and was reported only when volunteered by subjects. Drowsiness was reported by 32.6 percent of the patients receiving diphenhydramine 50 mg, by 20.9 percent of the patients receiving diphenhydramine 25 mg, and by 7 percent of the patients receiving placebo. The differences are statistically significant (G–25, p. 5). WL/PD contends that the drowsiness shown in these studies is due to the antitussive effects of diphenhydramine rather than to its sedative properties: by alleviating their cough, the drug allowed the chronically ill, sleep-deprived patients to sleep. (P19–10, p. 3, P19–13, pp. 3–4; P19–15, p. 3; P19–16, p. 7). The Bureau disagrees (T356–58, 512, 560, 574–75).

7. Protocols 184–40 and 184–41 (P16–1, P17–1) were two double-blind, crossover, randomized studies conducted under identical protocols, in which a total of 154 volunteers received diphenhydramine 25 mg, dextrodiphenhydramine 25 mg, dextromethorphan 20 mg, codeine 20 mg, or placebo. Protocol 184–40 was conducted at Tulane University using 54 volunteers (physicians, medical students, house officers, and adult members of their families). Protocol 184–41 was conducted in the Parke-Davis facility using 100 volunteers (scientists, technicians, and secretarial staff). In the briefing of subjects to obtain informed consent, all were told that drowsiness may be a side effect of the medications administered in the studies. In Protocol 184–40, the incidence of drowsiness was 38 percent in the diphenhydramine group, 10 percent in the placebo group, 6 percent in the dextromethorphan group, and 4 percent in the codeine group (P16–3, p. 4). In Protocol 184–41, the incidence of drowsiness was 34.3 percent in the diphenhydramine group, 10.1 percent in placebo group, 7.1 percent in the dextromethorphan group, and the 5.1 percent in the codeine group (P17–2, p. 4). The results of these studies were statistically significant (G–31, p. 16).

Order

It is hereby ordered that approval of NDA 6–514 S–007, submitted by Warner-Lambert/Parke-Davis & Co. for the drug Benylin, is refused.

Dated: June 29, 1979.

Donald Kennedy,

Commissioner of Food and Drugs.

[FR Doc 79-27010 Filed 8-30-79; 8:45 am]
BILLING CODE 4110-03-M