Buehler

OCT 26 1994

Steven M. Rodin, M.D.
Medical Officer

Food and Drug Administration

Division of Cardio-Renal Drugs
Tel 301-443-0320; FAX-9283

Medical Review of NDA Efficacy and Patient Clinical Pharmacology data

1 General information

NDA #:	**20-420**
Drug:	**i.v. arbutamine**
Sponsor:	Gensia Inc.
Proposed indication:	diagnostic adjunct
related IND:	34-895 (arbutamine)
Pharmacologic type:	nonselective β-adrenergic agonist, and low potency α-adrenergic agonist
Date of NDA submission:	23 December 1993
Latest data submission:	7 October 1994
Reviewer:	Steven M. Rodin, M.D.
Review last revised:	26 October 1994

2 Table of contents

Table of contents [continued]

3 Background

3.1 *General background*:

This NDA is for a hemodynamic stressor agent proposed for use as an adjunct to various cardiovascular diagnostic techniques (electrocardiography (EKG), 2-dimensional (2D) echocardiography, ventricular perfusion scintigraphy) for the noninvasive diagnosis of coronary artery disease (CAD). Arbutamine is a nonselective β-adrenergic agonist with low potency α-adrenergic agonism. It is proposed to be marketed as part of a drug-device system called the GenESA® System. In this system a "closed-loop" (i.e. feedback controlled, as opposed to "open-loop" or nonfeedback controlled) device monitors a patient's heart rate (HR) response to iv arbutamine, and attempts to adjust the rate of infusion accordingly (every 15 seconds) to achieve a pre-specified rate of HR rise and maximum ("target") HR.

This review focuses on the clinical pharmacology data obtained in CAD patients, and the clinical efficacy data. Elsewhere in the Cardiorenal division the safety data are being reviewed (by Dr. Karen Frank), and the clinical pharmacology data from healthy subjects are being reviewed (by Dr. Igor Cerny). Reviewers at the Center for Devices and Radiologic Health are evaluating the performance of the infusion device.

3.2 *Operational definitions*:

The following is a glossary of the terminology which is found in this review.

- **dyskinesis**: paradoxical systolic wall motion away from the center of the cardiac chamber.

- **dysynergy**: any of several types of cardiac wall motion abnormality including dyskinesis, hypokinesis (reduced wall motion), and akinesis (absent wall motion).

- **false negative (FN)**: a negative test result in a patient with confirmed disease.

- **false negative fraction**: this is the fraction of diagnoses which are incorrect among patients with disease. The false negative fraction (FNF) is calculated as the number of false negative (FN) test results divided by the number of subjects with disease. The FNF is equal to 1 minus sensitivity.

- **false positive (FP)**: a positive test result in a patient confirmed to be free of the disease of interest.

- **false positive fraction**: among patients without disease, the fraction which is incorrectly diagnosed is described by the false positive fraction (FPF). The FPF is calculated as the number of false positive (FP) test results divided by the number of subjects without disease. The FPF is equal to 1 - specificity.

Background

3.2 *Operational definitions* [continued]:

- **gold standard**: the defining criteria for the presence or absence of disease. In this NDA the presence of significant CAD is defined as coronary angiographic evidence of ≥50% reduction in cross-sectional diameter of any "major" epicardial coronary artery ("major" refers to the left main, left anterior descending, left circumflex, or right coronary arteries or any of their major branches).

- **kappa statistic**: for a given diagnosis kappa describes the overall agreement, beyond that which can be attributed to chance, between 2 diagnostic methods. A kappa value ≥ 0.75 is conventionally taken to represent a high degree of agreement between two techniques.

- **predictive value of a negative test**: among patients with negative test results, the fraction which is correctly diagnosed is described by the predictive value of a negative test (also called the negative predictive value). This is calculated as the ratio of the number of true negative (TN) test results to all negative test results, i.e. TN/(TN + FN).

- **predictive value of a positive test**: this describes the fraction of patients correctly diagnosed among those with positive test results. Also called the positive predictive value, this is calculated as the ratio of true positive (TP) test results to all positive test results, i.e. TP/(TP + FP).

- **sensitivity**: among patients with disease, the fraction which is correctly diagnosed is described by the diagnostic sensitivity of the test. It is calculated as the number of TP test results divided by the number of patients with disease. When there are no indeterminate test results the number of patients with disease equals the sum of the number of TP and FN test results, and sensitivity is calculable as TP/(TP + FN).

- **specificity**: among patients without disease, the fraction which are correctly diagnosed is described by the diagnostic specificity of the test. It is calculated as the number of true negative (TN) test results divided by the number of patients without disease. When there are no indeterminate test results the number of patients without disease equals the sum of the number of TN and FP test results, and specificity is calculated as TN/(TN + FP)].

- **stress**: a hemodynamic perturbation provoked either by exercise or a cardioactive agent.

- **target HR**: 85% of a patient's age-predicted maximum HR, calculated by the formula: 0.85 x (220 minus the patient's age in years).

- **test**: a screening procedure of putative diagnostic value.

- **true negative (TN)**: a negative test result in a patient confirmed to be free of the disease of interest.

- **true positive (TP)**: a positive test result in a patient with confirmed disease.

Background

3.3 *Logic of the metrics used in diagnostic assessment*:

This discussion is intended to illuminate the somewhat arcane metrics used in the assessment of diagnostic methodologies.

Seven metrics are considered in this review:

- sensitivity
- specificity
- false negative fraction
- false positive fraction
- positive predictive value
- negative predictive value
- kappa statistic

The kappa statistic stands apart from the others because it provides no information about the diagnostic validity of a test. It simply describes the overall agreement between two diagnostic methods. A high degree of agreement can be obtained with two diagnostic techniques that are each highly prone to error. I refer the reader to the Statistical Reviewer's report for computations of the kappa statistic.

The 6 remaining metrics (these being sensitivity, specificity, false negative fraction, false positive fraction, postive predictive value, and negative predictive value) are derived with data which is validated against a diagnostic gold standard. Of these there are 4 metrics which provide characterizations of the inherent error of the test, i.e. characterizations which are independent of the prevalence of disease in the population tested. These 4 prevalence-independent metrics are sensitivity, false negative fraction, specificity, and false positive fraction. Since these 4 prevalence-independent metrics are actually composed of two pairs the components of which are each related by simple algebraic expressions[1].

As shown in the following table these 4 prevalence-independent metrics can be grouped according to their relationship to two factors (correctness of diagnosis, and presence of disease) each with dichotomous levels (correct vs incorrect, disease vs no disease).

APPEARS THIS WAY
ON ORIGINAL

[1]false positive fraction = 1 - specificity; and false negative fraction = 1 - sensitivity.

Background

Table: 1

Conceptual organization of the 4 prevalence-independent metrics

	Descriptors of test performance in Patients WITH disease	**Descriptors of test performance in Patients WITHOUT disease**
Descriptors of the Fraction CORRECTLY Diagnosed	*Sensitivity*	*Specificity*
Descriptors of the Fraction INCORRECTLY Diagnosed	*False negative fraction*	*False positive fraction*

Note that the descriptors of the correctness of the diagnosis among patients *with disease* are: sensitivity (which refers to the fraction *correctly* diagnosed), and false negative fraction (referring to the fraction *incorrectly* diagnosed).

The descriptors of the correctness of the diagnosis among patients *without disease* are: specificity (which refers to the fraction *correctly* diagnosed) and false positive fraction (referring to the fraction *incorrectly* diagnosed).

Of the prevalence-independent descriptors there are two which provide direct means for conveying the adequacy of the performance of a diagnostic test. These two terms are sensitivity and specificity, particularly in their adjectival forms. An insensitive test performs just as the term suggests, i.e. it tends to fail to detect disease when it is present. These "missed" diagnoses obviously occur in patients with disease (i.e. they are false negatives), and it follows that an insensitive test is associated with a large false negative fraction. A nonspecific test also performs just as the term suggests, i.e. it detects signals that are not specific indicators of the presence of the disease in question. In the process of nonspecific detection the test "overdiagnoses" disease. These "overdiagnoses" obviously occur in patients without disease (i.e. they are false positives), and it follows that a nonspecific test is associated with a large false positive fraction.

Background

Those metrics which are dependent on the prevalence of disease in the population tested are positive predictive value and negative predictive value. Because of this property they are not useful descriptors of the inherent discriminative value of a test. However, they can be calculated merely with knowledge of the outcomes of the test, in the absence of knowledge of the true presence or absence of disease in individuals. Hence they do have a certain relevance to the usual clinical scenario in which the disease status of the tested subject is an unknown.

The negative predictive value (the fraction of patients correctly diagnosed among those with negative test results) is inversely proportional to the disease prevalence in the tested population whereas the positive predictive value (the fraction who are correctly diagnosed[2] among patients with positive test results) is directly proportional to the disease prevalence. These relationships between predictive values and prevalence can be illustrated as follows [with thanks to Professor Alvan Feinstein[3]]. Consider a sample (with 9% disease prevalence) comprised of 10 people with confirmed disease and 100 people confirmed to have no disease. Diagnoses are rendered by a test with both sensitivity and specificity equal to 50% (i.e. with both false negative fraction and false positive fraction equal to 50%). Testing would yield 5 TP results, 5 FN results, 50 FP results, and 50 TN results. Note that the negative predictive value (calculated as TN/(TN + FN)= 50/(50 + 5) = 91%) is quite high in the presence of a low disease prevalence, despite mediocre test specificity and sensitivity (in contrast the positive predictive value (calculated as TP/(TP + FP) = 5/(5+50) = 9%) is low).

Alternatively, consider a sample (with 91% disease prevalence) comprised of 100 people with confirmed disease and 10 people confirmed to have no disease. Again diagnoses are provided by a test with both sensitivity and specificity equal to 50%. Testing in this case yields 50 true positive results, 50 false negatives, 5 false positives, and 5 true negatives. Note that the positive predictive value (calculated as 50/(50+5)= 91%) is quite high in the presence of a high disease prevalence, despite mediocre test specificity and sensitivity (in contrast the negative predictive value (calculated as 5/(5+50)= 91%) is low.

[2]in describing predictive values there are no commonly used metrics which connote the fraction *incorrectly* diagnosed.

[3]Clinical Biostatistics. 1977; C.V. Mosby, St. Louis.

4 Chemistry

Arbutamine is formally named R (1) -1-(3,4-dihydroxyphenyl)-2-[4-hydroxyphenyl) butylamino] ethanol hydrochloride. See the chemist's review for a detailed discussion of chemistry.

5 Preclinical Pharmacology

See the pharmacologist's review for a detailed discussion of pre-clinical pharmacology. Briefly, this drug is a nonselective β-adrenergic agonist (and low potency α-adrenergic agonist). Arbutamine produced positive chronotropic and positive inotropic effects in animals. The disappearance halftime for arbutamine-induced tachycardia was about 7-17 minutes.

6 Clinical Pharmacokinetics

See the Biopharmaceutical review for detailed discussions of clinical pharmacokinetics. Briefly, the sponsor claims that the mean (± standard deviation) half-life of i.v. arbutamine was 8.2 ± 1.6 minutes in 13 healthy volunteers when drug was infused at dose rates up to approximately 0.3 μg/kg/min. However, it is not clear whether the kinetics are linear up to the maximum infusion rate used in clinical studies (0.8 μg/kg/min).

The kinetic data from normal subjects do not provide evidence that HR continues to rise after discontinuation of arbutamine infusion. Nonetheless, the sponsor adopted the view that hysteresis would be present to produce HR overshoot after discontinuation of arbutamine, and for the patient efficacy studies they included a parameter in the device's algorithm which determined that infusions would be stopped prior to the attainment of the desired "target HR".

[CONTINUED NEXT PAGE]

7 Clinical Pharmacology in presumably healthy subjects

The following study (128) was submitted with the clinical efficacy data because of its claimed relevance for the estimation of diagnostic specificity. I find that it does not have this relevance and classify it instead as a clinical pharmacology study. Note that the remainder of the clinical pharmacology studies in healthy subjects are to be covered in Dr. Cerny's review.

7.1 A characterizion of the proportion of negative arbutamine stress tests among subjects with a low likelihood of CAD (study 128)

SUMMARY:

This nonblinded study exposed 63 subjects (individuals with absent EKG ST-segment changes during exercise, and overall low pretest probability of CAD) to iv arbutamine on 3 separate occasions without sequence randomization. A "closed-loop" infusion device was used in an attempt to enable HR to exert feedback control of infusion rate. A different diagnostic modality was utilized during each infusion (surface EKG, transthoracic cardiac ultrasonography, and thallium-201 scintigraphy, respectively). The objective was to characterize the ratio of negative test outcomes to all test outcomes ("normalcy rate").

PROTOCOL:

- **Enrollment criteria:**

Enrolled subjects were adults of both sexes with absent EKG ST-segment changes during exercise, no previous diagnostic evidence of CAD, and an estimated low prior probability of CAD[4]. Excluded from enrollment were pregnant women and women of childbearing potential, subjects with hypokalemia, as well as those with a history of cardiovascular or cerebrovascular disease, renal or hepatic insufficiency, or a requirement for the uninterrupted use of a β-adrenoreceptor antagonist.

- **Test regimen:**

Intravenous arbutamine solution (50 µg/ml) was infused to supine subjects. A "closed-loop" infusion device was used in an attempt to enable HR to exert feedback control of infusion rate. The algorithm intended to maintain a HR rise of 8 beats per min (bpm) per minute. The range of intended dose rates was 0.056-0.8 µg/kg/min, and the intended maximum total dose was 10 µg/kg. The range of intended infusion flow rates was, for example, 56-800 µl/min in a 50 kg subject, and 112-1600 µl/min in a 100 kg subject.

A 12-lead surface EKG system was used to monitor patients during each test. Transthoracic 2D echocardiography was performed using standard image planes (apical two chamber, apical four

---

[4]this estimation was based on clinical grounds using a nomogram which incorporated the variables: age, sex, hyperlipidemia, diabetes mellitus, tobacco use, and EKG findings.

chamber, parasternal short and long axis views) prior to stress, and again at peak stress.

Approximately one minute prior to reaching "target HR" an iv injection of 2.5 to 3.5 mCi thallium-201 was undertaken. "Immediate" post-stress cardiac perfusion images were acquired within 10 to 15 minutes following the thallium-201 injection, using a rotating gamma camera with single photon emission computed tomography (SPECT) capability.

▸ **Endpoints**:

The objective was to characterize, for each diagnostic modality, the ratio of negative test results to all test results in a selected sample of patients with a low likelihood of CAD.

A positive EKG stress test for CAD was defined as one which displayed horizontal or downsloping ST segment depression ≥1 mm, or ≥1 mm ST segment elevation (changes were assessed relative to pre-stress, measurements were obtained 60 msec after the J point, and ST elevation was only assessed in leads without Q waves). The EKG interpretors were blinded to the source of the data, and to the results of other study procedures.

A positive echocardiographic stress test for CAD was defined as one in which a new or worsened (relative to pre-stress) ventricular wall motion abnormality was observed in at least one segment of the anterior wall, or in 2 adjacent segments of the posteroinferior wall. Echocardiograms were analyzed centrally by two independent observers who were blinded to the source of the images, and to the results of other study procedures. A third observer was used in cases of discrepancy between the first two observers.

The criterion for a positive thallium-201 stress test for CAD was the occurrence of at least two ventricular regions with moderately low tracer uptake, or the occurrence of a single region with severely reduced tracer uptake. The perfusion images were analyzed centrally by two independent observers who were blinded to the source of the images, and to the results of any other study procedures. A third observer was used in cases of discrepancy between the first two observers.

▸ **Statistical procedures**:

• **Dataset analyzed**:

One enrolled patient withdrew before receiving arbutamine.

• **Handling of missing data**:

Echocardiogram results were considered nondiagnostic ("indeterminate") when a non-positive test had 4 or more of the 16 defined ventricular segments not reliably visualized. It is not clear how missing or nondiagnostic data were handled.

Study 128

• **Analyses performed**:

Descriptive statistics were computed and a 95% confidence interval was constructed using the normal approximation to the binomial distribution.

RESULTS:

▸ **Demographics**:

The mean age of study subjects was 43 years.

▸ **Period effects**

For the sensitivity and specificity analyses it is not clear whether the hypothesis of equal carryover effectswas rejected.

▸ **Pharmacodynamic outcomes**:

Table: 2

Mean BP during arbutamine exposure
[study 128]

Systolic/Diastolic BP (mm Hg)	
pre-test	**maximum change from pre-test**
123/73	41/-22

[source: modification of tables 15 & 16, vol 89]

• **Rate of decline of tachycardia**:

By ten minutes following the discontinuation of arbutamine infusion, 36% of the subjects manifested a 50% or greater decline from the maximally attained HR.

Study 128

• **"Normalcy rates":**

Mean "normalcy rates" (the ratio of negative test outcomes to all test outcomes) are shown in the following table.

Table: 3

Mean "normalcy rate" during arbutamine stress tests [study 128]

Arbutamine-EKG	*Arbutamine-echocardiography*	*Arbutamine-thallium*
79% 95% CI= 67-89% [n=63]	96% 95% CI= 87-100% [n=52]	90% 95% CI= 79-96% [n=58]

[source: modification of table 18, vol 89; & addendum dated 6/8/94]

Shown are the results of an intent-to-treat analysis.

Study 128

COMMENTS (study 128):

1. Noninvasive diagnostic test methodologies interact with certain covariates related to the cardiac pathophysiology of patients. These interactions contribute to the occurrence of false positive test results (which reduce test specificity)[5,6]. Yet these would only be expected to reduce test specificity in patients, and not in subjects in whom these pathophysiologic variables are absent. For this reason the "normalcy rate" observed in subjects who are presumably devoid of cardiac pathophysiology plausibly overestimates the specificity (and underestimates the false positive fraction) of arbutamine-stress testing that would be found in patients with suspected CAD.

[5]one such covariate is the presence of resting ventricular hypokinesis. Imaging methods can falsely suggest the presence of critical coronary stenoses in chronically infarcted, hypokinetic regions which are subtended by recanalized arteries [see Marcovitz PA et. al. Am J Cardiol 1992; 69:1269-1273, and Eisner RL et. al. J Nucl Med 1994; 35:638-643].

[6]another such variable is the presence of subcritical coronary stenoses which do not meet standard clinical criteria for significant CAD. Such stenoses have the potential to confer stress-associated ventricular blood flow inhomogeneities which do meet conventional criteria for hypoperfusion [Gould KL, Am J Cardiol 1978; 42: 761-768].

8 Clinical Pharmacology in CAD patients

Among clinical pharmacology studies in this NDA, this review evaluates only the studies of CAD patients.

8.1 Arbutamine dose-finding investigation utilizing peripheral hemodynamic, EKG and ventricular wall motion endpoints in CAD patients (study 107)

SUMMARY:

In this nonblinded, uncontrolled study 4 subjects (patients with angiographic CAD, exercise-induced EKG ST-segment depression or elevation, and atrial pacing-induced ventricular dysynergy) received "open-loop" (not feedback controlled) iv arbutamine infusions (using ascending titration with 3 minute infusions of 0.025, 0.05, 0.1, 0.2, 0.4, and 0.6 µg/kg/min respectively, during each step). Surface EKGs and transesophageal echocardiograms were recorded during drug infusion. The objective was to assess the effects of arbutamine on peripheral hemodynamics, EKG-ST segments, and ventricular wall motion.

PROTOCOL:

- **Enrollment criteria**:

Enrolled subjects were patients with angiographic CAD, exercise-induced EKG ST-segment depression or elevation, and atrial pacing-induced ventricular dysynergy. It is not clear what atrial pacing protocol was used during the selection of patients.

- **Test regimen**:

Within 2-14 days after the qualifying atrial pacing study subjects received iv arbutamine while in the supine position. Arbutamine was given in an ascending titration regimen which involved 3 minute having intended dose rates of 0.025, 0.05, 0.1, and 0.2 µg/kg/min[7], respectively. An "open-loop" (not feedback controlled) device administered an arbutamine solution of 5 µg/ml. The range of intended infusion flow rates was, for example, 0.25-2 ml/min in a 50 kg subject, and 0.5-4 ml/min in a 100 kg subject. Twelve lead surface EKGs, and (after pretreatment with 0.5 mg iv atropine) transesophageal echocardiograms were recorded.

- **Endpoints**:

The objective was to assess the capacity of arbutamine to induce changes in EKG-ST segments and ventricular wall motion.

[7] one patient was also further titrated up to 0.4 and 0.6 µg/kg/min.

Study 107

A positive EKG stress test for CAD was defined as one which displayed horizontal or downsloping ST segment depression ≥1 mm, or ≥2 mm ST segment elevation (changes were assessed relative to pre-stress, measurements were obtained 80 msec after the J point, and ST elevation was only assessed in leads without pathological Q waves).

Ventricular wall motion abnormalities were assessed by means of 2D rest-stress echocardiography. It is not clear whether the necessary differentiation was made between pre-stress (resting) vs stress-induced wall motion abnormality.

▸ **Statistical procedures**:

• **Handling of missing data**:

There were no prespecified criteria for nondiagnostic echocardiograms, and it is not clear how such results (if any[8]) were handled.

RESULTS:

▸ **Pharmacodynamic outcomes**:

The mean maximal arbutamine-induced change in HR was 44 bpm, and the tachycardia tended to be dose-related.

In the small sample studied the maximal mean increase in systolic BP after a 3 minute arbutamine infusion at 0.2 µg/kg/min was 0.3 mm Hg. The maximal mean decrease in systolic BP after a 3 minute infusion at 0.1 µg/kg/min was -24.5[9]. The maximal mean decrease in diastolic BP after a 3 minute arbutamine infusion of 0.2 µg/kg/min was -30.3 mm Hg, on average, with comparable changes being produced by the 0.025 to 0.1 µg/kg/min doses. Diastolic BP was markedly reduced by the 0.4 and 0.6 µg/kg/min infusions (by 86 to 91 mm Hg) in the one such exposed subject.

The mean time for maximal HR to decrease by 50% after the discontinuation of arbutamine was reported to be approximately 6.5 minutes (for the 0.2 µg/kg/min dose), although it is not clear what this "offset time" was for the other doses studied.

---

[8]the transesophageal echocardiographic interrogation method is less prone to generating uninterpretable images than is the transthoracic method.

[9]this change was comparable to that produced by the lower doses.

Study 107

▸ **Efficacy outcomes:**

ST-segment depression was achieved in all four patients both during atrial pacing and during arbutamine infusions. The ventricular wall motion findings were not adequately reported.

COMMENTS (study 107):

1. The results of this small, nonplacebo-controlled study suggest (as did study 112) that arbutamine can produce substantial BP lowering in CAD patients. The reader is referred to Dr. Frank's safety review for details regarding the clinical outcomes associated with this effect.

8.2 Arbutamine dose-finding investigation utilizing peripheral hemodynamic, EKG, and ventricular wall motion endpoints in CAD patients (study 108)

SUMMARY:

In this uncontrolled, nonblinded, ascending dose study 10 subjects (patients with angiographic CAD and exercise-induced EKG ST-segment depression or elevation) received "open-loop" (not feedback controlled) iv arbutamine (using ascending doses involving 3 minute infusions of 0.025, 0.05, 0.1, 0.2, 0.4, 0.6, 0.8 and 1.1 µg/kg/min during each respective step). Surface EKGs and transesophageal echocardiograms were recorded during drug infusion. The objective was to assess the effects of arbutamine on peripheral hemodynamics, EKG-ST segments, and ventricular wall motion.

PROTOCOL:

- **Enrollment criteria:**

Enrolled subjects were candidates for percutaneous transluminal coronary angioplasty (PTCA) who had angiographic CAD and exercise-induced EKG ST-segment elevation or depression. Excluded from enrollment were pregnant or lactating women, and women of childbearing potential as well as those subjects manifesting:

- heart failure, chronic arrhythmias, myocardial infarction or unstable angina within the previous 3 months, cardiac conduction abnormalities, hypertension, idiopathic hypertrophic subaortic stenosis, or aortic stenosis
- a requirement for uninterrupted use of β-adrenoreceptor agonists or antagonists, or any antihypertensive drug
- cerebrovascular disease, thyroid disease, or glaucoma

- **Test regimen:**

An "open-loop" device delivered arbutamine infusions without exerting feedback control of infusion rate. Arbutamine solution (5 µg/ml) was administered to supine subjects via an ascending iv titration with 3 minute infusions during each step. The intended dose rates were 0.025[10], 0.05, 0.1, 0.2, 0.4, 0.6, 0.8 and 1.1 µg/kg/min, respectively, during each step. The intended infusion flow rates ranged, for example, from 0.25-11 ml/min in 50 kg subjects to 0.5-22 ml/min in 100 kg subjects.

Surface EKG and transthoracic 2D echocardiographic measurements were obtained during arbutamine infusions. Subsequently to the arbutamine test subjects were to have a PTCA (and then a second arbutamine stress test). A 12-lead surface EKG system was used to monitor patients during stress tests.

[10]two subjects were started at a higher arbutamine dose, i.e. 0.1 µg/kg/min.

Study 108

▸ **Endpoints:**

The principle endpoint was a comparison of the diagnostic sensitivity of arbutamine-induced vs exercise-induced changes in EKG-ST segments.
A positive EKG stress test for CAD was defined as one which displayed T wave inversion, horizontal or downsloping ST segment depression ≥1 mm, or ≥2 mm ST segment elevation (changes were assessed relative to pre-stress, and ST measurements were obtained 80 msec after the J point).

Another endpoint was arbutamine-induced ventricular dysynergy, as detected by 2D rest-stress echocardiography. The criteria for a positive echocardiogram were not detailed in the sponsor's report. Wall motion outcomes were apparently analyzed on a global (as opposed to regional) basis.

▸ **Statistical procedures:**

• **Dataset analyzed:**

The analysis reviewed here included all ten subjects who received arbutamine before PTCA procedures.

• **Handling of missing data:**

There were no prespecified criteria for characterizing echocardiogram results as nondiagnostic, and it is not clear how such results (if any) were handled.

• **Analyses performed:**

It appears that no statistical tests were performed.

RESULTS:

▸ **Demographics:**

The mean age of subjects was 53 years.

APPEARS THIS WAY
ON ORIGINAL

Steven M. Rodin, M.D.
Medical Officer

Food and Drug Administration
Division of Cardio-Renal Drugs
Tel 301-443-0320; FAX-9283

Addendum #1 to Medical Review of NDA:

1 General information

NDA #:	**20-420**
Drug:	**i.v. arbutamine**
Sponsor:	Gensia Inc.
Proposed indication:	diagnostic adjunct
Reviewer:	Steven M. Rodin, M.D.
Addendum last revised:	?

2 Errata:

a. The first paragraph on page 20 of my original review (dated 26 October 1994) contained several typographical errors. The correct text is as follows:

Study 108

▸ **Pharmacodynamic outcomes**:

The mean maximal arbutamine-associated increase from pretest systolic BP was 36.6 mm Hg whereas the mean maximal decrease was -8.8 mm Hg. With respect to diastolic BP the mean maximal increase from pretest values was 11.7 mm Hg and the mean maximal decrease was -11.6 mm Hg. The mean maximal increase from pretest HR was approximately 60 bpm, and the mean maximal decrease from pretest HR was -2.8 bpm [source: tables 5-6, pages 57-62, volume 44].

APPEARS THIS WAY
ON ORIGINAL

Study 108

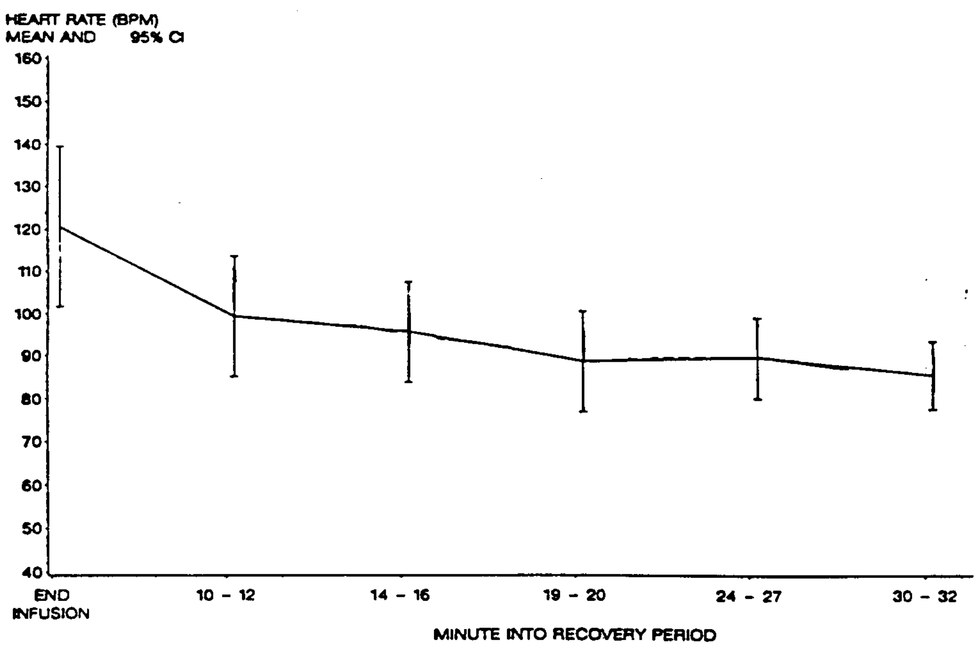
▸ Pharmacodynamic outcomes:

The mean maximal change from pretest diastolic BP was 11.7 mm Hg whereas the mean maximal change from pretest diastolic BP was -11.6 mm Hg. The mean maximal change from pretest HR was 62.7 bpm whereas the mean maximal change from pretest HR was -2.8 bpm.

Mean systolic and diastolic BPs generally returned to pre-treatment levels by approximately 15 minutes post-discontinuation of drug. In contrast, there was an average persistence of HR elevation (relative to pre-treatment) for the entire 30 minutes during which observations were obtained after drug discontinuation, and absolute sinus tachycardia (HR ≥100 bpm) remained observable at approximately 15 minutes into recovery [source: appendix 13, addendum dated 8/12/94]. See the following figure.

Figure: 1

Time course of tachycardia decline
following discontinuation of arbutamine (study 108)



[source: modification of figure in appendix 13, addendum dated 8/12/94]

Pretreatment mean HR was approximately 70 bpm, and the maximum arbutamine-associated HR was approximately 130 bpm. Data were obtained prior to PTCA.

Study 108

▸ **Efficacy outcomes**:

Among the 10 CAD patients who underwent arbutamine tests before PTCA procedures, 2 were reported to have positive arbutamine-EKG tests (yielding a mean sensitivity of 20%), and 5 were reported to have positive arbutamine-echocardiography tests (presumably based on a global ventricular wall motion analysis).

COMMENTS (study 108):

1. On average, there was a persistent HR elevation (relative to pre-treatment) for the entire 30 minutes of observation following discontinuation of arbutamine infusions, and absolute sinus tachycardia was still present approximately 15 minutes into recovery. These findings, although not placebo-controlled, are consistent with the reported finding that the drug has at least an 8 minute elimination half-life. The prolonged half-life of tachycardia could plausibly contribute to a persistence of myocardial hypoxia-related adverse events after discontinuation of the drug infusion in CAD patients. Refer to the safety review for insights into patient outcomes during recovering from arbutamine stress tests.

2. The submitted data are inadequate for critical evaluation of the diagnostic aspects of this study, but at their face the reported sensitivities for the arbutamine-assisted diagnosis of CAD were poor. The use of global (as opposed to regional) ventricular wall motion criteria for a positive echocardiogram could have plausibly contributed to a suboptimal sensitivity.

8.3 Arbutamine effect on hemodynamics in subjects with known systolic hypotensive responses to iv dobutamine (study 115)

SUMMARY:

This was a small, uncontrolled, crossover study which primarily examined the safety of iv arbutamine in 3 subjects known to exhibit a systolic hypotensive response to iv dobutamine.

Briefly, subjects crossed over between 2 sequence-randomized iv administrations of arbutamine solution (1 mg/ml) which were separated by at least 4 hours. One administration was via a force-titrated "open-loop" (not feedback controlled) infusion, and the other was via a "closed-loop" infusion. The "open-loop" regimen involved 3 minute infusions at each of several rates (0.025, 0.05, 0.1, 0.2, 0.4, 0.6 and 0.8 μg/kg/min) whereas the "closed-loop" regimen intended to achieve a HR increase of 4 bpm/min using a maximum total arbutamine dose of 6.5 μg/kg and a maximum dose rate of 0.4 μg/kg/min.

RESULTS:

Briefly, the "open-loop" regimen was associated with symptomatic, marked systolic hypotension (defined as a ≥40 mm Hg decrease from pre-treatment systolic BP to an absolute level <100 mm Hg) in 2 of 3 patients. Similarly, the "closed-loop" regimen resulted in symptomatic, marked systolic hypotension in 2 of 3 patients. Following the hypotensive episodes, HR and BP were reported to have returned to pre-event values within 3 minutes of the event in each patient.

COMMENTS:

See Dr. Frank's safety review for a detailed discussion of the results of this study.

8.4 Preliminary data on ventricular perfusion, and central hemodynamics following arbutamine exposure (study 130)

SUMMARY:

The sponsor submitted a prelimary report of this nonblinded, uncontrolled study in which iv arbutamine had been administered to 11 subjects (patients with suspected CAD). A "closed-loop" infusion device was used in an attempt to enable HR to exert feedback control of infusion rate. The objectives were to assess the effects of arbutamine on central hemodynamics, and to estimate the sensitivity and specificity (for the diagnosis of CAD) of arbutamine-induced EKG-ST segment changes, and arbutamine-associated ventricular perfusion abnormalities (using Tc^{99m} sestamibi as the myocardial blood flow tracer).

PROTOCOL:

- **Enrollment criteria:**

The enrollment criteria were generally comparable to those described below for study 120.

- **Test regimen:**

Coronary angiography was performed, and after the return of hemodynamic parameters to within 20% of pre-treatment values, iv arbutamine was administered. A 50 μg/ml arbutamine solution was administered by means of a "closed-loop" infusion device. The range of intended dose rates was 0.008-0.8 μg/kg/min, and the intended maximum total dose was 10 μg/kg. The range of intended infusion flow rates was, for example, 8-800 μl/min in a 50 kg subject, and 16-1600 μl/min in a 100 kg subject.
During arbutamine infusions (which were performed in the supine position) hemodynamic and surface EKG (3 lead) data were obtained by standard methods. In some cases ventricular perfusion images were acquired after infusing 15-25 mCi dose of Tc^{99m} sestamibi into the right ventricle during the maximum HR elevation. SPECT ventricular perfusion images were collected 30-60 minutes following tracer injection. Resting re-injection images were collected at 24-48 hours post stress. Short axis, long axis and horizontal long axis tomograms were divided into a total of 20 ventricular segments for visual scoring.

Study 130

▸ **Endpoints**:

The objectives were to assess the effects of arbutamine on central hemodynamics, and to estimate the diagnostic sensitivity and specificity (for the diagnosis of CAD) of arbutamine-associated[11] ventricular perfusion abnormalities, and arbutamine-induced EKG-ST segment depression or elevation.

A positive EKG stress test for CAD was defined as one which displayed horizontal or downsloping ST segment depression ≥1 mm, or ≥1 mm ST segment elevation (changes were assessed relative to pre-stress, measurements were obtained 60 msec after the J point, and ST elevation was only assessed in leads without Q waves).

The criterion for a positive perfusion scintigram for CAD was retrospectively defined. The criterion excluded cases where the regional deficit in scintillations (as observed on immediate post-stress images) became less marked after re-injection of tracer.

▸ **Statistical procedures**:

• **Dataset analyzed**:

Seven of the analyzed patients had angiographic evidence of significant CAD and 4 had negative coronary angiograms. Only 8 of the 11 subjects enrolled at the time of submission of this preliminary report were studied with ventricular perfusion imaging.

• **Handling of missing data**:

No perfusion images were classified as nondiagnostic (by application of retrospective criteria).

• **Analyses performed**:

The sponsor asserts that there are no plans to change the conduct of this study based upon the results of this interim look. Only descriptive statistics were thusfar submitted.

RESULTS:

▸ **Demographics**: The mean age of the subjects was 56 years.

---

[11] I use the word "associated" because the immediate post-arbutamine scintigrams do not distinguish arbutamine-induced hypoperfusion from hypoperfusion which was manifest prior to arbutamine exposure.

Study 130

▸ **Pharmacodynamic outcomes:**

The following are the conventions used for the description of hemodynamic parameters.

Table: 4

Terminology: units and abbreviations for hemodynamic parameters

parameter	*abbreviation*	*units*
heart rate	HR	bpm
pulmonary capillary wedge pressure	PCWP	mm Hg
cardiac output	CO	L/min
left ventricular dP/dt	LV dP/dt	mm Hg/sec
left ventricular end diastolic pressure	LVEDP	mm Hg
systemic vascular resistance	SVR	dyn/s/cm^{-5}

Study 130

▸ **Pharmacodynamic outcomes** [continued]:

Hemodynamic changes associated with arbutamine exposure are shown in the following table.

Table: 5

Change from pre-treatment mean hemodynamics after arbutamine infusion (study 130)

Parameter	*Patients with negative coronary angiograms [n=4]*		*Patients with angiographic CAD [n=7]*	
	pre-treatment value	*change at peak stress*	*pre-treatment value*	*change at peak stress*
HR (bpm)	77	39	73	39
Systolic BP (mm Hg)	135	-28	144	-31
Diastolic BP (mm Hg)	73	-23	77	-15
PCWP (mm Hg)	12	-4	10	-0.4
CO (L/min)	7.0	3.7	6.7	3.4
LV dP/dt (mm Hg/sec)	1253	1214	1376	850
LVEDP (mm Hg)	23	-6	19	-8
SVR (dyn/s/cm^{-5})	1349	-685	1255	-647

[source: modification of tables 8, 10, 12, 13, 14, & 17; vol 65]

The blood pressures (BP) shown were obtained invasively via a femoral arterial line.

• **Rate of decline of tachycardia**:

After the discontinuation of arbutamine infusions in the 3 CAD patients with reported data the mean time for HR to decrease by 50% was approximately 15 minutes.

Study 130

▸ **Efficacy outcomes:**

The preliminary point estimates (based on n=11) for the sensitivity and specificity (for the diagnosis of CAD) of arbutamine-EKG tests were 14% and 100%, respectively.

COMMENTS (study 130):

1. In association with iv arbutamine administration to CAD patients (at intended dose rates of 0.008-0.8 μg/kg/min) there were acute hemodynamic changes observed which were consistent with the expected positive chronotropic and positive inotropic actions of a nonselective β-adrenoreceptor agonist. However, the vasodepressor effect of drug could have plausibly been overestimated (with resultant overestimation of its effect to lower SVR and BP, and to raise LV dP/dt) on the basis of an artefact attributable to the spontaneous resolution of catheter-related systemic vasoconstriction.[12,13] The protocol neither used a placebo control to enable estimation of this artefact, nor employed the 24 hour delay (between the placement of the right heart catheter and the measurement of drug-associated hemodynamics) which is apparently necessary to obviate it.

2. The use of a 3 lead EKG recording system plausibly resulted in an underestimation of test sensitivity for the diagnosis of CAD (and an overestimation of the false negative fraction), relative to the 12 lead data-aquiring systems used elsewhere in this NDA.

3. The definition of a positive perfusion image for the diagnosis of CAD was not valid in this study. It incorrectly defined CAD to be absent when the hypoperfusion observed on immediate post-stress images became apparently less marked on delayed post-stress images (obtained after re-injection of tracer). The generally accepted definition of a positive Tc^{99m} sestamibi test for CAD is based solely on the finding of hypoperfusion on immediate post-stress images (irrespective of delayed post-stress data). The contrived definition used in this study plausibly resulted in an underestimation of sensitivity (and overestimation of the false negative fraction) for the diagnosis of CAD.

4. Seven additional patients have finished this study since the sponsor's interim report, but analyses of the final data have not been submitted.

APPEARS THIS WAY
ON ORIGINAL

[12]Packer M, et. al. (1985a) Circulation 71: 761-166.

[13]Packer M, et. al. (1985b) J Am Coll Cardiol 5:461.

9 Studies of the diagnostic utility of arbutamine in patients with known or suspected CAD:

9.1. Nonpivotal studies

Three diagnostic studies (120, 112, and 121) are nonpivotal insofar as they generated only isolated estimates of the sensitivity for the diagnosis of CAD without concomitant specificity measures.

9.1.1 Arbutamine-assisted detection of echocardiographic signs of ventricular wall motion abnormality, and EKG ST-segment changes (study 112)

SUMMARY:

This nonblinded, 2 period crossover study sequence-randomized 13 subjects (patients with angiographic CAD and exercise-induced EKG ST-segment depression or elevation) to crossover between "open-loop" iv arbutamine infusion and exercise. Both stress tests employed surface EKG and transthoracic echocardiographic measurements. The objective was to assess the sensitivity, for the diagnosis of CAD, of arbutamine-induced changes in EKG-ST segments, and induced ventricular dysynergy. No specificity estimates were obtained since the study excluded patients who were free of CAD.

PROTOCOL:

- **Enrollment criteria:**

The enrollment criteria were generally comparable to those described below for study 122, except that patients in this study had to have both angiographic evidence of CAD, and exercise-induced EKG ST-segment depression or elevation.

- **Test regimen:**

Patients were sequence-randomized to crossover between exercise and an arbutamine stress test. A 50 µg/ml arbutamine solution was administered (via "open-loop" iv infusion) to supine patients for 8 minutes at each of four intended dose rates (0.1, 0.2, 0.3, and 0.4 µg/kg/min). The range of intended infusion flow rates was, for example, 0.1-0.4 ml/min in a 50 kg subject, and 0.2-0.8 ml/min in a 100 kg subject.

Exercise and arbutamine tests were to be separated by 1-7 days. Exercise tests were performed according to the Bruce treadmill protocol. All tests employed 12-lead surface EKG monitoring, and transthoracic 2D echocardiography performed prior to and immediately following the hemodynamic stress. Standard imaging planes were analyzed centrally by blinded observers (apical two and four chamber views, parasternal short and long axis views).

Study 112

▸ **Endpoints:**

The principle endpoint was the sensitivity, for the diagnosis of CAD, of arbutamine-induced changes in EKG-ST segments, and arbutamine-induced ventricular dysynergy. The criteria for a positive EKG stress test for CAD were more strict than in the other studies in this NDA in that there needed to be induced angina in addition to horizontal or downsloping ST segment depression ≥1 mm, or ≥1 mm ST segment elevation (changes were assessed relative to pre-stress, measurements were obtained 80 msec after the J point, and ST elevation was only assessed in leads without Q waves).

A positive rest-stress echocardiography test for CAD was defined as one in which there was a ventricular wall motion abnormality which was new or worsened (relative to pre-stress, or rest) in at least one segment of the anterior wall, or in at least two adjacent segments of the posterior wall.

▸ **Statistical procedures:**

• **Dataset analyzed:**

Analyses were undertaken on all randomized patients.

• **Handling of missing data:**

No echocardiograms were deemed to be nondiagnostic (according to unspecified retrospective criteria).

• **Analyses performed:**

The analyses assumed equal residual effects.

RESULTS:

• **Demographics:**

The mean patient age was approximately 64 years.

• **Exposure to arbutamine:**

The mean total arbutamine dose and infusion duration were 6.1 μg/kg, and 27 minutes, respectively.

• **Sequence effects:**

Sequence effects were not analyzed (yet the mean time between arbutamine and exercise tests was 4.8 days).

Study 112

▸ **Pharmacodynamic outcomes:**

The hemodynamic changes associated with stress tests are shown in the following table.

Table: 6

Heart rate and systolic BP during stress tests [study 112]

stress test type	**Proportion achieving "target HR"**	**Systolic BP (mm Hg)**	
		pre-test	**maximum increase from pre-test**
arbutamine [n=13]	23%	137	34
exercise [n=13]	85%	131	50

[source: modification of table 8, vol 47; and addendum dated 3/11/94]

Table: 7

Mean arbutamine-associated diastolic BP (DBP) reductions from pre-treatment levels [study 112]

	Arbutamine dose			
	0.1 µg/kg/min	**0.2** µg/kg/min	**0.3** µg/kg/min	**0.4** µg/kg/min
mean change in DBP (mm Hg)	-10.4	-10.5	-14.5	-27.9

[source: modification of table 10, page 60, vol 47]

Arbutamine was administered for ≤ 8 minutes at each dosage level shown. There was no placebo control.

Study 112

• **Rate of decline of tachycardia**:

After the discontinuation of arbutamine infusions (at 0.4 μg/kg/min) the mean time for HR to decrease by 50% was approximately 19 minutes.

▸ **Efficacy outcomes**

• ***Stress-induced* EKG ST-segment changes**:

For the diagnosis of CAD, the sensitivity and specificity of stress-induced EKG ST-segment depression or elevation was as follows:

Table: 8

Mean Sensitivity and False Negative fraction, for the diagnosis of CAD, of stress-induced EKG-ST segment depression or elevation [study 112]

Metric	*Arbutamine-EKG*	*Exercise-EKG*
Sensitivity	85% 95% CI= 55-98% [n=13]	92% 95% CI= 64-100% [n=13]
mean False negative fraction	15%	8%
Specificity	not estimated	not estimated
False positive fraction	not estimated	not estimated

[source: modification of table 14, page 63, vol 47]

The dataset for the above analysis was all randomized patients.

Study 112

▸ **Efficacy outcomes** (continued)

• ***Stress-induced*** **ventricular dysynergy**:

The sensitivity and specificity, for the diagnosis of CAD, of stress-induced ventricular dysynergy was as follows:

Table: 9

Mean Sensitivity and False Negative Fraction, for the diagnosis of CAD, of stress-induced ventricular dysynergy [study 112]

Metric	***Arbutamine-echocardiography***	***Exercise echocardiography***
Sensitivity	62% 95% CI= 32-86% [n=13]	92% 95% CI= 64-100% [n=13]
mean False Negative Fraction	38%	8%
Specificity	not estimated	not estimated
False Positive Fraction	not estimated	not estimated

[source: modification of table 16, page 65, vol 47]

For the above analyses of all-randomized patients, resting (pre-stress) ventricular wall motion abnormalities were not counted.

Study 112

▸ **Efficacy outcomes** [continued]:

• ***Resting vs Stress-induced* ventricular dysynergy**:

Stress echocardiography is a 2-step procedure in which resting ventricular wall motion data are acquired first. The resting data can be used for the diagnosis of CAD (and can be obtained in patients who are unable to exercise). I here present a comparison of the discriminative value of arbutamine-induced vs resting ventricular dysynergy.

Table: 10

Mean Sensitivity and False Negative Fraction, for the diagnosis of CAD, of resting vs stress-induced ventricular dysynergy [study 112]

Metric	*Resting (pre-stress) echocardiography*	*Arbutamine-stress echocardiography*
Sensitivity	31% [n=13]	62% [n=13]
mean False Negative Fraction	69%	38%
Specificity	not estimated	not estimated
False Positive Fraction	not estimated	not estimated

[source: modification of table on the sixth page of addendum dated 8/12/94]

COMMENTS (study 112):

1. This nonplacebo-controlled study suggests (as did study 107) that arbutamine can produce substantial diastolic BP lowering in CAD patients [see the safety review for a discussion of associated clinical outcomes].

2. No specificity estimates (or estimates of false positive fraction) were generated. There is little justification for relying on an isolated estimate of sensitivity made in the absence of a characterization of its associated specificity (neither is there much justification for relying on an isolated estimate of false negative fraction). In such an instance onecannot exclude that sensitivity was only as high as it was because it was obtained at the expense of an exceedingly low specificity. That is, one cannot exclude that the false negative fraction was only as low as it was because it was obtained at the expense of an exceedingly high false positive fraction.

Human observers utilize internal decision thresholds which are to some extent subjectively rendered (this is especially true for echocardiographic and scintigraphic data interpretation). When measuring both sensitivity and specificity, the interdependence of the two metrics (i.e. the interdependence between false negative fraction and false positive fraction) provides checks and balances against the potential biases of subjective decision making. If, as in this study, observers are faced with a situation wherein their positive diagnoses are not evaluated for their incorrectness (i.e. where no specificity or false positive fraction data are generated) it is plausible that an unchecked bias towards lowering of the internal threshold for a positive test could arise, however unintentionally. This would produce higher sensitivity diagnoses (lower false negative fractions), but at the undisclosed price of lower specificity (higher false positive fractions).

3. Compared to like detection methods in this NDA the mean sensitivity for the arbutamine-EKG diagnosis of CAD was here (in study 112) the highest reported, whereas the mean sensitivity for the arbutamine-echocardiographic diagnosis of CAD was the lowest. However, it may be incautious to analyze the data across studies because confounding variables are present to plausibly impact the relative discriminative value of the tests. In this study (112) the criterion for EKG test positivity was the most restrictive in the NDA (it required the presence of angina), and this would be expected to reduce sensitivity (i.e. to increase the false negative fraction). Yet a relatively high proportion of patients achieved "target HR" in this study, and this is generally expected to increase sensitivity (i.e. to lower the false negative fraction).

4. This study (along with study 121) generated pharmacodynamic parameter estimates for the "closed-loop" algorithm used in the pivotal studies. The accuracy of these estimates was said by the sponsor to depend on the unsupported assumption that steady-state was "approached". Infusions continued for 8 minutes at each dose rate, and according to the only kinetic estimate we have been submitted[14] this comprises only about 1 half-life.

[14]this estimate was obtained in healthy subjects administered up to about 0.3 μg/kg/min iv arbutamine.

9.1.2 Arbutamine-assisted detection of echocardiographic signs of ventricular wall motion abnormality, and EKG ST-segment changes (study 121)

SUMMARY:

This nonblinded, 2 period crossover study sequence-randomized 45 subjects (patients with angiographic CAD) to crossover between "open-loop" iv arbutamine infusions, and an exercise stress test. Arbutamine infusion was by ascending titration involving four consecutive steps with intended dose rates of 0.1, 0.2, 0.3, and 0.4 µg/kg/min, respectively. Diagnostic tests employed concomitant surface EKG and transthoracic 2D echocardiographic measurements. The objective was to assess the sensitivity (for the diagnosis of CAD) of arbutamine-induced EKG-ST segment depression or elevation, and induced ventricular dysynergy. No specificity estimates were obtained in this selected sample of patients with angiographic CAD.

PROTOCOL:

▸ **Enrollment criteria:**

Enrollment criteria were generally comparable to those described below for study 122, except that patients in this study had to have angiographic CAD.

▸ **Test regimen:**

Patients underwent a sequence-randomized crossover between arbutamine and exercise. A 50 µg/ml arbutamine solution was administered (via "open-loop" iv infusion device) to supine patients for a maximum of 32 minutes. Infusion was by ascending titration involving four consecutive 8 minute dosage steps with intended dose rates of 0.1, 0.2, 0.3, and 0.4 µg/kg/min, respectively. The range of intended infusion flow rates was, for example, 100-400 µl/min in a 50 kg subject, and 200-800 µl/min in a 100 kg subject.

The second noninvasive test was to take place 1-4 days after the first test. Exercise tests utilized the Bruce bicycle or treadmill protocol. Twelve-lead surface EKG monitoring was undertaken during each stress test, and transthoracic 2D echocardiography was performed (using standard image planes) prior to and during hemodynamic stress.

Study 121

▸ **Endpoints**:

Principal endpoints were the sensitivity (for the diagnosis of CAD) of arbutamine-induced EKG-ST segment changes, and arbutamine-induced ventricular dysynergy. A positive EKG stress test for CAD was defined as one which displayed horizontal or downsloping ST segment depression ≥1 mm, or ≥1 mm ST segment elevation (changes were assessed relative to pre-stress, measurements were obtained 80 msec after the J point, and ST elevation was only assessed in leads without Q waves).

A positive rest-stress echocardiographic test for CAD was defined as one in which there was a new or worsened ventricular wall motion abnormality (relative to pre-stress) observed in at least one segment of the anterior wall, or in at least two adjacent segments of the posterior wall. Echocardiograms were analyzed centrally by two independent observers who were unaware of the angiographic results. A third observer was to be used in cases of discrepancy between the first two observers.

▸ **Statistical procedures**:

• **Dataset analyzed**:

The dataset used for estimation of the sensitivity of arbutamine-echocardiography excluded eleven subjects who did not have echocardiograms.

• **Handling of missing data**:

No subjects had nondiagnostic echocardiogram results (retrospectively, a test would be deemed to be nondiagnostic when 3 or more of the 16 ventricular segments were not reliably visualized).

Study 121

- **Analyses performed:**

When the hypothesis of equal carryover effects was not rejected, the results from periods were used in the analyses. There was no formal testing for treatment by center interactions in this 5 center study.

RESULTS:

- **Demographics:**

The mean subject age was 59 years.

- **Exposure to arbutamine:**

The mean total arbutamine dose and infusion duration were 4.8 μg/kg, and 23 minutes, respectively.

- **Sequence effects:**

Unequal sequence effects were found in the analysis of EKG-ST segment depression or elevation, although this was not the case for the analysis of echocardiographic ventricular dysynergy. The mean time between arbutamine and exercise tests was 1.7 days.

APPEARS THIS WAY
ON ORIGINAL

Study 121

▸ **Pharmacodynamic outcomes**:

Hemodynamic changes associated with stress tests are shown in the following table.

Table: 11

Heart rate and systolic BP during stress tests [study 121]

Stress test type	*Proportion achieving "target HR"*	*Systolic BP (mm Hg)*	
		pre-test	*maximum increase from pre-test*
arbutamine [n=42]	19%	132	33
exercise [n=42]	57%	129	52

[source: modification of table 8, vol 58; and addendum dated 3/11/94]

Analyses are based on the maximum changes from pre-treatment at anytime during the infusion.

• **Diastolic BP changes**:

Diastolic BP responses were highly variable. The maximum diastolic BP change from pre-treatment ranged from -67 to +62 mm Hg, with approximately one fourth of the patients manifesting no change.

• **Rate of decline of tachycardia**:

After the discontinuation of arbutamine (at 0.4 µg/kg/min) the mean time for HR to decrease by 50% was approximately 16.5 minutes

Study 121

▸ **Efficacy outcomes:**

• ***Stress-induced*** **EKG-ST segment changes:**

For the diagnosis of CAD, the sensitivity and false negative fraction associated with stress-induced EKG-ST segment depression or elevation was as follows:

Table: 12

Mean sensitivity and false negative fraction (for the diagnosis of CAD) of stress-induced EKG-ST segment depression or elevation
[study 121]

metric	*Arbutamine-EKG*	*Exercise-EKG*
Sensitivity	73% 95% CI= 50-89% [n=22]	35% 95% CI= 16-57% [n=23]
mean False negative fraction	27%	65%
Specificity	not estimated	not estimated
False positive fraction	not estimated	not estimated

[source: modification of table 13, page 59, vol 58]

In the above table only period 1 results are shown because unequal sequence effects were found. When combining the 2 periods the estimated sensitivity was 58 and 51% for arbutamine vs exercise, respectively.

Study 121

▸ **Efficacy outcomes** [continued]:

• ***Stress-induced*** **ventricular dysynergy**:

The sensitivity and false negative fraction, for the diagnosis of CAD, of stress-induced ventricular dysynergy are shown in the following table.

Table: 13

Sensitivity and False negative Fraction (for the diagnosis of CAD) of stress-induced ventricular dysynergy [study 121]

Metric	*Arbutamine-Echocardiography*	*Exercise-Echocardiography*
Sensitivity	88% 95% CI= 72-97% [n=33]	79% 95% CI= 61-91% [n=33]
mean False negative fraction	12%	21%
Specificity	not estimated	not estimated
False positive fraction	not estimated	not estimated

[source: modification of table 14, page 60, vol 58]

The dataset for the above table was all patients with interpretable echocardiograms. There were no unequal sequence effects observed.

Study 121

▸ **Efficacy outcomes** [continued]:

• ***Resting vs Stress-induced* ventricular dysynergy:**

Stress echocardiography is a 2-step procedure in which resting ventricular wall motion data are acquired first. The resting data can be used for the diagnosis of CAD (and are obviously available in patients who are unable to exercise). I here present a comparison of the discriminative value of arbutamine-induced vs resting ventricular dysynergy.

Table: 14

Mean Sensitivity and False Negative Fraction, for the diagnosis of CAD, of resting vs stress-induced ventricular dysynergy [study 121]

Metric	*Resting (pre-stress) echocardiography*	*Arbutamine-stress echocardiography*
Sensitivity	70% [n=33]	88% [n=33]
False Negative Fraction	30%	12%
Specificity	not estimated	not estimated
False Positive Fraction	not estimated	not estimated

[source: modification of table on the seventh page of addendum dated 8/12/94]

Study 121

COMMENTS (study 121):

1. Compared to the other studies in this NDA the mean sensitivities for the arbutamine-echocardiographic and the arbutamine-EKG diagnoses of CAD were, in this study, the highest vs second highest reported, respectively (i.e. the mean false negative fraction was the lowest vs second lowest, respectively). However, reliance on an isolated sensitivity estimate made in the absence of the characterization of its associated specificity is not well founded (i.e. reliance on an isolated estimate of false negative fraction without knowledge of the associated false positive fraction is not well founded).

2. This study (along with study 112) generated pharmacodynamic parameter estimates for the "closed-loop" algorithm used in the pivotal diagnostic studies in this NDA. The sponsor considers the accuracy of these estimates to be dependent on the assumption that steady-state was "approached," and I find no support for this assumption. Infusions continued for only 8 minutes at each dose rate, and according to the only kinetic estimate which has been submitted this comprises only about about 1 half-life.

9.1.3 Arbutamine-assisted detection of echocardiographic signs of ventricular wall motion abnormality, and EKG ST-segment changes (study 120)

SUMMARY:

This nonblinded, 2 period crossover study sequence-randomized 70 subjects (patients with symptomatic, angiographically-documented CAD) to crossover between 2 infusion regimens of iv arbutamine. A "closed-loop" infusion device intended to achieve a high rate of rise of HR during one period, and a low rate of HR rise during the other period. Surface EKG, and transthoracic 2D echocardiographic measurements were acquired during drug infusions. The sensitivities of arbutamine-induced EKG-ST segment changes, and arbutamine-induced ventricular dysynergy were estimated. The specificity of the test was not estimated in the absence of subjects with negative coronary angiograms.

PROTOCOL:

- **Enrollment criteria:**

Enrolled subjects were patients with symptomatic, angiographically-documented CAD. Excluded from enrollment were pregnant or lactating women and women of childbearing potential as well as those subjects manifesting:
- heart failure, chronic arrhythmias, myocardial infarction or unstable angina within the previous month, cardiac conduction abnormalities, hypertension, idiopathic hypertrophic subaortic stenosis, or aortic stenosis
- cerebrovascular disease, thyroid disease, or glaucoma
- a requirement for β-adrenoreceptor antagonists
- renal or hepatic insufficiency

- **Test regimen:**

Subjects were sequence-randomized to crossover between 2 infusion regimens of iv arbutamine, administered in the supine or semi-recumbent position. A "closed-loop" infusion device intended to achieve a high rate of rise HR during one period, and a low rate of HR ("HR slope") rise during the other period. The intended high rate vs low rate of HR rise was 8-10 bpm/min vs 4-6 bpm/min, respectively.

The two infusions of a 50 µg/ml arbutamine solution took place within 2-7 days of each other. The intended maximum total dose was 10 µg/kg during each period. During the high HR slope treatment period the range of intended dose rates was 0.08-0.8 µg/kg/min, and the range of intended infusion flow rates was, for example, 80-800 µl/min in a 50 kg subject, and 160-1600 µl/min in a 100 kg subject. During the low HR slope treatment period the range of intended dose rates was 0.016-0.8 µg/kg/min, and the range of intended infusion flow rates was, for example, 16-800 µl/min a 50 kg subject, and 32-1600 µl/min in a 100 kg subject.

Study 120

During drug infusion a 12-lead surface EKG was obtained. At 6 of 10 centers a transthoracic 2D echocardiogram were acquired (utilizing standard imaging views) during the pre-treatment period and throughout the arbutamine infusion.

- **Endpoints:**

A positive EKG stress test for CAD was defined as one which displayed
T wave inversion, ST segment depression ≥1 mm, or ≥2 mm ST segment elevation (changes were assessed relative to pre-stress, and measurements were obtained 80 msec after the J point).

A positive rest-stress echocardiographic test for CAD was defined as one in which there was a new or worsened ventricular wall motion abnormality (i.e. any increase in ventricular wall motion score, relative to pre-stress, in at least 1 ventricular segment). Wall motion was assessed after defining 16 ventricular regional segments and applying a scale wherein scores of 1, 2, 3, and 4 represented no abnormality, hypokinesia, akinesia, and dyskinesia, respectively. Echocardiogram analyses were performed at a central site by a single observer who was blinded to the results of other study procedures.

Angina was also an endpoint, although the vaguely specified criteria for angina were questionably valid (e.g. symptoms appear to have been classified as angina even if they were new to the patient).

- **Statistical procedures:**

  - **Dataset analyzed:**

Of the 70 randomized subjects, 61 completed both infusions and were included in the EKG sensitivity analyses. Forty four subjects contributed echocardiographic data, and in 15 of these the results were nondiagnostic.

  - **Handling of missing data:**

Echocardiogram results were defined to be nondiagnostic when 13 or more of the 16 defined ventricular segments were not reliably visualized (and the test was not otherwise positive). The nondiagnostic results were handled by excluding them in one analysis, yet in alternative analyses they were included under the contingent assumptions that they represented a) negative, or b) positive outcomes.

Study 120

• **Analyses performed**:

Testing for equal carryover effects for primary efficacy variables was carried out using a continuity corrected chi-square test at a 0.10 level of significance.

RESULTS:

▸ **Demographics**:

The mean age of subjects was approximately 57 years. Of the 61 analyzed patients, 67% had prior histories of exercise-induced EKG ST segment changes consistent with CAD.

▸ **Arbutamine exposure**:

Although there was no difference in the total amount of arbutamine infused, the mean infusion rate was lower during low HR slope infusions (0.21 μg/kg/min) than it was during the low HR slope infusions (0.24 μg/kg/min).

▸ **Period effects**:

There were no statistically significant period effects (according to the Fisher's Exact test).

▸ **Pharmacodynamic outcomes**:

"Target HR" was achieved in 33% vs 29% of patients during the high vs low HR slope regimens, respectively. During the high vs low slope regimens the maximal mean change in HR was 53.6 vs 53.2 bpm, respectively, the maximal mean increase in systolic BP was 30.3 vs 33 mm Hg, respectively, and the maximal mean increase in diastolic BP was 11.3 vs 12.2 mm Hg, respectively[15]. After the discontinuation of arbutamine the mean time for maximal HR to decrease by 50% was approximately 13-15 minutes.

[15]no attempt was made to isolate the cases of arbutamine-induced hypotension. See Dr. Frank's safety review for discussion of this phenomenon.

Study 120

▸ **Efficacy outcomes:**

• ***Stress-induced* EKG ST-segment changes:**

For the diagnosis of CAD, the sensitivity and specificity of stress-induced EKG ST-segment depression or elevation was as follows:

Table: 15

Mean Sensitivity and False Negative Fraction, for the diagnosis of CAD, of stress-induced EKG ST-segment depression or elevation [study 120]

metric	*Arbutamine-EKG*	
	low HR slope	*high HR slope*
mean Sensitivity	52% [n=61]	51% [n=61]
mean False Negative Fraction	48%	49%
Specificity	not estimated	not estimated
False positive fraction	not estimated	not estimated

[source: modification of table 10, page 63, vol 55]

Study 120

▸ **Efficacy outcomes** [continued]:

• ***Stress-induced*** **ventricular dysynergy**:

The sensitivity and specificity, for the diagnosis of CAD, of stress-induced ventricular dysynergy is shown in the following table.

Table: 16

Mean Sensitivity and False negative fraction (for the diagnosis of CAD) of arbutamine-induced ventricular dysynergy, according to the method of handling of nondiagnostic data [study 120]

Method of handling of nondiagnostic data	***Mean metric***	***Arbutamine echocardiography***	
		low HR slope	***high HR slope***
exclude nondiagnostic tests	**Sensitivity**	83% [n=29]	79% [n=29]
	False negative fraction	17%	21%
assume nondiagnostictests to be negative	**Sensitivity**	66% [n=44]	61% [n=44]
	False negative fraction	34%	39%
assume nondiagnostic tests to be positive	**Sensitivity**	79% ~~82%~~ [n=44]	82% [n=44]
	False negative fraction	18%	18%

[source: modification of table 12, page 65, vol 55]

For the above analysis a test result could be deemed a diagnostic negative (as opposed to nondiagnostic) despite inadequate resolution of as many as 12 of the 16 ventricular regions of interest.

Study 120

▸ **Efficacy outcomes** [continued]:

Angina occurred during 40% of the low HR slope infusions and 35% of the high HR slope infusions. Angina was accompanied by positive EKG changes in 47-52% of the tests and by positive echocardiographic findings in 41% of tests. A negative EKG was associated with angina in 19-33% of the tests, but angina did not occur during any of the negative echocardiogram tests.

COMMENTS (study 120):

1. The reported sensitivities were characterized in the absence of concomitant specificity measures, and one cannot exclude the possibility that specificities were extremely low. In other words, one is unable to exclude the possibility that the reported false negative fractions (which were relatively low) were only attained with concomitantly high false positive fractions.

2. One aspect of this study's design plausibly contributed to a maximization of false positive error, i.e. a minimization of specificity, for the echocardiographic diagnosis of CAD. An apparent wall motion abnormality occurring in just 1 posterior ventricular segment was accepted as evidence of a positive test despite the generally accepted view that there is often times limited resolution of the posterior ventricular wall with transthoracic echocardiography. Investigators typically attempt to avoid false-positive interpretations by specifying that apparent abnormalities must be observed in 2 segments of the posterior wall before the region is declared to be abnormal (as was done in studies 123, 112, 121, and 128). Yet in this study no special measure was taken to avoid this source of false-positive errors.

3. Another aspect of the design of this study plausibly resulted in an underestimation of sensitivity (i.e. an overestimation of false negative fraction) for the echocardiographic diagnosis of CAD. The criterion for a diagnostic test outcome was not adequately restrictive of uncertain results. A test outcome was deemed to be a diagnostic negative (as opposed to nondiagnostic) even if there was inadequate resolution of as many as 12 of the 16 ventricular regions of interest. It would have been reasonable to have used a more restrictive criterion for a diagnostic test. In analyses which exclude nondiagnostic tests such an approach would plausibly result in a higher sensitivity by rendering fewer false negative outcomes.

4. Although EKG T wave inversion is generally accepted to be a nonspecific finding (i.e. one associated with a high false positive fraction), it was taken as evidence for CAD in this study. This criterion plausibly resulted in minimization of EKG specificity (i.e. maximization of false positive fraction).

9.2. Pivotal diagnostic studies

The results of 3 trials, which estimated both sensitivity and specificity in patients with suspected or known CAD (studies 122, 123, and 127), form the pivotal basis for the assessment of the diagnostic utility of stress testing via "closed-loop" arbutamine infusions.

9.2.1 Arbutamine-assisted detection of ST-segment abnormality using surface electrocardiography (study 122)

SUMMARY:

This nonblinded, 2 period crossover study sequence-randomized 243 subjects (patients with suspected or known CAD) to crossover between exercise and a "closed-loop" arbutamine infusion. The objective was to assess the sensitivity and specificity, for the diagnosis of angiographic CAD, of arbutamine-induced surface EKG ST segment changes.

PROTOCOL:

▸ **Enrollment criteria:**

Enrolled subjects were adults of both sexes who were able to exercise, had symptoms and/or signs suggestive of CAD, and had either a positive pre-test coronary angiogram (within 8 weeks of the study), or prestudy findings resulting in plans to perform angiography within 4 weeks subsequent to the study. Excluded from enrollment were pregnant women and women of childbearing potential as well as those subjects manifesting:

- resting EKG ST segment depression or elevation, left bundle branch block, or left ventricular hypertrophy with EKG ST-T changes
- chronic arrhythmias, clinically significant cardiac conduction defects, aortic or subaortic stenosis, or uncontrolled hypertension
- a requirement for the uninterrupted use of β-adrenergic receptor antagonists, atropine, drugs interfering with catecholamine metabolism, class 1 antiarrhythmics, tricyclic antidepressants, digitalis, or amiodarone
- contraindications to exercise or catecholamine exposure
- renal or hepatic insufficiency
- hypokalemia

▸ **Test regimen:**

Subjects were sequence-randomized to crossover between exercise and arbutamine. The two types of stress tests were separated by 1-14 days. A 50 µg/ml arbutamine solution was administered nonblindly to supine or semi-recumbent patients via a "closed-loop" iv infusion device.

Study 122

The device employed algorithm revision #4. The intended rate of rise of HR was 8 bpm/min. The intended dose rates were 0.04-0.8 μg/kg/min, and the intended maximum total dose was 10 μg/kg. The intended infusion flow rates were, for example, 40-800 μl/min in 50 kg subjects, and 80-1600 μl/min in 100 kg subjects.

For exercise tests one of 3 protocols were used (Bruce treadmill (the predominant technique), bicycle, or Naughton treadmill). Twelve-lead surface EKG measurements were obtained prior to and during exercise or arbutamine.

Tests were terminated if patient's manifested any of the following conditions:
- angina or exercise-limiting fatigue
- horizontal or downsloping ST segment depression ≥ 2.0 mm, as measured 60 msec after the J point, or ≥ 2.0 mm ST segment elevation in leads without Q waves
- "target HR" attainment
- adverse events of an "intolerable" [undefined] nature

▸ **Endpoints**:

The principal endpoints were the sensitivity and specificity, for the diagnosis of CAD, of arbutamine-induced EKG-ST segment changes. A positive EKG stress test for CAD was defined as one which displayed
horizontal or downsloping ST segment depression ≥1 mm, or ≥1 mm ST segment elevation (changes were assessed relative to pre-stress, measurements were obtained 60 msec after the J point, and ST elevation was only assessed in leads without Q waves).

The test results were not supposed to influence the decision to perform angiography. Angiograms were evaluated centrally by reviewers who were to have no knowledge of the test results or any other clinical data.

For the arbutamine test, HR and BP measurements were recorded by the device whereas for the exercise test these measurements were captured by the EKG and sphygmomanometer, respectively.

▸ **Statistical procedures**:

• **Dataset analyzed**:

Analyses were performed using a dataset which included those randomized patients with analyzable EKG results of both stress tests, and analyzable coronary angiograms ("intent-to-treat" dataset). An alternative "evaluable" dataset excluded data from patients with various protocol violations.

Study 122

The study randomized 243 patients, of whom 231 had "evaluable" coronary angiograms. Both an arbutamine and exercise test were performed in 222. Of patients with both an arbutamine and exercise test 208 had hemodynamic data recorded on disk by the device. Five patients had nondiagnostic arbutamine test results, 1 had a nondiagnostic exercise test results and 5 had nondiagnostic results of both tests. Among all randomized patients 5% provided no evaluable angiogram. Of the 231 patients with evaluable angiograms, 92% were positive for CAD.

There were 23 patients withdrawn from the study. The reasons for withdrawal included adverse events (AE), intercurrent illness, protocol violations, noncompliance, and laboratory abnormalities.

- **Handling of missing data**:

Patients with missing or nondiagnostic stress test results were excluded from the analyses.

- **Analyses performed**:

Testing for unequal carryover effects was undertaken with a continuity corrected chi-squared test. When the hypothesis of equal carryover effects was rejected, only data from the first period were to be analyzed, using the continuity corrected chi-squared test. No formal analyses of center-by-treatment interactions were undertaken.

RESULTS:

▸ **Timing of angiography**:

Approximately 53% of patients had angiography prior to both noninvasive tests, 40% had it after noninvasive testing, less than 1% had it between tests, and 5% had no angiograms.

▸ **Demographics**:

The randomized population was comprised of 95% white people, and 84% males.

▸ **Period effects**:

It appears that the hypothesis of equal carryover effects was not rejected.

APPEARS THIS WAY
ON ORIGINAL

Study 122

▸ **Pharmacodynamic outcomes**:

Prior to arbutamine tests the mean HRs were 11 bpm lower than prior to exercise [patients were supine for arbutamine tests and, for the most part, standing during exercise]. Pre-test diastolic BPs were comparable in the arbutamine vs exercise groups, as were the maximum mean changes from pre-test diastolic BP.

Hemodynamic changes associated with stress tests are shown in the following table.

Table: 17

Hemodynamic parameters during stress tests (study 122)

stress test type	*Proportion achieving "target HR"*	*Systolic/ Diastolic BP (mm Hg)*	
		pre-test	*maximum increase from pretest*
arbutamine [n=208]	6%	139/83	35/1
exercise [n=208]	52%	134/81	43/1

[source: modification of tables 15, 16 & 20, & figure 1, vol 68; and addendum dated 3/11/94]

The dataset includes patients who had both an arbutamine and an exercise test (for whom the device wrote hemodynamic data to disk).

• **Rate of decline of tachycardia**:

After the discontinuation of arbutamine infusions the mean time for maximal HR to decrease by 50% was approximately 15 minutes (vs 2 minutes following the end of exercise).

Study 122

▸ **Efficacy outcomes**:

• ***Stress-induced*** **EKG ST-segment changes**:

For the diagnosis of CAD, the sensitivity, specificity, and false fractions associated with stress-induced EKG ST-segment depression or elevation were as follows:

Table: 18

Mean sensitivity, specificity, and false fractions (for the diagnosis of CAD) of stress-induced EKG ST-segment depression or elevation [study 122]

metric	***Arbutamine-EKG*** *intent-to-treat*	***Arbutamine-EKG*** *"evaluable" dataset*	***Exercise-EKG*** *intent-to-treat*	***Exercise-EKG*** *"evaluable" dataset*
mean Sensitivity	49% [n=201]	51 [n=172]	46 [n=201]	49 [n=172]
mean False negative fraction	51%	49%	54%	51%
Specificity	74% 95% CI= 49-91 [n=19]	74% [n=19]	84% 95% CI= 60-97 [n=19]	84% [n=19]
mean False positive fraction	26%	26%	16%	16%

[source: modification of tables 22 & 23, pages 125 & 127, vol 68]

Study 122

• **Inter-test agreement between arbutamine and exercise EKG tests:**

The agreement between arbutamine and exercise tests, for the diagnosis of CAD using EKG, was presented by the sponsor, but only after pooling the results of this study with those of studies 123, and 127. Note that the trueness of "positive" or "negative" test outcomes is not validated (relative to the gold standard) in these assessments of inter-test agreement, irrespective of whether one makes the assessment semi-quantitatively or via kappa statistics (this statistic (to be presented in the Statistician's review) merely accounts for the degree of inter-test agreement that is attributable to chance alone).

Table: 19

Agreement between arbutamine and exercise tests, for the diagnosis of CAD using EKG (pooled results of pivotal studies 122, 123, and 127)

		Exercise-EKG		**Totals, for arbutamine**
		Negative results	**Positive results**	
Arbutamine-EKG	**Negative results**	174	48	222
	Positive results	60	135	195
Totals, for exercise		234	183	417

[source: modification of table on page 3 of addendum submission dated 10/7/94]

The dataset includes patients for whom diagnostic arbutamine, exercise, and angiography test results were available.

• **Predictive values:**

As expected in a population such as the one studied (i.e. one with a high prevalence of CAD[16]) the positive predictive values of arbutamine and exercise tests were high (95 and 97%, respectively), and the negative predictive values were low (12 and 13%, respectively).

[16]92% of patients with "evaluable" angiographic data demonstrated evidence of CAD.

Study 122

COMMENTS (study 122):

1. For the diagnosis of CAD using arbutamine-induced surface EKG ST segment changes a medicore mean sensitivity of approximately 50% was found at a moderately high mean specificity of 74% (for which the lower 95% confidence bound was 49%). That is, a mediocre mean false negative fraction of approximately 50% was found in association with a moderately low mean false positive fraction of 26% (for which the upper 95% confidence bound was 51%).

2. The "closed-loop" arbutamine regimen (controlled by algorithm #4) was largely ineffective at attaining "target HR". The algorithm included a parameter which caused the infusion to stop at a HR which was 10 bpm less than "target HR". This approach was based on the assumption that HR would continue to rise after the discontinuation of drug infusions, yet the kinetic studies (obtained in normal subjects) fail to support this assumption insofar as there was no evidence of hysteresis.

3. For the diagnosis of CAD, arbutamine was associated with a roughly comparable sensitivity (i.e. comparable false negative fraction), and roughly comparable specificity (i.e. comparable false positive fraction) to that of exercise.

4. Approximately 20% of the arbutamine tests were stopped because a plateau or decline in HR response was manifest. See the safety review for a detailed discussion of this issue.

APPEARS THIS WAY
ON ORIGINAL

APPEARS THIS WAY
ON ORIGINAL

9.2.2 Arbutamine-assisted detection of ventricular wall motion abnormality using 2-dimensional echocardiography (study 123)

SUMMARY:

This nonblinded, 2 period crossover study sequence-randomized 169 subjects (patients with suspected or known CAD) to crossover between a "closed-loop" iv arbutamine infusion, and an exercise stress test. The objective was to assess the sensitivity and specificity (for the diagnosis of CAD) of arbutamine-induced ventricular dysynergy[17] (as detected by means of transthoracic 2D echocardiography).

PROTOCOL:

- **Enrollment criteria:**

The enrollment criteria were generally comparable to those described above for study 122.

- **Test regimen:**

Patients were sequence-randomized to crossover between iv arbutamine and an exercise stress test. The two stress tests were separated by 1-14 days. Intravenous arbutamine (in 50 µg/ml solution) was administered to supine patients via a "closed-loop" device which employed algorithm revision #4. The intended rate of rise of HR was 8 bpm/min. The intended range of dose rates was 0.012-0.8 µg/kg/min, and the intended maximum total arbutamine dose was 10 µg/kg. The range of intended infusion flow rates was, for example, 12-800 µl/min in a 50 kg subject, and 24-1600 µl/min in a 100 kg subject.

Exercise tests were conducted according to one of 3 protocols (bicycle, Bruce treadmill or Naughton treadmill). Twelve lead surface EKG recording and transthoracic echocardiography was performed continuously during the tests. Standard echocardiography imaging views were used: parasternal long and short axis, apical 2 and 4 chamber.

- **Endpoints:**

The principal endpoints were the sensitivity and specificity, for the diagnosis of CAD, of arbutamine-induced ventricular dysynergy.

The rest-stress echocardiographic method was used in which ventricular wall motion data were first aquired under resting conditions, and again under hemodynamically stressed conditions. A positive resting echocardiographic test for CAD was presumably defined as one which revealed a ventricular wall motion abnormality in at least one segment of the anterior wall, or in at least two adjacent

[17]as noted at the outset of this review, dysynergy includes any of the following wall motion abnormalities: hypokinesis, akinesis, or dyskinesis (paradoxical systolic wall motion).

segments of the posteroinferior wall. A positive stress echocardiographic test for CAD was defined as one which revealed a ventricular wall motion abnormality which was new or worsened (relative to rest) in at least one segment of the anterior wall, or in at least two adjacent segments of the posteroinferior wall[18]. Echocardiograms were analyzed centrally by two independent observers who were unaware of the results of other procedures. A third observer was used in cases of discrepancy between the first two observers. For each set of images the left ventricle was divided into 16 segments and each interpretable segment was assigned a score according to the follow categories: 1 = normal or hyperkinetic, 2= hypokinetic, 3= akinetic, 4= dyskinetic.

Additional endpoints were the sensitivity and specificity, for the diagnosis of CAD, of arbutamine-induced EKG-ST segment elevation or depression. A positive EKG stress test for CAD was defined as one which displayed horizontal or downsloping ST segment depression ≥1 mm, or ≥1 mm ST segment elevation (changes were assessed relative to pre-stress, measurements were obtained 60 msec after the J point, and ST elevation was only assessed in leads without Q waves).

The test results were not supposed to influence any decisions to perform angiography. Angiograms were reported to have been evaluated without the knowledge of the test results or any other clinical data.

- **Statistical procedures:**

  - **Dataset analyzed:**

The study randomized 169 patients, of which 161 had an arbutamine test, 126 had both arbutamine and exercise tests, and 6 withdrew prior to hemodynamic stress testing. Ninety percent had a positive angiogram.
An intent-to-treat analysis excluded all nondiagnostic test results. An "evaluable" dataset excluded 20 patients, the majority of whom were excluded because of the absence of an arbutamine test, or the absence of an "evaluable" angiogram.

  - **Handling of missing data:**

Echocardiogram results were considered to be nondiagnostic when a nonpositive test had poor visualization of least 13 of the 16 defined ventricular segments. Only 9 of 169 randomized patients were deemed to have nondiagnostic echocardiograms.

---

[18]the posterior ventricular wall is sometimes difficult to resolve echocardiographically, thus the criterion for positivity was adjusted (as is conventional) to minimize false positive posterior wall diagnoses.

Study 123

• **Analyses performed**:

Continuity-corrected chi-squared tests were performed to assess for equal period effects.

RESULTS:

▸ **Demographics**:

Eighty three percent of the randomized patients were male, and 89% were white. The age range was 37-86 years.

▸ **Period effects**

For neither the sensitivity nor the specificity analyses was the hypothesis of equal carryover effects rejected.

▸ **Angiographic results**:

Approximately 90% of the patients with angiographic data were found to have CAD. Approximately 64% of patients had angiography performed prior to noninvasive tests, 28% had it performed after noninvasive testing, 4% had angiography performed between tests, and 4% had no angiograms.

[CONTINUED NEXT PAGE]

APPEARS THIS WAY
ON ORIGINAL

APPEARS THIS WAY
ON ORIGINAL

APPEARS THIS WAY
ON ORIGINAL

Study 123

▸ **Pharmacodynamic outcomes:**

Hemodynamic changes associated with stress tests are shown in the following table.

Table: 20

Hemodynamic parameters during stress testing (study 123)

stress test type	*Proportion achieving "target HR"*	*Systolic/ Diastolic BP (mm Hg)*	
		pre-test	*maximum increase from pretest*
arbutamine [n=152]	8%	133/75	37/16
exercise [n=128]	61%	135/81	48/12

[source: modification of tables 15-16, & 19-20, vol 75; and addendum dated 3/11/94]

• **Rate of decline of tachycardia:**

After the discontinuation of arbutamine infusions the mean time for HR to decrease by 50% was approximately 17 minutes (as compared to a mean of 2 minutes for exercise tests).

Study 123

▸ **Efficacy outcomes**:

• ***Stress-induced* ventricular dysynergy**:

For the diagnosis of CAD, the sensitivity, false negative fraction, specificity, and false positive fraction associated with stress-induced ventricular dysynergy was as follows:

Table: 21

Sensitivity, specificity, and false fractions (for the diagnosis of CAD) of *stress-induced* ventricular dysynergy [study 123]

metric	*Arbutamine-echocardiography*		*Exercise-echocardiography*	
	intent-to-treat	*"evaluable" dataset*	*intent-to-treat*	*"evaluable" dataset*
Sensitivity	76% [n=127]	76% [n=119]	77% [n=98]	75% [n=83]
mean False Negative Fraction	24%	24%	23%	25%
Specificity	31% 95% CI= 11-59% [n=16]	31% [n=16	55% 95% CI= 23-83% [n=11	55% [n=11
mean False Positive Fraction	69%	69%	45%	45%

[source: modification of tables 22 & 24, pages 119 & 121; vol 75]

▸ **Efficacy outcomes** [continued]:

• ***Resting vs Stress-induced*** **ventricular dysynergy**:

Stress echocardiography is a 2-step procedure in which the first step is the aquisition of resting ventricular wall motion data. The resting data can be used for the diagnosis of CAD (and is obviously available for patients who are unable to exercise). The rationality of proceeding with the arbutamine administration, in the face of known results of resting echocardiograms, depends on the relative sensitivities and specificities of the arbutamine-stressed vs resting data [this issue is discussed in detail in the "Analytic framework" section (i.e. section 11) of this review]. What follows is a comparison of the discriminative value of arbutamine-induced vs resting ventricular dysynergy.

Table: 22

Mean sensitivity, specificity and false fractions (for the diagnosis of CAD) of resting vs stress-induced ventricular dysynergy [study 123]

Metric	*Resting (pre-stress) echocardiography*	*Arbutamine-stress echocardiography*
Sensitivity	40% 95% CI= 32-49% [n=127]	76% 95% CI= 67-83% [n=127]
False negative fraction	60%	24%
Specificity	75% 95% CI= 48-93% [n=16]	31% 95% CI= 11-59% [n=16]
False positive fraction	25%	69%

[source: modification of table 22, page 119, vol 1.75; & tables in addenda dated 8/12/94, and 9/7/94]

The analyses are based on the intent-to-treat dataset of patients with diagnostic echocardiograms.

Study 123

▸ **Efficacy outcomes** [continued]:

• **Inter-test agreement between arbutamine & exercise echocardiography tests:**

The agreement between arbutamine and exercise echocardiography, for the diagnosis of CAD, is presented below. In these assessments of inter-test agreement the trueness of "positive" or "negative" test outcomes is not validated (relative to the angiographic gold standard), irrespective of whether one makes the assessment semi-quantitatively or via kappa statistics (see the Statistician's review for presentation of this statistic).

Table: 23

Agreement between arbutamine and exercise tests, for the diagnosis of CAD using echocardiography (study 123)

		Exercise-Echocardiogram		Totals, for arbutamine
		Negative results	**Positive results**	
Arbutamine-echocardiogram	**Negative results**	11	5	16
	Positive results	11	85	96
Totals, for exercise		22	90	112

[source: modification of table on page 3 of addendum submission dated 10/7/94]

The dataset includes patients for whom diagnostic arbutamine, exercise, and angiography test results were available.

Study 123

▸ **Efficacy outcomes** [continued]:

• ***Stress-induced* EKG ST-segment changes**:

The discriminative value, for the diagnosis of CAD, of stress-induced EKG ST-segment depression or elevation is shown in the following table.

Table: 24

Mean sensitivity, specificity, and false fractions (for the diagnosis of CAD) of *stress-induced* EKG ST-segment depression or elevation [study 123]

	Arbutamine-EKG	*Exercise-EKG*
metric	*intent-to-treat*	*intent-to-treat*
Sensitivity	49% 95% CI= 40-59% [n=112]	47% 95% CI= 38-57% [n=112]
False negative fraction	51%	53%
Specificity	82% 95% CI= 48-98% [n=11]	64% 95% CI= 31-89% [n=11]
False positive fraction	18%	36%

[source: modification of tables 42 & 43, vol 75]

Study 123

▸ **Efficacy outcomes** [continued]:

• **Inter-test agreement between arbutamine and exercise EKG tests**:

The agreement between arbutamine and exercise tests, for the diagnosis of CAD using EKG, was presented after pooling the results of this study (123) with those of studies 122 and 127. See table 19 of this review for a description of these pooled data.

• **Predictive values**:

As expected in a population of the type studied (i.e. one with an approximately 90% prevalence of CAD) the positive predictive values of arbutamine and exercise tests were high (92 and 94%, respectively), and the negative predictive values were low (16 and 21%, respectively).

COMMENTS (study 123):

1. For the diagnosis of CAD arbutamine stress-echocardiography had a statistically significantly higher mean sensitivity than did resting echocardiography (76 vs 40%, respectively). That is, adjuvant arbutamine had a significantly smaller mean false negative fraction than did resting echocardiography (24 vs 60%, respectively).

2. The design of this study plausibly resulted in an underestimation of sensitivity (i.e. an overestimation of false negative fraction) for the echocardiographic diagnosis of CAD (both under resting and stressed conditions). A test outcome was deemed to be a diagnostic negative (as opposed to nondiagnostic) even if there was inadequate resolution of as many as 12 of the 16 ventricular regions of interest. It would have been reasonable to have used a more restrictive criterion for a diagnostic test while excluding nondiagnostic tests from the principal analyses. Such an approach would have plausibly resulted in a higher sensitivity by rendering fewer false negative results.

3. For the diagnosis of CAD arbutamine-stress echocardiography tended to have a lesser mean specificity than did resting (untreated) echocardiography (31% (with a lower 95% confidence bound of 11%)) vs 75%, respectively) in a comparison which was quite limited in power. The basis for this apparent difference has not been elucidated. If real, it could plausibly be attributable (at least in part) to the induction of regional myocardial hypoxia via supra-physiologic arbutamine-mediated perturbations of coronary flow and/or myocardial oxygen demand in regions subtended by stenoses which fell minimally below the threshold for the definition of CAD (which denoted a 50% reduction in cross-sectional diameter).

4. For a given sensitivity (of approximately 76%) the echocardiographic diagnosis of CAD tended to be less specific with arbutamine than it was with exercise (31 vs 55% mean specificity, respectively). That is, for a given mean false negative fraction there was a nonsignificant tendency towards a higher mean false positive fraction with arbutamine, relative to exercise.

COMMENTS (study 123) [continued]:

5. The ideal pharmacologic echocardiography test would be expected to outperform exercise echocardiography (with respect to the correct discrimination of CAD vs absent CAD). Assuming all else being equal, only exercise is handicapped by the degradation in cardiac image quality that results from induced hyperventilation.[19] Yet arbutamine (as administered) did not meet this ideal since the pharmacologic stress tests did not demonstrate superior sensitivity or specificity to exercise tests for the echocardiographic diagnosis of CAD (i.e. arbutamine was not associated with lower false negative or false positive fractions than was exercise).

6. For the diagnosis of CAD via arbutamine-EKG tests a mediocre mean sensitivity (49%) was found at a high mean specificity (82%, with the lower bound of the 95% confidence interval being 48%). That is, a mediocre mean false negative fraction (51%) was found concomitant to a low mean false positive fraction (18%).

For a given sensitivity (of approximately 48%) the EKG diagnosis of CAD tended to be more specific with arbutamine than it was with exercise (82 vs 64%, respectively), however the confidence intervals around these point estimates were overlapping. That is, for a given false negative fraction there was a nonsignificant tendency towards a lesser false positive fraction with arbutamine-EKG tests, relative to exercise-EKG tests.

7. As in the other pivotal studies, the "closed-loop" arbutamine regimen (controlled by algorithm #4) was largely ineffective at attaining "target HR".

APPEARS THIS WAY
ON ORIGINAL

APPEARS THIS WAY
ON ORIGINAL

[19]intrapulmonary air is a poor medium for the transthoracic propagation of the interrogating ultrasound beam.

9.2.3 Arbutamine-assisted detection of ventricular hypoperfusion using thallium-201 scintigraphy (study 127)

SUMMARY:

This nonblinded, 2 period crossover study sequence-randomized 149 subjects (patients with suspected or known CAD) to crossover between "closed-loop" iv arbutamine infusion and an exercise stress test. Each test employed scintigraphic ventricular perfusion imaging using the stress-redistribution thallium-201 technique. The objective was to assess the sensitivity and specificity, for the diagnosis of CAD, of arbutamine-associated ventricular hypoperfusion.

PROTOCOL:

▸ **Enrollment criteria:**

The enrollment criteria were generally comparable to those described above for study 122. Enrolled patients were to have undergone angiography within 12 weeks prior to the study, or to have been scheduled for angiography within 4 weeks after the test.

▸ **Test regimen:**

Patients were sequence-randomized to crossover between arbutamine and exercise stress tests. Arbutamine in 50 µg/ml solution was administered iv to supine patients via a "closed-loop" device which employed algorithm revision #4. The intended rate of rise of HR was 8 bpm/min. The range of intended dose rates was 0.078-0.8 µg/kg/min, and the intended maximum total dose was 10 µg/kg. The range of intended infusion flow rates was, for example, 78-800 µl/min in a 50 kg subject, and 156-1600 µl/min in a 100 kg subject.

Exercise tests were performed according to one of 3 protocols (bicycle, Bruce treadmill or Naughton treadmill). Twelve-lead surface EKG measurements were obtained at rest and during stress tests.

Thallium-201 (2.5 to 3.5 mCi) was injected iv during peak tachycardia. Initial post-stress SPECT ventricular perfusion images were obtained beginning 10 minutes after peak tachycardia[20] (these images will be referred to as "immediate post-stress" images to distinguish them from the further delayed images which are used for the identification of thallium "redistribution"[21]).

[20]the brief lag was presumably introduced to reduce the "upward creep" artefact which has been attributed to the cephalad movement of the dome of the diaphragm during the immediate recovery from exercise.

[21]when the regional deficit of myocardial tracer activity (as observed on the immediate post-stress image) becomes less marked during later imaging (i.e. when the "image defect fills in") this is referred to as "thallium redistribution."

Study 127

▸ **Endpoints**:

The principle endpoints were the sensitivity and specificity (for the diagnosis of CAD) of the finding of ventricular hypoperfusion. Hypoperfusion was considered to be detected when abnormally low regional ventricular tracer uptake was observed (by at least 2 of 3 blinded observers) on the immediate post-stress perfusion image, irrespective of subsequent changes in apparent tracer distribution. Low ventricular tracer uptake was defined as a moderate deficit in observed tracer scintillations (relative to normally perfused regions) in at least 2 of 20 defined ventricular regions, or a severe deficit in 1 ventricular region.

A positive EKG stress test for CAD was defined as one which displayed horizontal or downsloping ST segment depression ≥1 mm, or ≥1 mm ST segment elevation (changes were assessed relative to pre-stress, and measurements were obtained 60 msec after the J point).

Test results were not supposed to influence decisions to perform angiography. Angiograms were analyzed centrally by blinded reviewers.

▸ **Statistical procedures**:

• **Dataset analyzed**:

Among 149 randomized patients, 136 performed an arbutamine test, 75 performed both an arbutamine and an exercise test, and 20 withdrew (for reasons including adverse events, intercurrent illnesses, and noncompliance). The intent-to-treat dataset included all randomized patients with available data, and excluded unevaluable or missing results. The "evaluable" dataset excluded data from patients who did not comply with various protocol provisions (e.g. those for whom there were no evaluable arbutamine or angiographic data).

• **Handling of missing data**:

Patients with missing data were excluded from the intent-to-treat dataset.

• **Analyses performed**:

Continuity corrected chi-squared tests were performed to assess for equal period effects.

Study 127

RESULTS:

• Demographics:

Eighty four percent of randomized patients were male, and 97% were white. The mean age was 59 years.

▸ Period effects

The hypothesis of equal carryover effects was rejected for the intent-to-treat analyses of exercise-thallium tests for CAD. For these, only period 1 data were used to generate estimates of sensitivity and specificity.

▸ Angiographic results:

Approximately 90% of the patients with angiographic data were found to have CAD. Approximately 77% of patients had angiography prior to noninvasive tests, 14% had it after noninvasive testing, and 9% had no angiograms.

[CONTINUED NEXT PAGE]

APPEARS THIS WAY
ON ORIGINAL

APPEARS THIS WAY
ON ORIGINAL

APPEARS THIS WAY
ON ORIGINAL

Study 127

▸ Pharmacodynamic outcomes:

• *Stress-associated* hemodynamic changes:

Hemodynamic changes associated with stress tests are shown in the following table.

Table: 25

Hemodynamic parameters associated with stress testing [study 127]

stress test type	*Proportion achieving "target HR"*	*Systolic/ Diastolic BP (mm Hg)*	
		pre-test	*maximum increase from pre-test*
Arbutamine [n=133]	3%	138/81	35/8.9
Exercise [n=75]	68%	135/81	40/5.8

[source: modification of tables 17-20, vol 82; and addendum dated 3/11/94]

• Rate of decline of tachycardia:

The proportion of patients who had at least a 50% decrease in HR by 10 minutes post stress was 20 vs 96% for the arbutamine vs exercise tests, respectively [after 10 minutes the EKG leads were removed from patients' chests, presumably to reduce image artefacts resulting from photon attenuation].

Study 127

▸ **Efficacy outcomes:**

• ***Stress-associated* ventricular hypoperfusion:**

For the diagnosis of CAD, the disease-discriminative value of the finding of ventricular hypoperfusion (as observed immediately after peak stress) was as follows:

Table: 26

Mean sensitivity, specificity, and false fractions (for the diagnosis of CAD) of regional ventricular hypoperfusion observed immediately after peak hemodynamic stress [study 127]

metric	*Arbutamine-Thallium perfusion imaging*		*Exercise-Thallium perfusion imaging*	
	intent-to-treat	*"evaluable" dataset*	*intent-to-treat*	*"evaluable" dataset*
Sensitivity	87% [n=112]	86% [n=87]	94% [n=33]	95% [n=44]
mean False negative fraction	13%	14%	6%	5%
Specificity	25% 95% CI= 3-65% [n=8]	25% 95% CI= 3-65% [n=8]	80% 95% CI= 28-100% [n=5]	80% 95% CI= 28-100% [n=5]
mean False positive fraction	75%	75%	20%	20%

[source: modification of tables 22-23, vol 82; and addendum dated 3/11/94]

The intent-to-treat analysis of exercise-thallium data included only period 1 results because there was a significant carryover effect. In the subset of arbutamine-exposed patients who also had exercise tests the 95% CI for arbutamine sensitivity (in the diagnosis of CAD) was 88-100%, and the 95% CI for specificity was 7-93% [intent to treat].

Study 127

▸ **Efficacy outcomes:**

• **Inter-test agreement between arbutamine and exercise thallium tests:**

The agreement between arbutamine and exercise tests, for the diagnosis of CAD using thallium perfusion imaging, is presented below. In these assessments of inter-test agreement the trueness of "positive" or "negative" test outcomes is not validated (relative to the gold standard), irrespective of whether one makes the assessment semi-quantitatively or via kappa statistics (see the Statistician's review for presentation of this statistic).

Table: 27

Agreement between arbutamine and exercise tests,
for the diagnosis of CAD using thallium perfusion imaging (study 127)

		Exercise-Thallium		**Totals, for arbutamine**
		Negative results	**Positive results**	
Arbutamine-Thallium	**Negative results**	3	3	6
	Positive results	0	61	61
Totals, for exercise		3	64	67

[source: modification of table on page 4 of addendum submission dated 10/7/94]

The dataset includes patients for whom diagnostic arbutamine, exercise, and angiography test results were available.

• **Predictive values:**

As expected in a population of the type studied (i.e. one with a high prevalence of CAD) the positive predictive values of arbutamine and exercise-thallium tests were high (94 and 95%, respectively), and the negative predictive values were low (12 and 33%, respectively).

Study 127

▸ **Efficacy outcomes** [continued]:

• ***Stress-induced* EKG ST-segment changes**:

For the diagnosis of CAD, the disease discriminative value of stress-induced EKG ST-segment changes was as follows:

Table: 28

Mean sensitivity, specificity and false fractions (for the diagnosis of CAD) of stress-induced EKG ST-segment depression or elevation [study 127]

metric	*Arbutamine-EKG*	*Exercise-EKG*
Sensitivity	51% [n=124]	43% [n=69]
mean False negative fraction	49%	57%
Specificity	80% (44-98% [n=10]	80% (28-100%) [n=5]
mean False positive fraction	20%	20%

[source: modification of table 37, page 150, vol 82; and addendum dated 3/11/94]

The analyses are based on the intent-to-treat dataset. Ninety five percent confidence intervals are shown in parentheses.

• **Agreement between arbutamine and exercise EKG tests**:

The agreement between arbutamine and exercise tests, for the diagnosis of CAD using EKG, was presented only after pooling the results of this study (127) with those of studies 122 and 123. See table 19 of this review for a description of these pooled data.

Study 127

COMMENTS (study 127):

1. For the diagnosis of CAD via arbutamine-thallium perfusion imaging the mean sensitivity was high (87%), and the mean specificity was low (25%, with the lower bound of the 95% confidence interval being only 3%). That is, the mean false negative fraction was low (13%), and the mean false positive fraction was high (75%).

2. For the CAD diagnosis, the arbutamine-thallium tests had a roughly comparable mean sensitivity to that of exercise-thallium tests (87 vs 94%, respectively), but arbutamine tended to have a lesser mean specificity (25 vs 80%, respectively), although the confidence intervals for this comparison overlapped. That is, at roughly comparable false negative fractions there was a nonsignificant tendency towards a higher false positive fraction with arbutamine, relative to exercise.

3. For the diagnosis of CAD via arbutamine-EKG testing the mean sensitivity was mediocre (approximately 51%) as was the lower bound of the 95% confidence interval for specificity (44%). That is, the mean false negative fraction was mediocre (49%) as was the upper bound of the 95% confidence interval for false positive fraction (56%).

4. As in the other pivotal studies, the "closed-loop" arbutamine regimen (controlled by algorithm #4) was largely ineffective at attaining "target HR."

10 Results overview:

10.1 *Pivotal studies overview (122, 123, and 127)*:

10.1.1 *Overall pivotal results*:

The results of the 3 trials (studies 122, 123, and 127) which provided estimates of both the sensitivity and specificity (i.e. both the false negative and false positive fractions) of "closed-loop" arbutamine testing for CAD form the pivotal basis for the evaluation of diagnostic utility. The following 2 tables provide summaries of the results of arbutamine tests for the presence of CAD in pivotal patient studies. Table 29 expresses the findings in terms of sensitivity and specificity, and table 30 expresses results in terms of false negative and false positive fractions.

Table: 29

Overall results of arbutamine tests for the presence of CAD in pivotal patient studies, expressed in terms of **Sensitivity and Specificity** (studies 122, 123, and 127)

study	*sample sizes*	*% achieving "target HR"*	*EKG*		*Echocardiography*		*Thallium perfusion imaging*	
			SENS	*SPEC*	*SENS*	*SPEC*	*SENS*	*SPEC*
#127	n= 10-124	3%	51%	80% (44-98%)	no data	no data	87%	25% (3-65%)
#122	n= 19-201	6%	49%	74% (49-91%)	no data	no data	no data	no data
#123	n= 11-127	8%	49%	82% (48-98%)	76%	31% (11-59%)	no data	no data

Shown are the mean sensitivities (SENS) and specificities (SPEC) for the diagnosis of CAD, as determined in the intent-to-treat datasets. The 95% confidence intervals are depicted in parentheses. Each study utilized a "closed-loop" arbutamine administration.

Results overview

10.1.1 *Overall pivotal results* [continued]:

In the following table the pivotal results are expressed in terms of false negative fraction and false positive fraction.

Table: 30

Overall results of arbutamine tests for the presence of CAD in pivotal patient studies, expressed in terms of **False Negative Fraction and False Positive Fraction** (studies 122, 123, & 127)

study	*sample sizes*	*% achieving "target HR"*	*EKG*		*Echocardiography*		*Thallium perfusion imaging*	
			False Negative fraction	*False Positive fraction*	*False Negative fraction*	*False Positive fraction*	*False negative fraction*	*False positive fraction*
#127	n= 10-124	3%	49%	20% (2-56%)	no data	no data	13%	75% (35-97%)
#122	n= 19-201	6%	51%	26% (9-51%)	no data	no data	no data	no data
#123	n= 11-127	8%	51%	18% (2-52%)	24%	69% (41-89%)	no data	no data

Shown are the mean false fractions as determined in the intent-to-treat datasets. The 95% confidence intervals are depicted in parentheses. Each study utilized a "closed-loop" arbutamine administration.

Results overview

10.1.2 *EKG stress test results, according to demographic subgroup*:

The results of the pivotal studies were pooled by the sponsor for analysis of EKG sensitivity according to age, race, and sex (pooled specificity and false negative fraction data were not submitted). These results are summarized in the following table.

Table: 31

Pooled EKG Sensitivity and False Negative Fraction (for the diagnosis of CAD) according to sex, age, and race in pivotal studies [studies 122, 123, and 127]

	Sex		*Age*		*Race*		
	Female	*Male*	*≥70 years*	*<70 years*	*Black*	*Caucasian*	*Other*
sample size	n= 71	n= 396	n= 70	n= 397	n= 11	n= 442	n= 14
Sensitivity	55% (43-67%)	48% (43-53%)	63% (51-74%)	47% (42-52%)	55% (23-83%)	48% (44-53%)	64% (35-87%)
False Negative Fraction	45% (33-57%)	52% (47-57%)	37% (26-49%)	53% (48-58%)	45% (17-77%)	52% (47-56%)	36% (13-65%)

[source: modification of tables 8.1-8.3, pages 179-182, volume 95]

Shown are the mean sensitivities and mean false negative fractions (for the diagnosis of CAD). The 95% confidence intervals are depicted in parentheses.

10.1.3 *Comparisons between arbutamine and exercise stress tests*:

Both gold standard-validated comparisons between arbutamine and exercise tests (for the diagnosis of CAD), and assessments of inter-test agreement (between arbutamine and exercise tests) will be overviewed in this section.

Results overview

10.1.3 *Comparisons between arbutamine and exercise tests* [continued]:

10.1.3.1 *Gold standard-validated comparisons of arbutamine and exercise*:

Gold standard-validated comparisons between arbutamine and exercise tests (for the diagnosis of CAD in pivotal studies) are shown in the following two tables. In table 32 the data are expressed in terms of sensitivity and specificity, and in table 33 they are expressed in terms of false positive and false negative fractions.

Table: 32

Overall results of arbutamine vs exercise tests for CAD in pivotal patient studies, expressed in terms of **Sensitivity and Specificity** (studies 122, 123, and 127)

study	*metric*	*EKG*		*Echocardiography*		*Thallium scintigraphy*	
		Arbutamine	*Exercise*	*Arbutamine*	*Exercise*	*Arbutamine*	*Exercise*
#122	**Sensitivity**	49%	46%	no	data	no	data
	Specificity	74% (49-91%)	84% (60-97%)	no	data	no	data
#123	**Sensitivity**	49%	47%	76%	77%	no	data
	Specificity	82% (48-98%)	64% (31-89%)	31% (11-59%)	55% (23-83%)	no	data
#127	**Sensitivity**	51%	43%	no	data	87%	94%
	Specificity	80% (44-98%)	80% (28-100%)	no	data	25% (3-65%)	80% (28-100%)

Shown are the mean sensitivities and specificities for the diagnosis of CAD (intent-to-treat). The 95% confidence intervals are depicted in parentheses.

As shown in the above table, for the EKG diagnosis of CAD the mean sensitivities and specificities were roughly comparable for arbutamine and exercise tests. For both the echocardiographic and thallium-based scintigraphic diagnoses of CAD, arbutamine had roughly comparable mean sensitivity and nonsignificantly lesser specificity than did exercise.

Results overview

10.1.3.1 *Gold standard-validated comparisons* [continued]:

Gold standard-validated comparisons between arbutamine and exercise tests are shown below, expressed in terms of false negative and false positive fractions.

Table: 33

Overall results of arbutamine vs exercise tests (for the diagnosis of CAD) in pivotal patient studies, expressed in terms of **False negative and False Positive fractions** (studies 122, 123, and 127)

study	*metric*	*EKG*		*Echocardiography*		*Thallium perfusion imaging*	
		Arbutamine	*Exercise*	*Arbutamine*	*Exercise*	*Arbutamine*	*Exercise*
122	**False Negative fraction**	51%	54%	no data	no data	no data	no data
	False Positive fraction	26% (9-51%)	16% (3-40%)	no data	no data	no data	no data
#123	**False Negative fraction**	51%	53%	24%	23%	no data	no data
	False Positive fraction	18% (2-52%)	36% (11-69%)	69% (41-89%)	45% (17-77%)	no data	no data
#127	**False Negative fraction**	49%	57%	no data	no data	13%	6%
	False Positive fraction	20% (2-56%)	20% (0-72%)	no data	no data	75% (35-97%)	20% (0-72%)

Shown are the mean false negative and false positive fractions for the diagnosis of CAD (intent-to-treat). The 95% confidence intervals are depicted in parentheses.

As shown above, for the EKG diagnosis of CAD the mean false negative and false positive fractions were roughly comparable (or statistically indistinguishable) for arbutamine and exercise. For both the echocardiographic and thallium-based scintigraphic diagnoses of CAD, arbutamine had roughly comparable mean false negative fraction, and nonsignificantly higher false positive fraction than did exercise.

Results overview

10.1.3 *Comparisons between arbutamine and exercise tests* [continued]:

10.1.3.2 *Inter-test agreement between arbutamine and exercise tests*:

Note that the trueness of "positive" or "negative" test outcomes is not validated (relative to the gold standard) in the following assessments of inter-test agreement, irrespective of whether the assessment is undertaken semi-quantitatively or via kappa statistics. The kappa statistic (to be presented in the Statistician's review) merely accounts for the the degree of inter-test agreement that is attributable to chance alone.

The agreement between arbutamine and exercise tests, for the diagnosis of CAD using EKGs, is summarized below.

Table: 34

Agreement between arbutamine and exercise tests, for the diagnosis of CAD using EKG (pooled results of pivotal studies 122, 123, and 127)

		Exercise-EKG		Totals, for arbutamine
		Negative results	Positive results	
Arbutamine-EKG	Negative results	174	48	222
	Positive results	60	135	195
Totals, for exercise		234	183	417

[source: re-presentation of table 19 of this review]

APPEARS THIS WAY
ON ORIGINAL

APPEARS THIS WAY
ON ORIGINAL

Results overview

10.1.3.2 *Inter-test agreement between arbutamine & exercise tests* [continued]:

For the echocardiographic diagnosis of CAD, the agreement between arbutamine and exercise tests is depicted in the following table.

Table: 35

Agreement between arbutamine and exercise tests, for the diagnosis of CAD using echocardiography (study 123)

		Exercise-Echocardiogram		Totals, for arbutamine
		Negative results	Positive results	
Arbutamine-echocardiogram	Negative results	11	5	16
	Positive results	11	85	96
Totals, for exercise		22	90	112

[source: re-presentation of table 23 of this review]

For the thallium scintigraphic diagnosis of CAD, the agreement between arbutamine and exercise tests is depicted in the following table.

Table: 36

Agreement between arbutamine and exercise tests, for the diagnosis of CAD using thallium scintigraphy (study 127)

		Exercise-Thallium		Totals, for arbutamine
		Negative results	Positive results	
Arbutamine-Thallium	Negative results	6	6	12
	Positive results	2	53	55
Totals, for exercise		8	59	67

[source: re-presentation of table 27 of this review]

Results overview

10.2 *Nonpivotal study overview (112, 120, and 121)*:

In these 3 studies isolated sensitivity estimates were generated in the absence of specificity measures (i.e. estimates of false negative fraction were obtained without estimates of false positive fraction). There is little justification for relying on an isolated estimate of sensitivity made in the absence of a characterization of its associated specificity (neither is there much justification for relying on an isolated estimate of false negative fraction). In such an instance one cannot exclude that sensitivity was only as high as it was because it was obtained at the expense of an exceedingly low specificity (i.e. one cannot exclude that the false negative fraction was only as low as it was because it was obtained at the expense of an exceedingly high false positive fraction)[22].

The results of nonpivotal studies are summarized as follows:

Table: 37

Sensitivity and False Negative fraction (for the diagnosis of CAD) of arbutamine tests in Nonpivotal studies (studies 112, 120, 121)

study	*sample size/ infusion type*	*proportion achieving "target HR"*	*Mean Sensitivity*			*Mean False Negative fraction*		
			EKG	*Echo*	*Thallium*	*EKG*	*Echo*	*Thallium*
#121	n= 42; open loop	19%	73%	88%	no data	27%	12%	no data
#112	n= 13; open loop	23%	85%	62%	no data	15%	38%	no data
#120	n= 30-60; closed loop	33%	51%	79%	no data	49%	21%	no data

See the main body of this review for the 95% confidence intervals around these point estimates.

[22]this issue is further discussed above in the comments attached to the review of study 112.

11 Analytic framework

The following viewpoints characterize my conceptual schema for analyzing whether a diagnostic adjunct is approveable for marketing (for the sake of the discussion the adjunct is assumed to be adequately safe, and the infusion system is assumed to be adequately safe and functioning according to engineering specifications).

11.1 Regulatory importance of clinical benefit:

My view is that an approveable diagnostic adjunct is one which facilitates the making of a diagnostic finding which has a reasonably convincing net clinical impact on outcome.[23] I do not consider it to be justified to approve a diagnostic adjunct in the absence of a compelling basis for inferring that there is a net clinical benefit attendant to its use. Direct demonstrations of such benefit would be definitive, yet these have not in the past been requested of sponsors during the development of diagnostic agents.[24]

11.2 Clinically beneficial cardiac imaging-based diagnoses:

11.2.1 *Diagnostic exclusion of epicardial CAD*:

There is a convincing basis for inferring that the diagnostic exclusion of significant epicardial coronary stenosis (i.e. that producing ≥50% reduction in cross-sectional diameter) confers clinical benefit. The correct "ruling out" of such disease provides a rational basis for withholding therapeutic interventions[25] and sparing patients from the adverse events attributable to such interventions. Additionally, this diagnostic exclusion provides a rational basis for deciding to spare such patients from the risks of coronary angiography. The available data (from sources outside of this NDA) fully support the view that coronary angiography is, in the absence of a reliably predictive negative result of a noninvasive test, the rational diagnostic step to be taken (in order to determine the location of CAD and/or the number of diseased vessels[26]). Survival trials in patients with stable angina indicate that initial treatment choices (between coronary artery bypass graft (CABG) surgery vs medical therapy)

---

[23]that is, where the attendant benefit outweighs the attendant risk.

[24]this policy appears to be in flux insofar as our sister center, the Biologics group, has published the view that the development of diagnostic imaging agents should include clinical outcome data [see "Points to Consider in the Manufacture and Testing of Monoclonal Antibody Products for Human Use" in the 8/3/94 Federal Register].

[25]there may prove to be exceptions to this. For example, it may be that some patients with angiographically inapparent CAD have abnormal coronary microvascular that will prove to be an appropriate target of the same therapeutic interventions demonstrated to be of value in patients with epicardial CAD.

[26]neither arbutamine nor any other noninvasive stress test methodology has been proven to reliably determine these anatomic parameters.

are optimized on the basis of *a priori* determinations of the angiographic parameters of disease location and multiplicity of vessel involvement. In their recently published meta-analysis of CABG trials Yusuf et. al.[27] found that a strategy of initial CABG surgery significantly extended the survival of stable angina patients with epicardial CAD (relative to the survival of initially medically treated patients who were offered delayed CABG surgery as needed), but the mortality benefit appeared to be limited to patients with triple-vessel, left main artery, or proximal LAD disease.

11.2.2 *Diagnostic inclusion of epicardial CAD*:

Under some clinical circumstances (albeit relatively narrow ones) there is a reasonable basis for inferring that a test would plausibly provide net benefit merely by identifying the presence of CAD (without characterizing disease location and/or the numbers of vessels stenosed).[28] Consider the patient who experiences the morbidity of stable chest pain, for whom there is a contraindication to CABG surgery, in whom tests for a noncardiac etiology of symptoms are equivocal, and in whom symptoms are not alleviated by nonpharmacologic therapy. The rational decision to introduce drug therapy for such a patient introduces the risk of adverse events. An empirically-based choice of therapeutic agent (i.e. a choice made without consideration of noninvasive stress testing data) will, at least in some scenarios, be associated with a risk/benefit ratio which would plausibly be less favorable than that associated with therapy chosen on the basis of the results of an ideal noninvasive stress test (i.e. one which safely and reliably[29] identifies the presence and absence of CAD). For example, if the empiric therapy is an antianginal and the patient does not actually have CAD there would be expected to be little-to-no symptomatic benefit yet nonzero adverse effects of the antianginal.[30] For such a patient, the *a priori* consideration of a true negative outcome of an ideal stress test would result in the benefit of the patient being spared the risk of adverse effects attributable to antianginal exposure.

On the other hand, if CABG surgery is not contraindicated, noninvasive test-derived knowledge is not sufficient for determining whether a stable angina patient with some sort of CAD is likely to experience a survival benefit of CABG surgery. The above cited meta-analysis of Yusuf et. al., despite providing greater statistical power than the individual trials had provided, was unable to demonstrate that patient survival hinged on the initial treatment strategy (medical vs surgical) unless three vessel disease, or a proximal left anterior descending artery (LAD) stenosis were present. Since localizing CAD to this anatomic segment or identifying that 3 vessels are diseased requires

---

[27]Yusuf S, et. al. Lancet 1994; 344: 563-570.

[28]these anatomic parameters are not reliably determined by noninvasive imaging methods.

[29]i.e. with low false positive fraction, and low false negative fraction.

[30]a plausible exception is the patient who experiences a beneficial placebo-type (or otherwise poorly understood) effect, and for whom the symptom reduction outweighs the risk of adverse drug effects.

angiography, noninvasive diagnostic tests do not, by and large, confer the benefit of sparing a patient the risks of angiography.

11.2.3 *Diagnosis of ischemia*:

Various imaging methods are purported to make the diagnosis of ischemia (i.e. to differentiate 2 subgroups of CAD-related pathology: ischemia vs infarction). There is not widespread agreement as to which (if any) method is the gold standard for the diagnosis. Nonetheless, data-dependent inferences about the pathophysiologic validity of the inclusion diagnosis of ischemia (i.e. the exclusion of the infarct diagnosis) are obtainable if a sponsor undertakes prospective studies of the test's ability to predict the improvement of ventricular wall motion after revascularization of putatively noninfarcted ischemic regions, or perhaps studies controlled with various positron-labelled diffusable tracer methodologies.

11.3 Regulatory importance of comparisons between pharmacologic stress tests and exercise stress tests:

In my analytic framework the approvability of a pharmacologic stress test is evaluated without consideration of the correlation between the pharmacologic test and the exercise test. This conceptual model is adopted for the following reasons:

a) only the gold standard comparisons (sensitivity, specificity, false negative fraction, false positive fraction), and not the degree of agreement between pharmacologic tests and exercise tests, provide insights into the diagnostic validity of the test. The kappa statistic can indicate that the test results are in clearcut agreement even if neither the pharmacologic test nor the exercise test diagnoses what it purports to diagnose.

b) with at least some types of data aquisition (e.g. electrocardio-graphic monitoring), studies of exercise testing for CAD often demonstrate a lack of validity (i.e. exercise-EKG tests often fail to yield appreciably lower proportions of incorrect diagnoses than do random tosses of a coin). There is no greater rational for basing decisions about the approvability of a pharmacologic-EKG test on positive control exercise tests than there is for basing regulatory decisions about the efficacy of an investigational antidepressant on studies which utilize a marketed antidepressant[31] as positive control.

c) with some types of data aquisition (e.g. thallium perfusion imaging) the validity of exercise testing for CAD is generally considered to be higher when patients achieve a moderate-to-high exercise workload than it is when they achieve only a low workload. Studies of pharmacologic stressors rarely (if ever) employ a control group which exercises to the low capacity which is representative of the

[31]these agents being notorious for the inconsistency with which they demonstrate more efficacy than does placebo.

potential market population (i.e. "patients who cannot exercise adequately"[32]). There is no greater rational for deciding about the approvability of a pharmacologic-thallium test on the basis of typical positive control exercise studies (in which patients exercise to moderate or high workloads) than there is for basing approval decisions for an investigational therapeutic agent upon trials in which the unproven agent is administered to an unselected population while the positive control group is enriched with drug responders. Even if the investigational agent were to not compare favorably to the positive control in such a trial, it could plausibly prove to be as good or better than the control when compared in a population which is representative of the group proposed for market exposure.

11.4 Regulatory importance of high false positive or false negative fractions:

If the goals of testing are to both reliably identify disease when it is present, and to reliably exclude disease when it is absent (as are commonly the dual goals of a diagnostician) then the only test which has utility is one which provides both high sensitivity and high specificity (i.e. low false negative fraction, and low false positive fraction).

However, the clinical goal of diagnostic testing is sometimes focused on obtaining evidence of the absence of a disease. For this "exclusion diagnosis", in its narrowest sense, the goal is to be confident that a negative test has excluded the disease (one in pursuit of this finite goal would be content to conclude, in the face of a positive test outcome, that one has failed to exclude disease despite being unsure as to whether this failure reflects a limitation of the test). A test with low sensitivity (i.e. high false negative fraction) would fail to meet even this narrow goal. Yet for this purpose a test with high sensitivity (i.e. low false negative fraction) can be useful despite a low specificity (i.e. despite a high false positive fraction). Thus the finding that a test has a high false positive fraction (low specificity) is not a sufficient basis for its non-approval since it may still have utility for those whose diagnostic need is limited to having confidence in a negative test outcome.

In certain sample populations (in particular, those with a very low prevalence of disease) the exclusion diagnosis can be reliably obtained even if sensitivities and specificities are no higher than those achieved by the toss of a coin. In such a sample the correct diagnosis will still be achieved in a high fraction of patients with negative test outcomes (i.e. the predictive value of a negative test result will remain high) because negative predictive value is inversely proportional to disease prevalence.

Alternatively, the goal of diagnostic testing is other times focused on obtaining evidence of the presence of a disease. For this "inclusion diagnosis", in its narrowest sense, the goal is to be confident that a positive test has identified the disease (one in pursuit of this circumscribed goal would be satisfied to conclude, in the face of a negative test outcome, that one has failed to include disease despite being unsure as to whether this failure reflects a limitation of the test). A test with low specificity (i.e. high false positive fraction) fails to meet even this narrow goal. Yet for the inclusion diagnosis a test with high specificity (i.e. low false positive fraction) can be useful even if sensitivity

[32]as per the approved indication for i.v. dipyridamole.

is low (i.e. even if the false negative fraction is high). Thus the finding that a test has a high false negative fraction (low sensitivity) is not a sufficient basis for its non-approval since it may still have utility for those wishing to apply it for the limited purpose of diagnostic inclusion.

In a sample with very high prevalence of disease the inclusion diagnosis can even be reliably obtained despite sensitivities and specificities which are as low as those achieved by the toss of a coin (i.e. 50%). In such a sample the correct diagnosis will still be achieved in a high fraction of patients with positive test outcomes (i.e. the predictive value of a positive test result will remain high) because positive predictive value is proportional to disease prevalence.

11.5 Rational employment of a stressor in rest-stress echocardiography:

If one seeks only the *exclusion diagnosis* of CAD (as defined above in section 11.4) by means of the rest-stress echocardiographic method, and one observes a negative resting echocardiogram, it is not under all circumstances rational to proceed with the administration of the stressor. It would only be rational to do so under those conditions in which a negative result of the stress test is more highly indicative of the absence of disease than is a negative result of the resting test. Such conditions are present when the disease prevalence is not extremely low in the population being tested, and when the sensitivity of stress echocardiography is higher than that associated with resting echocardiography (i.e. when the false negative fraction for stress-echocardiography is lower than that associated with rest echocardiography). In contrast, when the disease prevalence is extremely low the predictive value of a negative resting test would itself be high (even if the resting test had relatively mediocre sensitivity), and it would not be rational to administer the stressor.

Similarly, if one seeks only the *inclusion diagnosis* of CAD by means of rest-stress echocardiography, and one observes a positive resting echocardiogram, it is not under all circumstances rational to proceed with the administration of the stressor. It would only be rational to do so under those conditions in which a positive result of the stress test is more highly indicative of the presence of disease than is a positive result of the resting test. Such conditions are present when the disease prevalence is not extremely high, and the specificity of stress echocardiography is higher than that associated with resting echocardiography (i.e. the false positive fraction for stress-echocardiography is lower than that associated with rest echocardiography). In contrast, when the disease prevalence is extremely high the predictive value of a positive resting test would itself be high (even if the resting test has relatively mediocre specificity), and it would not be justified to administer the stressor.

Analytic framework

11.6 Regulatory importance of data replication for each imaging modality:

There was no attempt on the part of the sponsor to replicate the evidence obtained with any particular imaging modality (EKG, echocardiography, or thallium scintigraphy). At least with respect to the echocardiographic and thallium data, the variability of clinical responses and diagnostic interpretation make it undesirable to make decisions on the basis of single studies per imaging modality. However, the minutes from previous meetings with the sponsor indicate that FDA agreed that such replication would not be required. In deference to this previous judgement, the present analytic scheme assumes that the reported findings are replicable.

11.7 Regulatory importance of the "normalcy rate" metric:

The "normalcy rate" metric is not a greatly useful estimator of test specificity in the population for whom any indication might be approveable (i.e. in patients with suspected CAD). This metric (defined as the ratio of negative test outcomes to all test outcomes) is measured in healthy subjects who are presumed to be devoid of cardiac pathophysiology. It plausibly overestimates the specificity (and underestimates the false positive fraction) of pharmacologic-stress tests in patients with suspected CAD [see the previous comments attached to the review of study 128].

[CONTINUED NEXT PAGE]

12 Conclusions

The following conclusions are rendered on the basis of the 3 pivotal studies (studies 122, 123, and 127) in which arbutamine was administered via the "close-loop" infusion device (controlled by algorithm revision #4). These conclusions are contingent upon the findings that clinical issues related to safety, device-performance, or data integrity do not prove to adversely impact the approvability of this NDA.

12.1 Arbutamine-EKG testing:

12.1.1 *Data basis for approval judgments*:

Across the 3 pivotal studies of approximately 470 patients (studies 122, 123, and 127) the values of the prevalence-independent measures of arbutamine-EKG tests for the diagnosis of epicardial CAD were concordant. A mediocre mean sensitivity (ranging from 49-51%) was found at a mean specificities of 74-82%, and the lower bound of the 95% confidence interval for specificity was mediocre (ranging from 44-49%). That is, a mediocre false negative fraction (49-51%), was found to be associated with a mediocre lower 95% confidence bound for false positive fraction (51-56%).

Compared to exercise, for the EKG diagnosis of CAD, arbutamine had roughly comparable mean sensitivity (and false negative fraction), and roughly comparable mean specificity (and false positive fraction)

Although the sample populations generally lacked demographic heterogeneity in the pivotal studies, subgroup comparisons of sensitivity were undertaken by the sponsor. In low powered analyses there was overlap of the 95% confidence intervals for sensitivity between subgroups of age, race, and sex. That is, the false negative fraction for the EKG diagnosis of CAD was not statistically differentiable between demographic subgroups.

12.1.2 *Approval of arbutamine-EKG testing for the diagnostic Exclusion of CAD*:

The arbutamine-EKG test did not manifest adequately high sensitivity (i.e. adequately low false negative fraction) to make it clinically useful even as an agent to be used for the narrowly defined goals of an exclusion diagnosis of CAD.

12.1.3 *Approval of arbutamine-EKG testing for the diagnostic Inclusion of CAD*:

Neither did the arbutamine-EKG test have adequately high specificity (i.e. adequately low false positive fraction) to make it clinically useful even as an agent to be used for the narrowly defined goals of an inclusion diagnosis of CAD [although the mean specificity findings may appear to suggest otherwise, these point estimates are not adequately reliable given the wide confidence intervals].

It is true that in a population with very low prevalence of disease the goals of the exclusion diagnosis could be met by even the mediocre sensitivity provided by the arbutamine-EKG test, and it is likewise true that in a population with very high disease prevalence the goals of the inclusion diagnosis could be met by the mediocre specificity of the arbutamine-EKG test. However, each of these goals would be adequately achieved by the predictive values rendered by a random guess about the presence or absence of CAD. Since exposure to arbutamine would undoubtedly be associated with a higher frequency of adverse events than would be the process of tossing a coin, there is no convincing basis for the approval of the arbutamine-EKG test for either the inclusion diagnosis of CAD or the exclusion diagnosis of CAD.

In conclusion, I find the submitted evidence for the utility of arbutamine-EKG testing for the diagnosis of CAD to be inadequate to support its approval for marketing.

12.2 Arbutamine-echocardiography testing:

12.2.1 *Data basis for approval judgments*:

Study 123 provided data on the sensitivity and specificity (and associated false fractions), for the diagnosis of epicardial CAD, of arbutamine-induced ventricular dysynergy (as detected by 2D rest-stress echocardiography) in 169 patients with suspected or known CAD.

Arbutamine stress-echocardiography had a reported mean sensitivity of 76% for the diagnosis of CAD (i.e. a mean false negative fraction of 24%). It is plausible that this was an underestimate of true sensitivity (i.e. that the false negative fraction was overestimated). A test outcome was deemed to be a diagnostic negative (as opposed to nondiagnostic) even if there was inadequate resolution of as many as 12 of the 16 ventricular regions of interest. The use of a more restrictive criterion for a diagnostic test, and the exclusion of nondiagnostic tests from the principal analyses, would have been reasonable. Such an approach would have plausibly resulted in a higher sensitivity by the rendering of fewer false negative results.

Arbutamine stress-echocardiography had a mean sensitivity for the diagnosis of CAD which was statistically significantly higher than that of resting echocardiography (76 vs 40%, respectively). That is, the mean false negative fraction for arbutamine stress-echocardiography was significantly lower than that associated with resting echocardiography (24 vs 60%, respectively).

The mean specificity of arbutamine-echocardiography tests, for the diagnosis of CAD, was quite low (31%, with a 95% confidence bound of 11%). That is, the arbutamine-echocardiography tests had a high mean false positive fraction (69%).

The mean specificity of arbutamine-echocardiography (31%) at least tended to be lower than that of resting echocardiography tests (75%), although the difference could not be readily distinguished statistically given the small sample sizes. That is, the arbutamine tests tended to have a higher mean false positive fraction (69%) than did resting echocardiography (25%).

Conclusions

Compared to exercise, for the echocardiographic diagnosis of CAD, arbutamine had roughly comparable mean sensitivity (false negative fraction) and tended to have lesser specificity (higher false positive fraction).

12.2.2 *Approval of arbutamine-echocardiography for the diagnostic Exclusion of CAD*:

Arbutamine rest-stress echocardiography manifested clinically useful properties as an agent for use in excluding the diagnosis of CAD. The test sensitivity for the diagnosis of CAD was reasonably high,[33] and this sensitivity was significantly higher than that associated with resting echocardiography (in other words, arbutamine echocardiography had a false negative fraction for the diagnosis of CAD which was reasonably low, and significantly lower than that associated with resting echocardiography). This evidence is adequate to support the approval of arbutamine-echocardiographic for use in excluding the diagnosis of CAD. Nonetheless, in populations with extremely low disease prevalence the administration of arbutamine affords no incremental utility for the diagnostic exclusion of CAD when the resting echocardiogram test is negative for CAD (in such circumstances the resting test is itself highly indicative of the absence of disease).

12.2.3 *Approval of arbutamine-echocardiography for the diagnostic Inclusion of CAD*:

There is no convincing basis for approving arbutamine-echocardiography for use in obtaining the inclusion diagnosis of CAD. Arbutamine-echocardiography did not manifest the high specificity for CAD which is necessary to provide utility for the inclusion diagnosis. In other words, arbutamine-echocardiography did not demonstrate the low false positive fraction for the diagnosis of CAD that would make it useful for this diagnostic purpose.

Approving arbutamine-echocardiography for diagnostic confirmation of a positive resting echocardiogram finding would only be supportable if the arbutamine-induced findings had higher specificity for CAD (i.e. lower false positive fraction) than did resting echocardiography, but they did not.

Although the application of arbutamine-echocardiography to a population with very high disease prevalence would, despite mediocre test specificity, result in high positive predictive values (which would provide utility for the inclusion diagnosis of CAD), in such a population this diagnostic goal would be adequately achieved by the high positive predictive value rendered by the safer diagnostic method of tossing a coin.

[33]there is an additional margin of comfort with this judgement insofar as it is plausible that the reported mean sensitivity of 76% was an underestimate.

Conclusions

12.2.4 *Approval of arbutamine-echocardiography for the dual Inclusion and Exclusion of CAD*:

The property of low specificity makes unsupportable any approval for marketing the arbutamine-echocardiography test for the dual purpose of including and excluding the diagnosis of CAD.

12.2.5 *Approval of arbutamine-echocardiography for the diagnosis of ischemia*:

There is no adequate basis to support the approval of arbutamine-echocardiographic imaging for the putative diagnosis of myocardial ischemia (i.e. for the putative differentiation of ischemia from infarct). There have been no data submitted to support the pathophysiologic validity of this diagnostic differentiation, and drawing inferences about possible attendant clinical benefit is incautious in the absence of such data.

12.3 Arbutamine-thallium testing:

12.3.1 *Data basis for approval judgments*:

Study 127 provided data on the sensitivity and specificity (and associated false fractions), in the diagnosis of CAD, of arbutamine-associated ventricular hypoperfusion (as detected with thallium perfusion imaging). For the diagnosis of CAD a high mean sensitivity of 87% (i.e. a low false negative fraction of 13%) was obtained with thallium-201 images acquired immediately after arbutamine-stress. Yet for this diagnosis a low mean specificity of 25% (i.e. a high false positive fraction of 75%) was also observed.

Compared to exercise, for the thallium-based scintigraphic diagnosis of CAD, arbutamine had roughly comparable mean sensitivity (and false negative fraction) and nonsignificantly lesser specificity (i.e. a tendency towards higher mean false positive fraction).

12.3.2 *Approval of arbutamine-thallium testing for the diagnostic Exclusion of CAD*:

The submitted data support the approval of arbutamine-thallium testing for the diagnostic exclusion of CAD.

12.3.3 *Approval of arbutamine-thallium testing for the diagnostic Inclusion of CAD*:

The submitted data are not adequate to support the approval of arbutamine-thallium testing for the diagnostic inclusion of CAD. The method did not manifest an adequately high specificity for CAD to provide utility for this purpose. In other words arbutamine-thallium testing did not manifest the adequately low false positive fraction which would give it utility for this purpose.

Conclusions

12.3.4 *Approval of arbutamine-thallium testing for the dual Inclusion and Exclusion of CAD*:

The property of low specificity makes unsupportable any proposal to market the arbutamine-thallium test for the dual purpose of including and excluding the diagnosis of CAD.

12.3.5 *Approval of arbutamine-thallium testing for the diagnosis of Ischemia*:

There is not an adequate basis to support the approval of arbutamine-thallium imaging for the putative diagnosis of myocardial ischemia (i.e. the putative differentiation of ischemia from infarct). There have been no data submitted which describe the validity of this diagnostic differentiation, and the thallium literature points to shortcomings of accuracy in this regard (at least when non-late redistribution imaging, and non-reinjection techniques are used[34]). Developing inferences about possible attendant clinical benefit is incautious in the absence of basic validation data.

12.4 Approval of arbutamine for use without the GenESA® System device):

It is unlikely that the submitted data will be adequate to support approval for the administration of arbutamine by conventional infusion devices which are not controlled by the studied "closed-loop" algorithm. The "open-loop" data were inadequately few for detection of anything other than high frequency adverse events. The implication of this view is that if the device data fail to support approval of the "closed-loop" infusion device, then there would be no support for the approval of arbutamine[35].

12.5 Approval of the infusion algorithm as revised since the pivotal trials:

In an attempt to improve the rate of attainment of "target HR" the sponsor has, since conducting the pivotal trials with algorithm revision #4, again revised the infusion algorithm. Although it is plausible that the sensitivity of arbutamine testing could improve with this revision (algorithm revision #5), it is also plausible that there would be a higher rate of dose-related adverse events, relative to the rate observed in the pivotal trials. Despite this, the sponsor proposes algorithm revision #5 for marketing. Since the clinical safety of arbutamine infusions produce by algorithm revision #5 has not been empirically studied, there is no adequate basis for approving this marketing proposal.

[34]Tillisch J, et. al. N Eng J Med 1986; 314:884-888.

[35]nonetheless I do not find the rationale for using a closed-loop device to be very convincing. I know of no data which supports the sponsor's concept that it could be more harmful to expose CAD patients to a rate of rise of HR which is less constant than that which an algorithm can produce. Exercise (which appears to be safer than "closed-loop" arbutamine administration) routinely produces nongradual rates of HR rise, particularly in hemodynamically deconditioned subjects.

Conclusions

12.6 Administration of arbutamine soon after exercise testing:

It should be anticipated that arbutamine, if marketed, will in some cases be used relatively immediately subsequent to an exercise test. It is plausible, because of persisting endogenous adrenergic activation post-exercise, that there are increased risks of arbutamine when administered soon after recovery from exercise (relative to the risks described in fully resting patients in this NDA).

The rapid sequential administration of exercise and arbutamine tests thallium would conceivably occur for several reasons. Large numbers of patients referred for exercise testing are unable to reach "target HR" (witness the nearly 50% rate of failure to reach this target in the largest of the pivotal studies (122) in this NDA).[36] The sensitivity of exercise testing is believed to be reduced when a subject fails to attain "target HR." Accordingly, when the sensitivity of a negative exercise test is suspected to be low because of the failure to achieve "target HR" during exercise, clinicians not uncommonly seek alternative noninvasive means (such as pharmacologic stress tests) to reliably exclude the presence of CAD.[37] Patient "throughput" being a concern of diagnostic laboratory faciltites, some clinicians will undoubtedly be inclined to perform the pharmacologic stress test relatively immediately subsequent to an exercise test which they suspected to yield a false negative result.

[CONTINUED]

[36]my own unpublished data confirms a rate of this magnitude (and even somewhat higher). In my series, the duration of exercise was commonly limited on the basis of dyspnea developing in a hemodynamically deconditioned subject. Despite there being a relatively higher maximum HR than a well-conditioned subject would have manifested at a like exercise duration, the duration of exercise in these subjects was sufficiently limiting to result in failure to achieve 85% of the age-predicted maximum HR.

[37]if not prone to spontaneously behave this way clinicians may end up doing so in response to published advertisements (put forth by the sponsors of i.v. dipyridamole) which have encouraged such use of a pharmacologic stress test.

13 **Recommendations**:

I make the following recommendations contingent upon our finding, in the final analysis, that clinical issues related to safety, device-performance, or data integrity do not prove to adversely impact the approvability of this NDA. These recommendations pertain specifically to arbutamine as administered by the GenESA® device utilizing infusion algorithm revision #4.

13.1 Recommendations regarding Approval:

13.1.1. *Arbutamine-Echocardiography test*:

I recommend that arbutamine **be approved** as an adjunct for use with 2 dimensional echocardiography for the purpose of the diagnostic exclusion of epicardial CAD in patients who cannot exercise adequately.

13.1.2 *Arbutamine-Thallium test*:

I recommend that arbutamine **be approved** as an adjunct for use with thallium-201 perfusion imaging for the purpose of the diagnostic exclusion of epicardial CAD in patients who cannot exercise adequately.

13.1.3 *Arbutamine-EKG test*:

I recommend **nonapproval** of arbutamine as an adjunct for use with electrocardiography for the diagnostic inclusion or exclusion of CAD.

13.1.4 *Infusion algorithm revision #5*:

I recommend **nonapproval** of the sponsor's request to market their device with infusion algorithm revision #5.

13.2 General recommendation:

I recommend that prior to any market launch of this product the sponsor should be encouraged to demonstrate the degree of effectiveness of potentially useful pharmacologic intervention(s) aimed at reversing serious toxic sequelae of arbutamine use (e.g. arterial hypotension).

13.3 Recommendations regarding Labelling:

13.3.1 *Language of each indication*:

I recommend the we avoid use of the language, "indicated for the evaluation of CAD" which is found in the i.v. dipyridamole label. By use of the nonspecific term "evaluation of CAD" we give the appearance of advocating the multitude of applications of stress testinf that are found in clinical use,[38] most of which are not adequately supported by the submitted arbutamine data. We should specify exactly which diagnostic use has been adequately supported by the submitted data.

13.3.2 *Arbutamine-echocardiography test*:

13.3.2.1 *Description of the echocardiography method*:

I recommend that the descriptor, "2-dimensional echocardiography" be used in the indication for the diagnostic exclusion of CAD, as opposed to the term, "echocardiography". With the recommended language we avoid the appearance of advocating the use of arbutamine with diagnostic Doppler velocimetry[39] which, in imaging parlance, often falls under the general category of "echocardiography".

13.3.2.2 *Caveat regarding populations with low CAD prevalence*:

I recommend that the label provide a caveat (with appropriate emphasis) which describes the following:

- in populations with low disease prevalence, when the resting echocardiogram is negative for CAD-diagnostic ventricular wall motion abnormality this finding itself reliably established CAD exclusion diagnosis.

- since the resting echocardiogram test is safer than arbutamine echocardiography, and the latter method provides no marginal utility in this setting, the administration of arbutamine is not warranted under these circumstances.

---

[38] for example, uses such as for the putative diagnoses of ischemia, myocardial viability, "area-at-risk", and coronary flow reserve deficiency, as well as various prognostic applications in patients with known CAD diagnoses.

[39] a method used to generate putative diagnoses of coronary flow reserve deficit.

13.3.2.3 *CAD-diagnostic ventricular wall motion abnormalities*:

The label should define in simple terms the range of pertinent ventricular wall motion abnormalities which defined the diagnosis of CAD in the NDA database: these encompass abnormally reduced regional wall motion (hypokinesis), totally absent regional wall motion (akinesis), and paradoxical systolic wall motion away from the center of the ventricle (dyskinesis).

13.3.3 *Noncitation of Predictive Values of arbutamine tests*:

I recommend that the label should not cite the predictive values observed in the pivotal clinical trials. These indices are population-specific, for example, the reported positive predictive values were biased towards high numbers by the extremely high disease prevalence in the studied populations. The inclusion of positive predictive value invites mistaken expectations on the part of clinicians who apply the test to less highly selected populations, and invites the misleading use of the metric in drug advertising.

13.3.4 *Arbutamine testing immediately following exercise*:

I recommend that the label describe the potential risks of rapid sequential exposure to exercise stress followed by arbutamine. The label should suggest the imposition of a delay between exercise and the subsequent administration of arbutamine.

Steven M. Rodin, MD   10/26/94

Steven M. Rodin, MD   Date
Medical Officer

cc: BuehlerG/HFD-110; division file (NDA 20-420); ***no copy to SRodin**