4. See Serious Adverse Events which follows.

Serious Adverse Events:

A 77 yr. old F was hospitalized for **hypotension, syncope, fever, pain and redness in her neck**. These symptoms developed 3-4 days after she received her second Thyrogen injection. Thyroiditis was ruled out as the cause of her neck pain and erythema. She responded to IV fluids and a course of steroids and was discharged. The hypotension was determined by the investigator to be either possibly related to Thyrogen or to volume depletion.

Overall Adverse Events During the Thyrogen Phase:

42/152 (28%) patients reported adverse events (AES) during the Thyrogen phase. The most commonly reported AES were nausea (25/152= 16%), headache (7/152= 5%), asthenia (5/152= 3%), nausea/vomiting (3/152= 2%) and dizziness/light-headedness (3/152= 2%). These were also the most commonly reported AES considered by the investigators to be related to Thyrogen (possibly, probably or definitely).

The following AES occurred in 2 patients each (1%): chills, diarrhea, vomiting and rash.

The following AES occurred in 1 patient each (<1%):

General: pain, abdominal pain, neck pain
Skin: maculopapular rash
Respiratory: dyspnea
CV: pulmonary embolus, hypotension, vasodilation
GI: diarrhea, oral monilia and thirst
Nervous system: confusion, insomnia and nervousness

Vital Signs:

The only notable change was the mean pulse rate was lower after withdrawal of thyroid hormone than after administration of Thyrogen (78 $\pm$ 12 bpm vs. 68 $\pm$ 11 bpm, $p< 0.001$).

Serum Hematology, Chemistries and Urine Results:

These laboratory parameters were measured at enrollment and prior to the administration of radioiodine for the Thyrogen and WD scans. Changes in these laboratory parameters were categorized as low to normal, low to high, normal to high, normal to low, high to low and high to normal. These shift tables were used to illustrate the changes in the laboratory parameters from baseline to the Thyrogen phase (i.e. visit 1 to visit 3 which was 1 day after the second Thyrogen injection), from baseline to the WD phase (visit 1 to visit 6 which is ~2 weeks after d/c T3 or ~ 4 weeks after d/c T4) and from the Thyrogen phase to the WD phase (visit 3 to visit 6). Mean and

median values during each study phase (Thyrogen and WD) and changes in mean/median values from baseline and between study periods were not provided.

Results:

Changes in labs during the Thyrogen phase:

No clinically significant trends in abnormal laboratory values were identified during and immediately after the Thyrogen phase.

Changes in labs during the WD phase:

Approximately, 15-30% patients experienced increases in cholesterol, triglycerides, uric acid, SGOT, creatinine and LDH, consistent with the known effects of hypothyroidism. The following information is derived from the published review of this study by Ladenson et al, NEJM 337(13): 888-896, 1997: "The patients had significantly higher mean serum concentrations of cholesterol (by 66%), triglycerides (by 70%), uric acid (by 24%), and creatinine (by 44%) after withdrawal of thyroid hormone than after administration of thyrotropin ($P<0.001$ for all)."

Development of Thyrogen Antibodies:

No patient had detectable serum antibodies specific to Thyrogen. This included 7 patients who had received Thyrogen 7 to 16 months earlier.

Protocol TSH95-0101:

The main differences between this second phase III study and the first (TSH 92-0601) were:

1. Patients with hemi- or partial thyroidectomy were excluded in study TSH95-0101 because, in TSH92-0601, the presence of residual thyroid tissue may have confounded interpretation of the diagnostic scans and Tg results.
2. In TSH95-0101, patients were monitored more closely for documentation of suppressed TSH levels ($\leq$ 0.5 mU/L) prior to Thyrogen administration.
3. In TSH95-0101, diagnostic scans were to be conducted for a minimum acquisition time or number of counts using a gamma camera. For whole body images, the recommended scanning speed was < 10 cm/min so that a target count of 140,000 counts for the anterior image could be reached. For spot images, an attempt was made to reach 60,000 counts for a large anterior image and 35,000 counts for small. Due to limits of patient comfort, a maximum of 15 minutes was established for acquiring spot images. Scans deemed inadequate by the independent reviewers were not included in the primary efficacy analyses.
4. The diagnostic activity of ^{131}I administered for scanning was increased from 2-4 mCi (74- 148 Mbq) in study TSH92-0601 to 4 mCi ($\pm$ 10%) (148 Mbq) in TSH95-0101 to ensure that sufficient ^{131}I was available for uptake during Thyrogen imaging.

This parameter was chosen based on results of an independent study conducted in a subset of 33 patients from protocol TSH92-0601 which indicated that the clearance rate of ^{131}I was ~2 fold greater in the euthyroid study (i.e. on Thyrogen) than when patients were hypothyroid during THST withdrawal. These differences in clearance are consistent with the well-documented decrease in creatinine clearance that accompanies the hypothyroid state. Furthermore, the whole body ^{131}I retention was 50% less in the euthyroid state than the hypothyroid state. This decreased retention may have resulted in less bioavailability of ^{131}I in the euthyroid state compared to the hypothyroid state.

5. In study TSH92-0601, the diagnostic scans were rated "concordant" if they showed the same stage, number and distribution of lesions. The reviewers stated that since a difference in the exact number of lesions was not always clinically meaningful, concordance should be defined based on location and extent of disease. For TSH95-0101, the sponsor devised a classification system which they felt reflected clinically meaningful differences and it was this classification system that was used to rate the scans. If, for example, one scan documented the presence of more clinically relevant thyroid remnant (note: this was a subjective assessment made by the independent reviewers), then the scans were rated "discordant".

6. In study TSH92-0601, only a 48 hr. whole body scan (WBS) was obtained. In TSH95-0101, a WBS could also be obtained 72 hrs. after the radioiodine scanning dose was administered if the additional 24 hr. clearance of ^{131}I from the serum and normal sites of uptake would help delineate an actual focus of uptake in thyroid remnant or cancerous tissue from background or from normal sites of uptake.

7. In study TSH92-0601, each diagnostic scan was independently assessed prior to the paired comparison of the scans for each patient. The independent reviewers were blinded as to which scan of a given pair was the Thyrogen scan and which was the withdrawal scan. However, they did know which scan pair belonged to a given patient. To reduce the number of observations that needed to be made by the independent reviewers, only a comparison between the scans for a given pair was conducted in TSH95-0101.

8. If a patient were treated with radioiodine, a post-therapy scan was to be performed in study TSH95-0101 or if surgery was performed post withdrawal, an histology report was to be provided to confirm the presence of cancer. This requirement was included to help identify false positive scans.

9. The TSH92-0601 study was performed in the US. The TSH95-0101 study involved 11 US sites and 3 European sites (France, Germany and Italy). However, in all sites, it was agreed that the images would be obtained using only the gamma camera.

STUDY OBJECTIVES:

PRIMARY:

1. The primary objective was to demonstrate that Thyrogen provided adequate stimulation for the detection of post-thyroidectomy remnants and cancer by diagnostic ^{131}I imaging.

2. To determine the superior Thyrogen dosing regimen (a 2-dose regimen: Thyrogen 0.9 mg IM q 24 hr. x 2 or a 3-dose regimen: Thyrogen 0.9 mg IM q 72 hrs. x 3).

3. To confirm that patients experience fewer hypothyroid signs and symptoms after Thyrogen than during withdrawal.

Secondary:

1. To determine the diagnostic utility of Thyrogen Tg alone and combined with the diagnostic scan to detect metastatic disease in patients who have been successfully ablated and are Tg antibody negative. Post-hoc, this was changed to detection of remnants and/or cancer.

2. To establish the kinetic profile of the serum Thyrogen Tg response.

3. To compare quality of life after Thyrogen to that on withdrawal.

4. To collect additional Thyrogen safety data.

STUDY DESIGN:

This was a multi-center, open-label, randomized, 2-arm, parallel study.

The main inclusion criteria were:

1. Patients were $\geq$ 18 yrs. old, of either gender, with well-differentiated thyroid cancer, including papillary (including follicular-variant), follicular or Hurthle cell.

2. All patients were s/p near total or total thyroidectomy and by $\geq$ 6 weeks.

3. Patients were $\geq$ 4 mos. post ^{131}I ablation.

4. Recent thyroidectomy patients must have been on THST for $\geq$ 1 month.

5. TSH on THST must have been suppressed to $\leq$ 0.5 mU/L within 7 days prior to Thyrogen administration.

6. Participants in the prior Thyrogen study were eligible for this study.

The main exclusion criteria were:

1. Hypopituitary patients.

2. Patients with undifferentiated thyroid cancer.

3. Patients with non-thyroidal conditions known to decrease radioiodine uptake (e.g. CHF or renal failure) or to cause false positive scans (e.g. inflammatory lung disease,

renal cyst or scrotal hydrocele).

4. Patients taking drugs that may affect thyroid function.

5. Patients who had IV iodinated contrast material within 4 weeks of study enrollment, 2 mos. for intrathecal administration and 3 mos. for cholecystographic agents.

6. Patients who previously received bovine TSH.

PROTOCOL SCHEDULE:

Within 7 days of Thyrogen administration, serum TSH, Tg and Tg antibody was measured. If the serum TSH was ≤ 0.5 mU/L, the patient was randomized to one of two treatment arms: Arm I- Thyrogen 0.9 mg IM q 24h x 2 doses or to Arm II- Thyrogen 0.9 mg IM q 72 h x 3 doses (Thyrogen phase). 24 hrs. after the final Thyrogen dose, patients received a diagnostic activity of 4 mCi (± 10%) (148 Mbq) ^{131}I and underwent diagnostic scanning 48 h later (the whole body scan consisted of composite anterior and posterior views of the head, chest, abdomen, pelvis and thighs,and/or one anterior and one posterior whole body image). The investigator had the option of performing a 72 h whole body scan. Thyroidal ^{131}I uptake measurements were made for all patients. In arm I, serum TSH and Tg levels were measured on the same day as the last Thyrogen injection and on days 1, 2, 3 and 7 after the final injection; in arm II, there was one additional TSH and Tg measurement- on the day of the second Thyrogen injection. Subsequently, the patients remained on THST for 2 additional weeks before being withdrawn for the period required for endogenous TSH levels to rise ≥ 25 mU/L (generally 2 weeks off T3 medication and 4-6 weeks off T4 medications) (Hypothyroid or Withdrawal Phase). When TSH rose to ≥ 25 mU/L, serum TSH and Tg levels were measured and a diagnostic activity of 4 mCi (± 10%) (148 Mbq) ^{131}I was administered; diagnostic whole body scanning was done 48 h later. The investigator again had the option of performing a 72 h scan. The decision to treat the patient was based on the withdrawal scan and the serum Tg measured at the investigational site. However, the Tg levels on all patients were also assayed centrally, at the [redacted]. This central laboratory used a sensitive Tg radioimmunoassay with a lower limit of detection of 0.5 ng/ml. To control for interassay variability, Endocrine Services Laboratory ran all the samples in a singular assay. It was this central Tg level that was used to compute the diagnostic utility of Tg on Thyrogen vs. withdrawal vs. Tg on THST.

An assessment of hypothyroid symptomatology was performed at baseline, after Thyrogen administration and after withdrawal using two validated quality of life instruments-

Billewicz Scale and the SF-36 QOL instrument. The Billewicz Scale is an observer-rated evaluation of 14 signs and symptoms of hypothyroidism. The SF-36 QOL instrument is a patient self-administered scale that measures 8 health concepts: general health †scale ranges from poor to excellent), physical functioning (ranging from very limited in performing all physical activities such as bathing or dressing to being able to perform all types of physical activities) , role-physical (scale ranges from problems to no problems with work or other daily activities as a result of physical health), bodily pain, vitality (scale ranges from tired all the time to full of energy all the time), social functioning (scale ranges from frequent interference with normal social activities due to emotional problems to performs normal social activities), role-emotional (scale ranges from problems to no problems with work or other daily activities as a result of emotional problems), mental health (ranges from depressed all the time to happy all the time) and reported health transition (ranges from believes general health is much better than one year ago to much worse now than one year ago).

EFFICACY ENDPOINTS, METHODOLOGY AND STATISTICAL METHODS USED FOR ANALYSIS:

PRIMARY:

a. Within Patient Equivalence of Diagnostic ^{131}I Imaging with Thyrogen Compared to Withdrawal:

The primary efficacy analysis was a comparison of the uptake classification rating for the 48 h whole body scan obtained on Thyrogen to that on withdrawal for each scan pair in each treatment arm. The following scan classification system was utilized:

Class 0- no uptake in the thyroid bed, no evidence of metastases i.e. negative scan
Class 1- uptake limited to the thyroid bed
Class 2- 2A: solitary focus of uptake in the neck
2B: multiple foci of uptake in the neck
Class 3- 3A: uptake in mediastinum but not lungs
3B: nodular foci of uptake in the lungs
3C: diffuse uptake in the lungs
3D: any combination of 3A, B &/or C.
Class 4- 4A: solitary focus in the skeleton
4B: more than one focus in the skeleton
4C: one or mor foci of uptake in liver
4D: one or more foci of uptake in brain
4E: any combination of 4A orB with 4C orD

Each of three independent reviewers (IR) conducted a blinded (didn't know which scan was the Thyrogen scan and which, the withdrawal scan) side-by-side comparison of each scan pair. Based on the above classification system, each IR assigned

a class to each scan of the pair for each patient. The scan pair went to panel review (consisting of the same 3 IRs brought together to conduct a blinded review of the scan pair) if 2/3 IRs could not agree on the scan class, or an IR had difficulty determining the rating of the scan, or one or more of the reviewers agreed that there was a difference in the number and distribution of the lesions which could alter the clinical management of the patient. If a consensus was not reached by the panel (2/3 IRs could not agree), the scan pair was not included in the efficacy analysis.

The scans of a given pair were rated as "concordant" if both 48 h scans were given the same classification rating by 2/3 IRs and there were no clinically important differences in the number and distribution of the lesions which could potentially alter the clinical management of the patient. The scans were considered "discordant" if a higher classification was assigned to one of the scans in a given pair or there were clinically important differences in the # and distribution of the lesions which could potentially alter the clinical management of the patient. Hence, the scan pair from each patient was categorized as follows:

Thyrogen scan class $\geq$ withdrawal scan
Thyrogen scan < withdrawal scan

These proportions (i.e. proportion of patients in each uptake classification) were presented in terms of both a point estimate and a 95% confidence interval using the Fleiss formula. Also, a two-tailed sign test was used to test whether the discordances significantly favored the Thyrogen scan or the withdrawal scan. The 2-tailed Fisher's exact test was used to determine whether there was a significant difference between the two treatment arms.

b. The mean change in hypothyroid signs/symptoms from baseline to the Thyrogen scan were compared to mean changes from baseline to the withdrawal scan using the Billewicz Scale. The Wilcoxon Signed Rank test was used for within-treatment arm comparisons and the Mann Whitney test for between-treatment arm comparisons.

SECONDARY:

a. The kinetic profile of the Tg response to Thyrogen was evaluated to determine the optimal time for Tg testing on Thyrogen for each dosing regimen. The optimal time was determined by the day on which Thyrogen Tg levels peaked. Treatment arm comparisons were performed using analysis of variance (ANOVA) on the change of Tg from baseline for days 1, 2, 3 and 7 after the final Thyrogen injection.

b. Prospectively, the study was designed to

compare the diagnostic utility of Thyrogen Tg alone and combined with the scan to detect metastatic disease to withdrawal Tg -/+ scan and to Tg on THST using the same Tg cut-offs for Thyrogen Tg, withdrawal Tg and Tg on THST (2, 5 and 10 ng/ml, all of which have been reported in the literature to be indicators of cancer in patients who have been surgically and radioiodine ablated). The reference standard was a post-therapy scan positive for metastatic disease. This analysis was to be conducted in successfully ablated, Tg antibody negative patients, both high risk and low risk patients together and separately- high risk = TNM stage III and IV, and low risk= TNM stage I and II. Successfully ablated was prospectively defined as post a total or near total thyroidectomy (TT or NT), with or without subsequent ^{131}I ablation, and with radioiodine thyroidal uptake < 1% by ROI or by probe on the most recent scan prior to the study.

These diagnostic utility analyses entailed calculation of the following operating characteristics (note: true positive= TP, true negative= TN, false positive= FP, false negative= FN):

-prevalence= # patients who had disease as defined by the reference standard/total # evaluable patients

-sensitivity= probability that the test was positive if disease was truly present: TP/TP + FN

-specificity= probability that the test was negative if disease was truly not present: TN/TN + FP

-positive predictive value= probability the patient truly did have disease given a positive test result: TP/TP + FP

-negative predictive value= probability the patient truly did not have disease given a negative test result: TN/TN + FN

-accuracy= proportion of patients in whom the test correctly detected disease: TP + TN/total # evaluable patients

In addition, the diagnostic utility of Thyrogen Tg to detect metastatic disease was to be compared to Tg on THST using, as the reference standard, a post therapy scan positive for metastatic disease or a WD Tg $\geq$ 10 ng/ml in a successfully ablated patient who was treated (treated, e.g., with ^{131}I or external radiation). (Note: in this latter analysis, a direct comparison to withdrawal cannot be made, because withdrawal, being part of the reference standard, cannot be compared against itself).

Post-hoc (and without notifying FDA), the diagnostic utility analysis of Thyrogen Tg alone and combined with the scan was changed from detection of metastatic disease to detection of remnants and/or cancer (rationale for this change by the sponsor can be found on pages 32-33 of this review). The reference standard now included withdrawal Tg (using cut-offs of 2, 5 and 10 ng/ml) and a withdrawal scan (class $\geq$ 1) in addition to the post therapy scan (class $\geq$ 1). Withdrawal and Thyrogen Tg

levels were no longer compared using the same Tg cut-off for both. Receiver operator curves (ROCs) were used to determine what Thyrogen Tg cut-off levels best correlated i.e. gave the best post-hoc fit, in terms of sensitivity and specificity, to the WD - Tg and WD/post-rx.scan data in this study. These best fit Thyrogen Tg cut-offs were then used to conduct the diagnostic utility analyses. A 95% confidence interval for the sensitivity and specificity was presented using the Fleiss formula. In addition, post-hoc, the definition of a successfully ablated patient was expanded to include patients who had undergone a NT or TT, -/+ subsequent radioiodine ablation, and had < 1% radioiodine uptake in the thyroid bed on the WD study scan by ROI or by probe.

c. The mean thyroidal percent ^{131}I uptake measured during the Thyrogen scan was compared to that measured at the time of the withdrawal scan using the Wilcoxon Signed Rank test within each treatment arm.

d. Patient quality of life during the Thyrogen study phase was compared to the withdrawal study phase using the Standard Form-36 (SF-36) QOL instrument. The Wilcoxon Signed Rank test was used for within-treatment arm comparisons and the Mann Whitney test for between-treatment arm comparisons.

STUDY RESULTS:

PATIENT DISPOSITION:

Patient Population	Arm I	Arm II
Total enrolled and randomized	123	131
Total # Rx'd with Thyrogen	117	112
Total # completed study	116	110

There were 3 reasons why patients were randomized but not treated with Thyrogen:

withdrew from study (Arm I: 4; Arm II: 11)
non-compliant with study (Arm I: 1; Arm II: 6)
did not meet inclusion criteria(Arm I: 1; Arm II: 2)

There were 3 patients (1 in Arm I and 2 in Arm II) who did not complete the study (none were stated to be due to adverse events):

- Arm I: due to disease progression requiring external beam radiation

Arm II: 1 withdrew from the study for prednisone treatment and 1 was non-compliant

Patient Demographics (Note: NT/TT= near total/total thyroidectomy, Hemi= hemithyroidectomy):

Demographic Feature	Arm I (N=117)	Arm II (N= 112)
Gender: Female Male	74 (63%) 43 (37%)	74 (66%) 38 (34%)
Mean Age (years)	44 yrs.	50.5 yrs.
Histology: Papillary Papillary/follicular variant Follicular Hurthle cell	 78 (67%) 19 (16%) 17 (15%) 3 (3%)	 64 (57%) 21 (19%) 22 (20%) 5 (5%)
Status at Enrollment Recent: NT/TT Hemi Follow/up: NT/TT Hemi	 18 (15%) 16 NT/TT 2 Hemi 99 (85%) 89 NT/TT 10 Hemi	 21 (19%) 20 NT/TT 1 Hemi 91 (81%) 80 NT/TT 11 Hemi
TNM Stage: I II III IV	 70 (60%) 22 (19%) 16 (14%) 6 (5%)	 54 (48%) 22 (20%) 22 (20%) 10 (9%)
Mean Time Since Most Recent Surgery(mos.)	44 mos. range: 1-233 mos.	50 mos. range: 1-341 mos.
Mean Time Since Most Recent ^{131}I Rx.(mos.)	41 mos. range: 4-231 mos.	38 mos. range: 2-259 mos.

A distict effort was made to recruit patients with a history of metastatic disease into the study. 120/229 (arm I + arm II) (52%) had a history of metastatic cancer. In addition, 52/229 (23%) patients had evidence of metastatic disease on their most recent scan prior to study enrollment.

All analyses were performed on the intent-to-treat population. The primary efficacy analyses were also performed in the efficacy evaluable patients.

These populations were defined as follows:

Intent-to-Treat population (ITT): consisted of all randomized patients who had at least one efficacy evaluation after receiving at least one Thyrogen injection. Patients with scans that were technically inadequate and, hence, not rated by the IRs were excluded from the ITT analysis.

Efficacy Evaluable (EEP): excluded patients who failed to meet inclusion/exclusion criteria, whose TSH levels were not suppressed to $\leq$ 0.5 mU/L prior to Thyrogen administration, failed to raise TSH to $\geq$ 25 mU/L during the withdrawal phase, failed to meet criteria for ^{131}I administration used for scanning, failed to complete the study or had ineligible scans for review.

PRIMARY EFFICACY:

Table 1: 48 hr. Scan Evaluable ITT Population (IR evaluations):

Category	Arm I	Arm II
Concordance	101/113 (89%)	94/107 (88%)
Discordance	12/113 (11%)	13/107 (12%)
Favoring Thyrogen (Thy scan > WD scan)	3[A]	5[C]
Favoring WD scan (WD scan > Thy scan)	9[B]	8[D]
p value	0.146	0.581

Footnotes A-D: The discordant scan pairs in each treatment arm will be grouped by scan class, patient's ablation status (R= recent thyroidectomy but pre- ^{131}I ablation; f/u= follow-up patient, s/p surgical and radioiodine ablation) and the patient's TNM stage (stages I & II= low risk for mortality, stages III and IV= high risk for mortality):

Arm I: (12 discordant scan pairs= 11%):

A. Thyrogen scan > WD scan (n= 3):

Thyrogen scan	WD scan	Post-rx. scan
3 class 1 (3 f/u), (2 low & 1 high risk)	3 class 0	1 rx'd- class 3A

B. WD scan > Thyrogen scan (n= 9):

Thyrogen scan	WD scan	Post-rx. scan
6 class 0	6 class 1 (6 f/u) (6 low risk)	3 rx'd- 2 class 1, 1 class 2B
1 class 1	1 class 2B (f/u) (high risk)	1 class 2B
1 class 0	1 class 3B (f/u, low risk)	1 class 3B
1 class 4A	1 class 4B (f/u, high risk)	1 class 4B

Arm II: (13 discordant scan pairs= 12%):

C. Thyrogen scan > WD scan (n= 5):

Thyrogen scan	WD scan	Post-rx. scan
3 class 1 (1 R, 2 f/u) (3 low risk)	3 class 0	1 rx'd- class 0
2 class 2A (1 R, 1 f/u) (2 low risk)	1 class 1 1 class 0	1 class 2A 1 not rx'd (WD & Thy Tg<1)

D. WD scan > Thyrogen scan (n= 8):

Thyrogen scan	WD scan	Post-rx. scan
6 class 0	6 class 1 (6 f/u) (5 low, 1 high risk)	5 rx'd- 3 class 1 1 class 3D 1 class 3B
1 class 1	1 class 2B (R, low risk)	1 class 2B
1 class 0	1 class 3A (f/u, high)	1 class 3A

Note: the analysis in the EEP was similar to the ITT. In the EEP, arm I, there were 11 discordant scan pairs (10%), with 2 favoring Thyrogen and 9 favoring WD (p= 0.065, 2-tailed sign test); in arm II, there were 12 discordant scan pairs (12%), with 4 favoring Thyrogen and 8 WD (p= 0.39).

In the following tables (tables 2-6), the concordance and discordance rates (based on the IR evaluations of the 48h scans in the ITT population) for Thyrogen and WD diagnostic scans will be presented for the following patient subgroups:

patients with positive scans (all scans class $\geq$ 1),
patients with scans class= 1,
patients with scans class $\geq$ 2,
patients in whom metastatic disease was confirmed by post-rx. scan, or a + lymph node bx. for cancer post-study or a + chest CT scan post-study,
patients at low risk vs. high risk for mortality (i.e. TNM stages I & II vs. III & IV),
patients recently thyroidectomized but pre ^{131}I ablation vs. follow-up patients- s/p surgical and radioiodine ablation):

Table 2: Scan Equivalence of Uptake Class: Scan Class ≥ 1:

Class ≥ 1	Arm I	Arm II
Concordance	36/48 (75%)	47/60 (78%)
Discordance Thy > WD WD > Thy p value	12/48 (25%) 3 9 0.146	13/60 (22%) 5 8 0.581

Table 3: Scan Equivalence of Uptake Class: Class= 1 and Class ≥2:

Class = 1	Arm I	Arm II	Class ≥ 2	Arm I	Arm II
Concordance	30/39 (77%)	35/44 (80%)	Concordance	6/9 (67%)	12/16 (75%)
Discordance Thy > WD WD > Thy p value	9/39 (23%) 3 6 0.508	9/44 3 6 0.508	Discordance Thy > WD WD > Thy p value	3/9 (33%) 0 3 0.250	4/16 (25%) 2 2 1.00

Table 4: Patients with Confirmed Metastatic Cancer (post-rx. scan class ≥ 2 or positive lymph node bx. for cancer post-study or + chest CT scan post-study):

Category	Arm I	Arm II
Concordance	10/15 (67%)	25/30 (83%)
Discordance Thy > WD WD > Thy p value	5/15 (33%) 1 4 0.375	5/30 (17%) 1 4 0.375

Table 5: Scan Equivalence by TNM Stage (stages I & II= low risk for mortality, stages III and IV= high risk for mortality) (IR evaluations):

Category	Arm I: low risk	Arm I: high risk	Arm II: low risk	Arm II: high risk
Concordance	82/91 (90%)	17/20 (85%)	62/73 (85%)	28/30 (93%)
Discordance Thy > WD WD > Thy p value	9/91 (10%) 2 7 0.180	3/20 (15%) 1 2 1.00	11/73 (15%) 5 6 1.000	2/30 (7%) 0 2 0.500

Table 6: Scan Equivalence by Ablation Status at Enrollment: (R= recent thyroidectomy but pre- ^{131}I ablation, f/u= follow-up patient, s/p surgical and radioiodine ablation) (IR evaluations):

Category	Arm I Recent	Arm I f/u	Arm II Recent	Arm II f/u
Concordance	17/17 (100%)	84/96 (88%)	16/19 (84%)	78/88 (89%)
Discordance Thy > WD WD > Thy p value	0	12/96 (13%) 3 9 0.146	3/19 (16%) 2 1 1.000	10/88 (11%) 3 7 0.344