THE DIAGNOSTIC UTILITY ANALYSES OF 72 h THYROGEN Tg ALONE, THE THYROGEN SCAN ALONE AND THE TWO COMBINED TO DETECT REMNANTS AND CANCER:

As stated above, the ROCs had evaluated the whole range of Tg cut-offs from 1-10 ng/ml; the diagnostic utility analyses used only those cut-offs which provided the optimal post-hoc fit, in terms of sensitivity and specificity, of Thyrogen Tg to the reference standard. Thyrogen Tg cut-offs of 1, 2 and 3 ng/ml had been identified as providing the best post-hoc fit to compare, respectively to the reference standard WD Tg cut-offs of 2, 5 and 10 ng/ml or a positive (class ≥1) WD or post-rx. scan.

Remnants/cancer was defined as present by the reference standard if the WD Tg was ≥ 2, 5 or 10 ng/ml **or** the WD or post-rx. scan was class ≥1. If the Thyrogen Tg was ≥1, 2 or 3 ng/ml, respectively to the reference standard, the Thyrogen Tg was classified as a true positive for the detection of remnants/cancer.

Remnants/cancer was absent by the reference standard if the WD Tg was < 2, 5 or 10 ng/ml **and** the study scan was negative. If the Thyrogen Tg was <1, 2 or 3 ng/ml, respectively to the reference standard, the Thyrogen Tg was classified as a true negative.

Due to the fact that the study had been prospectively designed to compare Thyrogen to WD using the same cut-offs for both, I conducted this 1:1 analysis as well.

The results of both types of comparisons follows (i.e. using the adjusted cut-off vs. the same cut-off):

Tables 3A&3B: Dx. Utility Analysis: 72h Thyrogen Tg-/+ scan(arm I
Tables 4A&4B: " " " " " -/+ " (arm II)
Table 5: " " " : Tg on THST (arm I)
Table 6: " " " : Tg on THST (arm II)

APPEARS THIS WAY
ON ORIGINAL

3.A. Arm I: Dx. Utility of 72 hr. Thyrogen Tg to Detect Remnants/Cancer @ Tg cut-offs between 1-10 ng/ml:

Reference Standard 1: uses same cut-off for Thyrogen Tg as used for WD
Reference Standard 2: uses a WD Tg cut-off of 2 ng/ml to compare to a Thyrogen Tg cut-off of 1 ng/ml, a WD Tg cut-off of 5 ng/ml to a Thyrogen Tg cut-off of 2, and a WD cut-off of 10 to a Thyrogen cut-off of 3 ng/ml.
Both ref. stds. 1 & 2 define a TP as a WD Tg ≥ the given cut-off or a + scan:

	Ref.1	Ref.2	Ref.1	Ref.2	Ref.1	Ref.2	Ref.1	Ref.1
	cut.1	cut.1	cut.2	cut.2	cut.3	cut.3	cut.5	cut10
Elig.	78	78	78	78	78	78	78	78
Prev.	64	57	57	51	54	49	51	49
FP	0	2	0	2	0	0	0	0
FN	17	12	16	12	18	13	17	23
TP	47	45	41	39	36	36	34	26
TN	14	19	21	25	24	29	27	29
Sens.	73%	79%	72%	76%	67%	73%	67%	53%
Spec.	100%	90%	100%	93%	100%	100%	100%	100%
+Pred	100%	96%	100%	95%	100%	100%	100%	100%
-Pred	45%	61%	57%	68%	57%	69%	61%	56%
Accur	78%	82%	79%	82%	77%	83%	78%	71%

3.B. Arm I: Dx. Utility of 72 hr. Thyrogen Tg + Scan to Detect Remnants/Cancer @ Tg cut-offs between 1-10 ng/ml:

	Ref.1	Ref.2	Ref.1	Ref.2	Ref.1	Ref.2	Ref.1	Ref.1
	cut.1	cut.1	cut.2	cut.2	cut.3	cut.3	cut.5	cut10
Elig.	77	77	77	77	77	77	77	77
Prev.	63*	56*	56*	50*	53*	48*	50*	48*
FP	0	3	0	3	0	2	0	0
FN	10	5	7	3	8	3	6	9
TP	53	51	49	47	45	45	44	39
TN	14	18	21	24	24	27	27	29
Sens.	84%	91%	88%	94%	85%	94%	88%	81%
Spec.	100%	86%	100%	89%	100%	93%	100%	100%
+Pred	100%	94%	100%	94%	100%	96%	100%	100%
-Pred	58%	78%	75%	89%	75%	90%	82%	76%
Accur	87%	90%	91%	92%	90%	94%	92%	88%

*= 1 patient less than in table 3A, because the scans were inadequate in 1.

4.A. Arm II: Same analysis as 3.A. above but using data from Arm II: Dx. Utility of 72 hr. Thyrogen Tg to Detect Remnants/Cancer @ Tg cut-offs between 1-10 ng/ml:

Reference Std. 1: uses same cut-off for Thyrogen Tg as for WD Tg
Reference Std. 2: uses the adjusted cut-offs to compare Thyrogen to WD

Both ref. stds. 1 & 2 define a TP as a WD Tg ≥ the given cut-off or a + scan:

	Ref. 1	Ref. 2	Ref. 1	Ref. 2	Ref. 1	Ref. 2	Ref. 1	Ref. 1
	cut. 1	cut. 1	cut. 2	cut. 2	cut. 3	cut. 3	cut. 5	cut10
Elig.	85	85	85	85	85	85	85	85
Prev.	71	68	68	64	65	59	64	59
FP	0	1	0	2	0	3	0	1
FN	7	5	16	14	20	17	24	27
TP	64	63	52	50	45	42	40	32
TN	14	16	17	19	20	23	21	25
Sensit	90%	93%	76%	78%	69%	71%	63%	54%
Specif	100%	94%	100%	91%	100%	89%	100%	96%
+ Pred	100%	98%	100%	96%	100%	94%	100%	97%
- Pred	67%	76%	52%	58%	45%	58%	47%	48%
Accur	93%	93%	82%	81%	77%	76%	73%	68%

4.B. Arm II: Dx. Utility of 72 hr. Thyrogen Tg + Scan to Detect Remnants/Cancer @ Tg cut-offs between 1-10 ng/ml:

	Ref. 1	Ref. 2	Ref. 1	Ref. 2	Ref. 1	Ref. 2	Ref. 1	Ref. 1
	cut. 1	cut. 1	cut. 2	cut. 2	cut. 3	cut. 3	cut. 5	cut10
Elig.	84	84	84	84	84	84	84	84
Prev.	70*	67*	67*	63*	64*	58*	63*	58*
FP	0	3	0	4	0	5	0	1
FN	3	1	5	3	5	2	6	6
TP	67	66	62	60	59	56	57	52
TN	14	14	17	17	20	21	21	25
Sens.	96%	99%	93%	95%	92%	97%	90%	90%
Specif	100%	82%	100%	81%	100%	81%	100%	96%
+ Pred	100%	96%	100%	94%	100%	92%	100%	98%
- Pred	82%	93%	77%	85%	80%	91%	78%	81%
Accur	96%	95%	94%	92%	94%	92%	93%	92%

*= 1 patient less than in table 4A because the scans were inadequate in 1.

5. Arm I: Dx. Utility of Tg on THST to Detect Rem/Ca @ cut-offs bet 1-3 ng/ml:
Reference Standard 1: uses same cut-off for Tg on THST as used for WD
Reference Standard 2: uses a WD cut-off of 2 ng/ml to compare to a Tg on THST cut-off of 1 ng/ml, a WD cut-off of 5 ng/ml to a Tg on THST cut-off of 2, and a WD cut-off of 10 to a Tg on THST cut-off of 3 ng/ml:

	Ref.1	Ref.2	Ref.1	Ref.2	Ref.1	Ref.2
	cut.1	cut.1	cut.2	cut.2	cut.3	cut.3
Elig.	79	79	79	79	79	79
Prev.	67	59	59	53	56	51
FP	1	5	1	1	0	0
FN	29	25	37	31	36	31
TP	37	33	22	21	20	19
TN	12	16	19	26	23	29
Sens.	56%	57%	37%	40%	36%	38%
Spec.	92%	76%	95%	96%	100%	100%
+Pred.	97%	87%	96%	95%	100%	100%
-Pred.	29%	39%	34%	46%	39%	48%
Accur	62%	62%	52%	59%	54%	61%

6. Arm II: Dx. Utility of Tg on THST to Detect Rem/Ca @ cut-offs bet 1-3 ng/ml:
Reference Standard 1: uses same cut-off for Tg on THST as used for WD
Reference Standard 2: uses a WD cut-off of 2 ng/ml to compare to a Tg on THST cut-off of 1 ng/ml, a WD cut-off of 5 ng/ml to a Tg on THST cut-off of 2, and a WD cut-off of 10 to a Tg on THST cut-off of 3 ng/ml:

	Ref. 1	Ref. 2	Ref. 1	Ref. 2	Ref. 1	Ref. 2
	cut. 1	cut. 1	cut. 2	cut. 2	cut. 3	cut. 3
Elig.	80	80	80	80	80	80
Prev.	68*	64*	64*	60*	61*	55*
FP	1	1	0	1	1	2
FN	25	21	33	30	35	30
TP	43	43	31	30	26	25
TN	11	15	16	19	18	23
Sens.	63%	67%	48%	50%	43%	46%
Spec.	92%	94%	100%	95%	95%	92%
+ Pred.	98%	98%	100%	97%	96%	93%
- Pred.	31%	42%	33%	39%	34%	43%
Accur.	68%	73%	59%	61%	55%	60%

*= 5 patients less at each cut-off compared to table 2B because there was no Tg value on THST in these 5 patients and they had remnants/cancer by the reference standard.

The key points regarding tables 3- 6 will be discussed here:

A. Key Points Regarding the Dx. Utility of Thyrogen Using the Adjusted Cut-Offs 1, 2 and 3 ng/ml:

1. 72h Thyrogen Tg alone-

a. In arm I, the operating characteristics (FP and FN rates, sensitivity, specificity, +/- predictive value and accuracy) were similar at the adjusted cut-offs of 1, 2 and 3 ng/ml. At these respective cut-offs, Thyrogen Tg did not detect 12/57 (21%), 12/51 (24%) and 13/49 (27%) patients identified with remnants/cancer by the reference standard.

Sensitivity, negative predictive value and accuracy were higher in arm II than arm I at the 1 ng/ml cut-off due to the lower number of FNS (at 1 ng/ml cut-off: arm II- 5/68 FNS vs. arm I- 12/57 FNS). However, at the 2 and 3 ng/ml cut-offs, the operating characteristics for arm II were comparable to or somewhat inferior compared to arm I.

2. 72h Thyrogen Tg + Scan-

The combination of the Tg and the scan is superior to Tg alone. The addition of the scan markedly improved the operating characteristics in both treatment arms by decreasing the number of false negatives. Thus, the sensitivity improved by ~20%; and negative predictive value, by this degree or even greater. Note the slight increase in false positives which lowered the specificity by 4-10%.

The exception was the 1 ng/ml cut-off in arm II, where the degree of improvement by the addition of the scan was not as great due to the smaller reduction in the number of FNS.

3. Scan alone-

The diagnostic utility of the Thyrogen scan alone was computed using the adjusted Thyrogen Tg cut-offs of 1, 2 and 3 ng/ml. The # of FN Thyrogen scans was, on average, ~2 fold higher (range was 1.2- 5.8 x higher, both treatment arms) than with Thyrogen Tg alone at the corresponding cut-off. Therefore, the diagnostic utility of the Thyrogen scan alone was inferior to Thyrogen Tg alone at all cut-offs for both treatment arms.

4. The sponsor also analyzed the diagnostic utility of Thyrogen Tg -/+ scan in patients at high risk vs. low risk of mortality and concluded that the Thyrogen Tg -/+ scan improves the ability to detect remnants/cancer compared to Tg testing on THST. It should be noted that there were only 13 high risk patients in arm I and only 26 in arm II compared to 64 low risk patients in arm I and 58 low risk in arm II. (Note: the TNM stage for several patients was not reported).

B. 72h Thyrogen Tg vs. Tg on THST (i.e. comparison between tables 3A/4A and tables 5/6):

The sponsor always used the best fit for Thyrogen in all comparisons to WD. Per the ROCs, a 1 ng/ml cut-off provided the best fit for Tg on THST compared to WD. Therefore, I will only comment on the comparison between Thyrogen Tg and Tg on THST at the 1 ng/ml cut-off. Using higher cut-offs artificially magnifies the differences between the two.

The main difference was the lower # of FNS on Thyrogen compared to Tg on THST, resulting in improved sensitivity and negative predictive value for Thyrogen Tg compared to Tg on THST for both treatment arms.

Note: Results were similar when the 48h Thyrogen Tg level was used rather than the 72h value in the above diagnostic utility analyses. Specifically, the main differences were:

For Tg alone:

arm I: with the 48h value, there were 2-3 more FNS at all cut-offs, resulting in ~5% decrease in sensitivity and negative predictive value.

arm II: @ the adjusted 1 ng/ml cut-off, the 48h Thyrogen Tg did not detect 9/67 patients (13%) considered to have remnants/cancer compared to 5/68 (7%)when the 72h level was used. Consequently, with the 48h Tg value, there was less difference in sensitivity and negative predictive value between the adjusted 1 ng/ml cut-off and the higher adjusted cut-offs.

For Tg plus scan:

Results were similar for both the 48 and 72h values in both treatment arms.

THERE ARE SEVERAL PROBLEMS WITH THE DIAGNOSTIC UTILITY ANALYSES:

1. The validity of the reference standard is questionable because the reference standard included the WD diagnostic scan and the reading of the diagnostic scans was questionable. With only a side-by-side comparison of the scans in a given pair for a given patient, the readings were biased in favor of not seeing a difference even if one were present. This is clearly illustrated by the results of the scan readings in the first phase III study, TSH92-0601. 11 scan pairs which had been rated as discordant by the IRS when read individually, became concordant when read side-by-side.

2. How disease was defined as present by the reference standard resulted in classifying several Tg levels on Thyrogen and on THST as FN even though they were equivalent to or similar to WD. This could occur when disease was defined as present by the reference standard based on a class $\geq$ 1 scan rather than the WD Tg level itself.

3. Using lower cut-offs for Thyrogen Tg to compare to

WD Tg rather than using the same cut-off to compare both, underestimates the number of FN Thyrogen Tg levels and overestimates the number of FP Thyrogen Tg levels.

The adjusted cut-offs underestimate the # of FN Thyrogen Tg levels because Thyrogen Tg may test as TP at a lower cut-off but be a FN were a higher cut-off used.

Conversely, using downward adjusted cut-offs to compare Thyrogen Tg to WD, will overestimate the # of FP Thyrogen Tg levels when the reference scan is negative because a lower Tg cut-off increases the likelihood of Thyrogen Tg testing as FP even though the level may be similar to the WD Tg.

Comparing the # of FPS and FNS in tables 3A/B and 4A/B, using the lower adjusted cut-off for Thyrogen (i.e. ref. 2) rather than the same cut-off (i.e. ref. 1) to compare to WD, indicates that when a lower Thyrogen Tg cut-off is used, there is a larger decrease in the # of FNS than there is an increase in FPS. Consequently, using a lower Thyrogen Tg cut-off to compare to WD rather than using the same cut-off to compare both, artificially improves the sensitivity and negative predictive value of Thyrogen. This is most readily apparent in comparing Thyrogen Tg to WD Tg using an adjusted cut-off of 3 ng/ml for Thyrogen vs. using a 10 mg/ml cut-off to compare to WD: in arm I, using an adjusted cut-off of 3 ng/ml, the Thyrogen Tg did not detect 27% (13/49) patients who had remnants/cancer; using a 10 ng/ml cut-off, Thyrogen Tg did not detect 47% (23/49) patients who had remnants/cancer by the reference standard. This decrease in FNS is accompanied by a 20% improvement in sensitivity and a 13% improvement in negative predictive value. The corresponding changes in arm II, were a decrease in FNS from 27 to 17 (46% to 29% decrease in patients not detected by Thyrogen Tg as having remnants/cancer) and an increase in sensitivity and negative predictive value by 17% and 10%, respectively.

<u>Assessment of the Clinical Relevance of the FN 72 h Thyrogen Tg -/+ Scan Values in the Diagnostic Utility Analyses:</u>

To assess the clinical relevance of the Thyrogen Tg -/+ scan values in the diagnostic utility analyses, the FNS will be grouped by reference scan class (post-rx. scan, if available, otherwise, WD scan) to determine the # (%) of successfully ablated, Tg antibody negative patients in whom reference scan demonstrated remnants/cancer was missed by Thyrogen Tg -/+ scan. The Thyrogen Tg -/+ scan FNS will be examined using both the adjusted Tg cut-offs and the same cut-off to compare to the reference standard.

Arm I: FN 72 h Thyrogen Tg -/+ scan patients grouped (gp.) by the **adjusted Thyrogen Tg cut-offs of 1-3 ng/ml** and the reference standard scan class:

[Note: the Ref. Std. for the Thyrogen Tg adj. cut-off of 1 ng/ml is a WD Tg $\geq$

2 ng/ml or a class ≥ 1 WD or post-rx. scan; for adjusted cut-off of 2 ng/ml, it is a WD Tg ≥ 5 or a class ≥ 1 scan; & for adjusted cut-off of 3, it is a WD Tg ≥ 10 or a class ≥ 1 scan)[note: R= recently thyroidectomized but pre- ^{131}I ablation, f/u= follow-up patient: s/p surgical and ^{131}I ablation]:

Adjusted 72h Thyrogen Tg cut-off (ng/ml)	# (%)*FN 72h Thyrogen Tg patients by reference scan class (Note: the denominator for # FN at each ref. scan class is # TP by ref. std.)	# (%)*FN 72h Thyrogen Tg + scan patients by reference scan class (Note: the denominator for # FN at each ref. scan class is # TP by ref. std.)
1 ng/ml	12/57* (21%) class 0 n= 4[A] /27 f/u class 1 n= 8/20 (3R, 5 f/u) class 2 n= 0/4 class ≥3 n= 0/5 inadequate scans n= 0/1	5/56** (9%) class 0 n= 4/27 f/u class 1 n= 1[B]/20 f/u class 2 n= 0/4 class ≥3 n= 0/5
2 ng/ml	12/51* (24%) class 0 n= 2[A]/21 f/u class 1 n= 10/20 (5R, 5 f/u) class 2 n= 0/4 class ≥3 n= 0/5 inadequate scans n= 0/1	3/50** (6%) class 0 n= 2/21 f/u class 1 n= 1[B]/20 f/u class 2 n= 0/4 class ≥3 n= 0/5
3 ng/ml	13/49* (27%) class 0 n= 1[A]/19 f/u class 1 n= 11/20 (5R, 6 f/u) class 2 n= 0/4 class ≥3 n= 0/5 Inadequate scans n= 1/1	3/48** (6%) class 0 n= 1/19 f/u class 1 n= 2[B]/20 f/u class 2 n= 0/4 class ≥3 n= 0/5

* = % patients in whom Thyrogen Tg -/+ scan failed to detect remnants/cancer identified by the reference standard= FN rate.

+= TP by ref. std were distributed by ref. scan class as follows:

	@1	@2	@3
class 0	27	21	19
class 1	20	20	20
class ≥2	9	9	9
inadequate	1	1	1

++= TP by. ref. std. were distributed by reference scan class as follows:

@1	@2	@3
27	21	19
20	20	20
9	9	9
0	0	0

Note there were only 9 patients in arm I who had metastatic disease confirmed by scan.

A: These Tg differences are significant despite a negative scan because Tg is more sensitive than the scan in detecting disease of thyroid origin; the FN rate of the WD dx. scan being reported to range from 10-35% (see refs. below).

B: In these patients, the Thyrogen scan was negative for cancer localized to thyroid bed which was detected by the WD scan.

Additional comments on above table regarding arm I:

Thyrogen Tg missed remnants/cancer identified by the reference standard in 21-27% patients @ the adjusted 1-3 ng/ml cut-offs, respectively.

In patients with class 1 disease detected by the WD reference scan (note: the WD scan detected all patients who had class 1 disease by the post-rx. scan), **Thyrogen Tg was a FN in 8-11 patients (14-22%) @ the adjusted 1-3 ng/ml cut-offs, respectively.**

The combination of the Thyrogen Tg and scan did not detect 6-9% patients with cancer identified by the reference standard @ the adjusted 1-3 ng/ml cut-offs.

The combination of the Thyrogen Tg and scan missed cancer identified by the WD reference scan as localized to the thyroid bed, in 1-2 pts. @ the adjusted 1-3 ng/ml cut-offs, respectively.

Specific FN 72h Thyrogen Tg levels Compared to WD Tg level (Thyrogen Tg/WD Tg) Grouped by Cut-off and Reference Scan Class:

@ 1 ng/ml cut-off: 72h Thyrogen Tg/WD Tg gp. by ref. scan class	@ 2 ng/ml cut-off 72h Thyrogen Tg/WD Tg gp. by ref. scan class	@ 3 ng/ml cut-off 72h Thyrogen Tg/WD Tg gp. by ref. scan class
class 0:	class 0:	class 0:
0.5/11.6 ng/ml	0.5/11.6 ng/ml	0.5/11.6 ng/ml
0.5/ 5.0 "	0.5/ 5.0 "	
0.9/ 2.6 "		
0.8/ 3.5 "		
class 1:	class 1:	class 1:
0.8/ 1.4 ng/ml	0.8/ 1.4 ng/ml	0.8/ 1.4 ng/ml
0.5/ 1.8 "	0.5/ 1.8 "	0.5/ 1.8 "
0.5/ 2.0 "	0.5/ 2.0 "	0.5/ 2.0 "
5 pts. both Thyrogen Tg & WD Tg <1 ng/ml[A, B]	1.1/ 0.9 "	1.1/ 0.9 "
	5 pts. Thy. Tg & WD Tg <1 ng/ml[A, B]	2.1/ 3.7[C] "
	1 pt. Thyr Tg & WD Tg >1-<2 ng/ml[A]	inadequate scan: 2.5/53.9 ng/ml
		5 pts. both Tgs <1[A, B]
		1 pt. both Tgs >1-<2[A]

A= These Thyrogen Tg levels were FNS despite similarity to the corresponding WD Tg because by the ref. std., disease was present (ref. scans were class 1), but these Thyrogen Tg levels were below the given cut-off.
B= Thyrogen scan was FN in 1 of these 5; corresponding WD scan was class 1.
C= Thyrogen scan was FN in this pt.; corresponding WD & post scans were class1

The above tabulation of FN Thyrogen Tg/WD Tg levels demonstrates that there is no correlation between Thyrogen Tg and WD Tg. For any given cut-off, Thyrogen Tg could be equivalent to, less than or greater than the corresponding WD Tg; and if greater than or less than, to highly variable degrees. Therefore, for a given Thyrogen Tg, we cannot predict what the corresponding WD Tg would have been, had the patient been withdrawn.

References to footnote A: Pacini et al, J Nuc Med 28:188, 1987; Van Herle and Brown, Thyroid Disease: Endocrinology, Surgery, Nuclear Medicine and

Radiotherapy: 473-483, 1990; Muller-Gartner et al, Cancer 61:976-981, 1988; Black and Sheppard, Clin Endocrinol 35:519-520, 1991 and Schlumberger, 1997.

Arm I: FN 72h Thyrogen Tg-/+ scan patients using the **same Tg cut-off** to compare to WD, grouped by Tg cut-off and the reference std. scan class:

[Note: the Ref. Std. is a WD Tg ≥ the given cut-off- 1, 2, 3, 5 or 10 ng/ml- or a class ≥ 1 WD or post-rx. scan] [note: R= recently thyroidectomized but pre- ^{131}I ablation, f/u= follow-up patient: s/p surgical and ^{131}I ablation]:

Tg cut-off (ng/ml)	# (%)*FN Thyrogen Tg patients by reference scan class (Note: the denominator for # FN at each ref. scan class is # TP by ref. std.)	# (%)*FN Thyrogen Tg + scan patients by reference scan class (Note: the denominator for # FN at each ref. scan class is # TP by ref. std.)
1 ng/ml	17/64* (27%) class 0 n= 9/34 f/u class 1 n= 8/20 (3R, 5 f/u) class 2 n= 0/4 class ≥3 n= 0/5 inadequate scans n= 0/1	10/63** (16%) class 0 n= 9/34 f/u class 1 n= 1^/20 f/u class 2 n= 0/4 class ≥3 n= 0/5 inadequate scans n= 0/1
2 ng/ml	16/57* (28%) class 0 n= 6/27 f/u class 1 n= 10/20 (5R, 5 f/u) class 2 n= 0/4 class ≥3 n= 0/5 inadequate scans n= 0/1	7/56** (13%) class 0 n= 6/27 f/u class 1 n= 1^/20 f/u class 2 n= 0/4 class ≥3 n= 0/5
3 ng/ml	18/54* (33%) class 0 n= 6/24 f/u class 1 n= 11/20 (5R, 6 f/u) class 2 n= 0/4 class ≥3 n= 0/5 inadequate scans n= 1/1	8/53** (15%) class 0 n= 6/24 f/u class 1 n= 2^/20 f/u class 2 n= 0/4 class ≥3 n= 0/5
5 ng/ml	17/51* (33%) class 0 n= 4/21 f/u class 1 n= 12/20 (6R, 6 f/u) class 2 n= 0/4 class ≥3 n= 0/5 inadequate scans n= 1/1	6/50** (12%) class 0 n= 4/21 f/u class 1 n= 2^/20 f/u class 2 n= 0/4 class ≥3 n= 0/5
10 ng/ml	23/49* (47%) class 0 n= 4/19 f/u class 1 n= 14/20 (7R, 7 f/u) class 2 n= 3/4 f/u class ≥ 3 n= 1/5 f/u inadequate scans n= 1/1	9/48** (19%) class 0 n= 4/19 f/u class 1 n= 3^/20 f/u class 2 n= 2^/4 f/u class ≥3 n= 0/5

* = % patients in whom Thyrogen Tg -/+ scan failed to detect remnants/cancer

identified by the reference standard= FN rate

	+= TP by ref. std. were distributed by reference scan class as follows:					++= TP by ref. std. were distributed by reference scan class as follows:				
	@1	@2	@3	@5	@10	@1	@2	@3	@5	@10
class 0	34	27	24	21	19	34	27	24	21	19
class 1	20	20	20	20	20	20	20	20	20	20
class 2	4	4	4	4	4	4	4	4	4	4
class ≥3	5	5	5	5	5	5	5	5	5	5
inadequate	1	1	1	1	1	0	0	0	0	0

Note there were only 9 patients in arm I who had metastatic disease confirmed by scan.

A= In these patients, the Thyrogen scan was negative for cancer localized to the thyroid bed which was detected by the WD scan.

B= In patients who were successfully ablated and Tg Ab negative, the Thyrogen scan missed **metastatic disease** detected only by the post-rx. scan in 4 pts. In only 1 of these 4 pts., was the Thyrogen Tg > 10 ng/ml. Of the remaining 3 pts., one was not rated a FN when the Thyrogen Tg and scan were combined because the Thyrogen scan was class 1 and the ref. scan std. was class ≥ 1. Had the ref. scan std. been class ≥ 2, then this pt. would have been a FN. The **Thyrogen Tg** levels in these **3 patients** ranged from **5.2- 9.5 ng/ml.** Although the WD scan alsc missed metastatic disease in these 3 pts., the **WD Tgs** were all **> 20 ng/ml.**

Additional comments relative to the above table:

The above table clearly depicts the # of remnants/cancers missed at each cut-off using Thyrogen Tg alone or combined with the scan. Specifically:

Thyrogen Tg missed remnants/cancer identified by the reference standard in 27-47% patients between the 1-10 ng/ml cut-offs, respectively.

In patients with class 1 disease detected by the WD reference scan (note: the WD scan detected all patients who had class 1 disease by the post-rx. scan), **Thyrogen Tg was a FN in 8-14 patients (13-29%) between the cut-offs of 1-10 ng/ml, respectively.**

Thyrogen Tg missed metastatic disease identified by the reference scan, in 4/49 patients (8%) @ the 10 ng/ml cut-off.

The combination of the Thyrogen Tg and scan missed cancer identified by the reference standard in 12-19% patients between the 1-10 ng/ml cut-offs.

The combination of the Thyrogen Tg and the scan missed cancer identified by the WD reference scan as localized to the thyroid bed, in 1-3 patients between the cut-offs of 1-10 ng/ml, respectively.

The combination of the Thyrogen Tg and the scan missed metastatic disease identified by the reference scan, in 2 patients at the 10 ng/ml cut-off.

The above table also clearly points out why the sponsor used receiver operator curves to adjust the Thyrogen Tg cut-off downward relative to WD rather than using the same cut-off to compare both:

- the # of FNS would decrease and, consequently, the sensitivity and negative predictive value of Thyrogen Tg -/+ scan would increase using downward adjusted Tg cut-offs for Thyrogen to compare to the ref. std.
- more importantly, using an adjusted cut-off of 3 ng/ml for Thyrogen Tg to compare to a WD Tg cut-off of 10 ng/ml (rather than using a 10 ng/ml cut-off for both), would capture (i.e. test as true positive) all the patients in whom metastatic disease was detected by the reference standard scan (WD and/or post-rx.).

Note: had the 48h Thyrogen Tg level been used here rather than the 72h value, 2 additional pts. with metastatic disease would have had FN Thyrogen Tgs, both at the 5 ng/ml cut-off.

Specific FN 72h Thyrogen Tg levels Compared to WD Tg (Thy Tg/WD Tg) Grouped by Cut-off and the Reference Scan Class:

n= 17 FN @ 1 ng/ml Thy Tg/WD Tg	n= 16 FN @ 2 ng/ml Thy Tg/WD Tg	n= 18 FN @ 3 ng/ml Thy Tg/WD Tg	n= 17 FN @ 5 ng/ml Thy Tg/WD Tg	n= 23 FN @ 10 ng/ml Thy Tg/WD Tg
class 0:	**class 0:**	**class 0:**	**class 0:**	**class 0:**
0.5/ 5.0ng/ml	0.5/ 5.0ng/ml	0.5/ 5.0ng/ml	0.5/ 5.0ng/ml	0.5/ 5.0ng/ml
0.5/11.6 "	0.5/11.6 "	0.5/11.6 "	0.5/11.6 "	0.5/11.6 "
0.8/ 3.5 "	0.8/ 3.5 "	0.8/ 3.5 "	4.8/21.6 "	4.8/21.6 "
0.9/ 2.6 "	0.9/ 2.6 "	1.8/ 4.6 "	2.8/ 7.6 "	8.4/24.6 "
0.8/ 1.5 "	1.8/ 4.6 "	2.8/ 7.6 "		5.1/12.4 "
0.5/ 1.9 "	1.8/ 2.3 "	2.7/ 3.9 "		
0.9/ 1.5 "				
0.8/ 1.1 "				
0.7/1.9 "				
class 1:	**class 1:**	**class 1:**	**class 1:**	**class 1:**
0.5/ 2.0ng/ml	0.5/ 2.0ng/ml	0.5/ 2.0ng/ml	0.5/ 2.0ng/ml	0.5/ 2.0ng/ml
0.5/ 1.8 "	0.5/ 1.8 "	0.5/ 1.8 "	0.5/ 1.8 "	0.5/ 1.8 "
0.8/ 1.4 "	0.8/ 1.4 "	0.8/ 1.4 "	0.8/ 1.4 "	0.8/ 1.4 "
5 pts. both Thyrogen Tg & WD Tg <1^	1.1/ 0.9 "	1.1/ 0.9 "	1.1/ 0.9 "	1.1/ 0.9 "
	5 pts. both Thy Tg & WD Tg <1^	2.1/ 3.7 "	2.1/ 3.7 "	2.1/ 3.7 "
		5 pts. both Thy Tg & WD Tg <1^	3.6/ 1.7 "	3.6/ 1.7 "
	1 pt. both Thy Tg & WD Tg >1-<2^	1 pt. both Thy Tg & WD Tg >1-<2^	5 pts. both Thy Tg & WD Tg <1^	8.5/22.2 "
			1 pt. both Thy Tg & WD Tg >1-<2^	5 pts. both Thy Tg & WD Tg <1^
				1 pt. both Thy Tg & WD Tg >1-<2^
				class 2:
				5.2/22.2ng/ml
				5.4/25.3
				9.5/27.3
				class >3:
				6.9/11.8ng/ml
		inadeq. scan	**inadeq. scan**	**inadeq. scan**
		2.5/53.9ng/ml	2.5/53.9ng/ml	2.5/53.9ng/ml

APPEARS THIS WAY ON ORIGINAL

A= The 5 patients in whom Thyrogen and WD Tg were both < 1ng/ml and the 1 patient in whom the Tg values were similar (Thyrogen Tg/WD Tg= 1.2/1.7 ng/ml) were FNS because by the ref. scan std., disease was present (ref. scans were class 1), but the Thyrogen Tg levels were below the given cut-off.

The above tabulation of FN Thyrogen Tg/WD Tg levels demonstrates that there is no correlation between Thyrogen Tg and WD Tg. For any given cut-off, Thyrogen Tg could be equivalent to, less than or greater than the corresponding WD Tg;
and if greater than or less than, to highly variable degrees. Therefore, for a given Thyrogen Tg, we cannot predict what the corresponding WD Tg would have been, had the patient been withdrawn.

Arm II: FN 72h Thyrogen Tg -/+ scan patients grouped by the **adjusted Thyrogen Tg cut-offs of 1-3 ng/ml** and reference standard scan class:

[Note: the Ref. Std. for the Thyrogen Tg adj. cut-off of 1 ng/ml is a WD Tg ≥ 2 ng/ml or a class ≥ 1 WD or post-rx. scan; for adjusted cut-off of 2 ng/ml, it is a WD Tg ≥ 5 or a class ≥ 1 scan; & for adjusted cut-off of 3, it is a WD Tg ≥ 10 or a class ≥ 1 scan)[note: R= recently thyroidectomized but pre- ^{131}I ablation, f/u= follow-up patient: s/p surgical and ^{131}I ablation]:

[note: R= recently thyroidectomized but pre- ^{131}I ablation, f/u= follow-up patient: s/p surgical and ^{131}I ablation]:

Adjusted 72h Thyrogen Tg cut-off (ng/ml)	# (%)*FN 72h Thyrogen Tg patients by reference scan class (Note: the denominator for # FN at each ref. scan class is # TP by ref. std.)	# (%)*FN 72h Thyrogen Tg + scan patients by reference scan class (Note: the denominator for # FN at each ref. scan class is # TP by ref. std.)
1 ng/ml	5/68* (7%) class 0 n= 0/15 class 1 n= 5/29 (1R, 4 f/u) class 2 n= 0/5 class ≥3 n= 0/18 inadequate scans n= 0/1	1/67** (1%) class 0 n= 0/15 class 1 n= 1[A]/29 f/u class 2 n= 0/5 class ≥3 n= 0/18 inadequate scans n= 0/1
2 ng/ml	14/64* (22%) class 0 n= 1/11 f/u class 1 n= 13/29 (3R,10 f/u) class 2 n= 0/5 class ≥3 n= 0/18 inadequate scans n= 0/1	3/63** (5%) class 0 n= 1/11 f/u class 1 n= 2[A]/29 f/u class 2 n= 0/5 class ≥3 n= 0/18
3 ng/ml	17/59* (29%) class 0 n= 0/6 class 1 n= 16/29 (5R, 11f/u) class 2 n= 1[B]/5 f/u class ≥3 n= 0/18	2/58** (3%) class 0 n= 0/6 class 1 n= 2[A]/29 f/u class 2 n= 0/5 class ≥3 n= 0/18

* = % patients in whom Thyrogen Tg -/+ scan failed to detect remnants/cancer identified by the reference standard= FN rate.

+= TP by ref. std were distributed by ref. scan class as follows:

	@ 1	@2	@3
class 0	15	11	6
class 1	29	29	29
class 2	5	5	5
class ≥3	18	18	18
inadequate	1	1	1

++= TP by. ref. std. were distributed by reference scan class as follows:

@1	@2	@3
15	11	6
29	29	29
5	5	5
18	18	18
0	0	0

Note: there were 23 patients total in arm II who had metastatic disease confirmed by scan.

A= In these patients, the Thyrogen scan was negative for cancer localized to the thyroid bed which was detected by the WD scan.
B= Thyrogen scan was negative for metastatic disease detected by the post-rx. scan in this patient.

Additional comments:

Thyrogen Tg missed remnants/cancer identified by the reference standard in 7-29% patients @ the adjusted 1-3 ng/ml cut-offs, respectively.

In patients with class 1 disease detected by the reference scan, Thyrogen Tg was a FN in 5 (7%), 13 (20%) and 16 (27%) patients @ the adjusted 1, 2 and 3 ng/ml cut-offs, respectively.

Thyrogen Tg was a FN @ the adjusted 3 ng/ml cut-off in 1 patient in whom metastatic disease was detected only by the post-rx. scan. The Thyrogen Tg was only 2.0 ng/ml. However, this patient was not included as a FN when the Thyrogen Tg and scan were combined because the Thyrogen scan was class 1 and the reference scan standard was class ≥ 1. Had the ref. scan std. been class ≥ 2, then this patient would have been a FN. Note, although the WD scan was also negative for metastatic disease in this pt., the **WD Tg was 16.5ng/ml.**

The combination of the Thyrogen Tg and scan missed cancer identified by the reference standard in 1-5% patients @ the adjusted 1-3 ng/ml cut-offs.

The Thyrogen Tg and scan combination missed cancer identified by the reference scan as localized to the thyroid bed, in 1-2 pts. at the adj. 1-3 ng/ml cut-offs, respectively.