UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS MÉDICAS:
ENDOCRINOLOGIA

IMPACTO DO TNM-8 E DO RASTREIO CORPORAL TOTAL PÓS-DOSE DE RADIOIODO NO MANEJO DE PACIENTES COM CARCINOMA DIFERENCIADO DE TIREOIDE

Dissertação de Mestrado

Carla Fernanda Nava

Porto Alegre, 2019

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS MÉDICAS:
ENDOCRINOLOGIA

IMPACTO DO TNM-8 E DO RASTREIO CORPORAL TOTAL PÓS-DOSE DE RADIOIODO NO MANEJO DE PACIENTES COM CARCINOMA DIFERENCIADO DE TIREOIDE

CARLA FERNANDA NAVA

Dissertação apresentada ao curso de Pós-Graduação em Ciências Médicas: Endocrinologia, UFRGS como requisito parcial para obtenção do grau de Mestre.

Orientador: Professor Dr. José Miguel Dora
Co-orientadora: Professora Dra. Ana Luiza Maia

Porto Alegre, 2019

AGRADECIMENTOS

Ao meu orientador Prof. Dr. José Miguel Dora, a quem considero um excelente professor e orientador, sempre auxiliando de forma bastante paciente e otimista nas incontáveis dúvidas e angústias que surgiram durante a realização deste trabalho.

À minha co-orientadora Profa. Dra. Ana Luiza Maia pela sua contribuição essencial para o aprimoramento do meu trabalho, guiando meu crescimento profissional e pessoal.

A todos os colegas do Grupo de Tireoide. Especialmente as amigas Carla Vaz Ferreira, Ana Patrícia Cristo e Shana Weber pelos ensinamentos, troca de experiência, amizade e palavras de incentivo nos momentos difíceis.

Aos Colegas do Grupo de Tireoide e do Ambulatório de Tireoide Rafael Scheffel e André Zanella pelos ensinamentos e oportunidade de crescimento constante.

A minha Família, principalmente meu Pai Arlindo, minha Mãe Aldemira e meu irmão Paulo, que sempre me incentivaram a alcançar caminhos cada vez mais distantes, sempre foram meu porto seguro e meu apoio em todos os momentos e, acima de tudo, pela família especial que formamos.

Ao meu noivo Alessandro, pelo amor, incentivo, companheirismo, paciência e compreensão pela minha ausência em momentos importantes também a ele.

A todos os mestres que tive na vida, por seu incentivo, exemplo e auxílio na minha formação.

A Deus, pelo dom da vida e do amor ao próximo.

A todas as pessoas e instituições que contribuíram direta ou indiretamente para a conclusão desta dissertação.

Esta dissertação de Mestrado segue o formato proposto pelo Programa de Pós-Graduação em Ciências Médicas: Endocrinologia, Metabolismo e Nutrição, Faculdade de Medicina, Universidade Federal do Rio Grande do Sul, sendo apresentada na forma de dois manuscritos:

- Artigo original referente ao trabalho de pesquisa: Impact of the update TNM staging criteria on prediction of persistent disease in a differentiated thyroid carcinoma cohort; publicado na *Archives of Endocrinology and Metabolism* 2019 Feb;63(1):5-11.

- Artigo original referente ao trabalho de pesquisa: Reappraising the Diagnostic Accuracy of Post-Treatment Whole-Body Scan for Differentiated Thyroid Carcinoma (não publicado)

Além dos artigos já citados, ao longo do período do mestrado foram desenvolvidos os seguintes trabalhos:

- Dynamic Risk Stratification in the Follow-up of Children and Adolescents with Differentiated Thyroid Cancer. Zanella AB, Scheffel RS, **Nava CF**, Golbert L, Laurini de Souza Meyer E, Punales M, Gonçalves I, Dora JM, Maia AL. *Thyroid*. 2018 Oct;28(10):1285-1292.

- Impact of neoadjuvant multikinase inhibitors in response to therapy in patients with locally advanced unresectable thyroid carcinoma: case report and systematic review. **Carla Fernanda Nava**; Rafael Selbach Scheffel; Ana Patrícia Cristo; Carla Vaz Ferreira; Shana Weber; André Borsatto Zanella; Francisco Costa Paixão; Alceu Migliavaca; José Ricardo Guimarães; Marcia Graudenz; José Miguel Dora; Ana Luiza Maia. *Submitted to Thyroid Journal*

LISTA DE ABREVIATURAS E SIGLAS

AJCC/UICC - American Joint Committee on Cancer/Union for International Cancer Control
ATA - American Thyroid Association
CAAE - certificado de apresentação para apreciação ética
CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CDT - carcinoma diferenciado de tireoide
CFT - carcinoma folicular de tireoide
CPT - carcinoma papilar de tireoide
CNPq - Conselho Nacional de Pesquisas
CT - computed tomography
DSS - disease specific survival
DTC - differentiated thyroid carcinoma
DxWBS - diagnostic whole-body scan
FAPERGS - Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul
FIPE - Fundo de Incentivo à Pesquisa
FN - false-negative
FP - false-positive
FTC - follicular thyroid carcinoma
HCPA - Hospital de Clínicas do Porto Alegre
I-131 - iodine-131
MRI - magnetic resonance
NA - not available
NC - not calculable
NCDB - National Cancer Database
neck US - neck ultrasom
NIS - sodium/iodine symporter
NR - non-reactive
PRONEX - Programa de Apoio aos Núcleos de Excelência
PTC - papillary thyroid carcinoma
ptWBS - post-treatment whole-body scan
RA - reactive
RAI - radioactive iodine

RCT - rastreio corporal total

RIT - radioiodoterapia

SD - standard deviation

SEER - Surveillance, Epidemiology and End Results

sPOTg - stimulated post-operatory thyroglobulin

sTg - thyroglobulin under stimulated TSH condition

Tg - thyroglobulin

Tg-T4 - thyroglobulin under TSH suppression

TgAb - anti-thyroglobulin antibody

TNM - 7 - 7th TNM system edition

TNM - 8 - 8th TNM system edition

TNM - tumor, node, metastasis system

TP - true-positive

TSH - thyroid stimulating hormone

US - ultrasound

WBS - whole-body scan

WGS - weighted generalized score

SUMÁRIO

RESUMO 9

ABSTRACT 12

INTRODUÇÃO 14

ANEXO - TABELA 1 16

CONCLUSÃO 18

REFERÊNCIAS BIBLIOGRÁFICAS 19

PARTE I - IMPACT OF THE UPDATE TNM STAGING CRITERIA ON PREDICTION OF PERSISTENT DISEASE IN A DIFFERENTIATED THYROID CARCINOMA COHORT (*publicado na Archives of Endocrinology and Metabolism 2019 Feb;63(1):5-11*) 21

PARTE II - REAPPRAISING THE DIAGNOSTIC ACCURRACY OF POST-TREATMENT WHOLE-BODY SCAN FOR DIFFERENTIATED THYROID CARCINOMA 44

RESUMO

O carcinoma diferenciado de tireoide (CDT) inclui os subtipos papilar (CPT) e folicular (CFT) e é responsável pela maioria das neoplasias malignas tireoidianas. Aproximadamente 85% dos CDT são CPT, atingindo principalmente mulheres, nas faixas etárias entre 40-59 anos. O CDT é geralmente um tumor indolente, usualmente diagnosticado em estágio I e com excelente prognóstico. O manejo do CDT tem mudado nos últimos anos, em janeiro de 2018 foi incorporado à prática clínica a oitava edição do sistema de estadiamento do *American Joint Committee on Cancer - Tumor, node, metastasis* (AJCC/TNM-8) que traz duas principais mudanças com relação à antiga classificação (sétima edição do tumor, linfonodos e metástases - TNM-7): o corte de idade ao diagnóstico passa dos 45 anos para os 55 anos e os critérios de classificação de T3 e T4. Essa atualização objetiva melhorar a predição de doença e sobrevida dos pacientes, separando os pacientes em risco de persistência e/ou recorrência de doença em estágios mais avançados do TNM. Com objetivo de avaliar o impacto da mudança do TNM-7 para o TNM-8 em uma população brasileira de CDT fizemos uma análise comparativa das classificações e dos desfechos relacionados à doença que se encontra no artigo intitulado "*Impact of the update TNM staging criteria on prediction of persistent disease in a differentiated thyroid carcinoma cohort*". Comparamos as classificações do TNM-7 e do TNM-8 em uma coorte de pacientes brasileiros, do sul do país, com carcinoma diferenciado de tireoide. Foram incluídos no trabalho 419 pacientes, quando comparadas às distribuições dos pacientes dentro das classificações notamos que as diferenças entre o TNM-7 e o TNM-8 são estatisticamente significativas e levaram os pacientes, com o TNM-8, a 37% de reclassificações para estágios de menor risco de mortalidade relacionada à doença. Mais da metade (56%) das reclassificações foi atribuída a mudança do corte de idade ao diagnóstico para 55 anos. Com a classificação TNM-8 os pacientes de menor risco foram alocados em estágios mais baixos, sugerindo que o novo sistema é melhor em distribuir os pacientes de acordo com suas categorias de risco. Durante o seguimento mediano de 4,4 anos, os achados referentes à resposta ao tratamento são coerentes com essa

interpretação. Evidenciando assim que o TNM-8 é melhor em estratificar os pacientes com CDT, alocando os pacientes dentro das categorias de risco correspondentes, o que leva ao tratamento mais adequado, menos agressivo, sem expô-los a tratamentos desnecessários e excessivos.

Corroborando com condutas mais conservadoras no manejo dos pacientes com CDT, o consenso Americano no Manejo de Pacientes Adultos com CDT publicado em dezembro de 2015 deixa de indicar tratamento com radioiodo (RIT) aos pacientes de baixo risco. Para todos os outros pacientes que recebem tratamento RIT, a realização do exame de rastreamento corporal total (RCT) após a dose de iodo é mandatória. Diante da incerteza dos benefícios da realização do RCT e dos potenciais risco do exame, fazia-se necessária uma releitura da sua utilidade para os pacientes com CDT que receberam RAI. No artigo intitulado *"Reappraising the Diagnostic Accuracy of Post-Treatment Whole-Body Scan for Differentiated Thyroid Carcinoma"* usamos a mesma coorte de pacientes para avaliar a acurácia diagnóstica do RCT, onde foram avaliados 268 pacientes após sua primeira dose de RAI. Foram revisadas todas as imagens de RCT e os pacientes com diagnóstico documentado de metástases à distância ou captação à distância no RCT foram revisados independentemente por dois especialistas em carcinoma de tireoide. Vinte e nove pacientes possuíam metástases à distância, destes 20 apresentaram captação à distância no RCT (verdadeiro-positivos) e 9 não apresentaram captação à distância (falso-negativos). Vinte e oito pacientes apresentaram captação à distância ao RCT, 9 deles falso-positivos. Estratificando o RCT de acordo com a classificação de risco da *American Thyroid Association* (ATA) notamos que para pacientes de risco baixo e intermediario o exame apresenta baixa sensibilidade no diagnóstico de metástáse à distância. Quando excluídos os pacientes de baixo risco, que atualmente não tem indicação de tratamento com RIT, a performance do exame mostrou-se ainda pior. No entanto, para o grupo de alto risco da ATA, o RCT foi melhor em predizer presença de metástase à distância, com boa sensibilidade, especificidade e valor preditivo positivo (82%, 100% e 100%, respectivamente), com significância estatística. Em resumo, em pacientes com baixa probabilidade pré-teste o exame deve ser reconsiderado, já para os pacientes de alto risco da ATA nos quais a probabilidade pré-teste é alta,

o RCT se mostra uma ferramenta útil para diagnóstico de metástases à distância.

ABSTRACT

Differentiated thyroid carcinoma (DTC) includes the papillary (PTC) and follicular (FTC) subtypes and is responsible for most thyroid malignancies. Approximately 85% of DTC are PTC, afflicting mainly women aged 40-59 years. DTC is usually an indolent tumor, diagnosed in stage I and with excellent prognosis.

The management of DTC has changed in recent years. Released in January 2018, the eighth edition of the *American Joint Committee on Cancer - Tumor, node, metastasis* (AJCC/TNM-8) staging system included two main changes from the previous classification (the seventh edition, TNM-7): the age range at diagnosis of 45 years to 55 years old and the classification criteria for T3 and T4. These updates aim to improve disease prediction and survival of patients, separating patients at risk of persistence and/or recurrence of disease in more advanced stages of TNM. In order to evaluate the impact of the change from TNM-7 to TNM-8 in a Brazilian population of DTC, we performed a comparative analysis of the classifications and outcomes related to the disease found in the article titled *Impact of the update TNM staging criteria on prediction of persistent disease in a differentiated thyroid carcinoma cohort*. We compared the TNM-7 and TNM-8 classifications in a cohort of 419 Brazilian DTC patients from the south of the country. The differences in the distributions of patients between TNM-7 and TNM-8 were statistically significant and led to 37% of patients classified under TNM-8 being moved to stages with a lower risk of mortality related to the disease. More than half (56%) of the reclassifications were attributed to changing the age cut-off to 55 years. With the TNM-8 classification, the lowest risk patients were placed in lower stages, suggesting that the new system is better at distributing patients according to their risk categories. During the median follow-up of 4.4 years, findings regarding treatment response were consistent with this interpretation and showed that TNM-8 is better at stratifying patients with DTC, which leads to more appropriate and less aggressive treatment, thereby not exposing them to unnecessary or excessive treatments.

Corroborating the more conservative management of DTC patients, the U.S. consensus given in the Management of Adult Patients with DTC

published in December 2018 no longer indicates treatment with radioiodine (RAI) for low-risk patients. For all other patients receiving RAI treatment, a whole-body scan (WBS) following the iodine dose is mandatory. Given the uncertainty of the benefits of WBS and the potential risk of the test, a re-reading of its usefulness was necessary for patients with DTC who received RAI. In the article entitled *Reappraising the Diagnostic Accuracy of Post-Treatment Whole-Body Scan for Differentiated Thyroid Carcinoma*, we used the same cohort of patients to assess the diagnostic accuracy of the WBS, where 268 patients were evaluated after their first dose of RAI. All WBS images were reviewed and the patients with documented diagnosis of distant metastases or remote uptake in the WBS were independently reviewed by two specialists in thyroid carcinoma. Twenty-nine patients had distant metastases, of which 20 had remote uptake in the WBS (true-positive) and 9 had no distant uptake (false-negative). Twenty-eight patients presented remote uptake in the WBS, nine of them false-positive. Stratifying the WBS according to the American Thyroid Association (ATA) risk classification, we note that for low-risk and intermediate-risk patients, the test presents low sensitivity in the diagnosis of distant metastases. When low-risk patients, who currently do not have an indication for RAI treatment, were excluded, the performance of the exam was even worse. However, for the high-risk ATA group, WBS was better at predicting presence of distant metastases with statistically significant sensitivity, specificity, and positive predictive value (82%, 100%, and 100%, respectively). In summary, for patients with a low pre-test probability the exam should be reconsidered, whereas for high-risk patients in whom the pre-test probability is high, the WBS is a useful tool for the diagnosis of distant metastases.

INTRODUÇÃO

O carcinoma de tireoide é a neoplasia endócrina mais comum com uma incidência global estimada de 3,1% em 2018 (Bray, et al. 2018). O carcinoma diferenciado de tireoide (CDT), incluindo os subtipos carcinoma papilar (CPT) e carcinoma folicular (CFT), é derivado das células foliculares da glândula tireoide, sendo responsável pela maioria das neoplasias malignas tireoidianas. Aproximadamente 84% dos CDT são CPT, atingindo principalmente mulheres, nas faixas etárias entre 40-59 anos (Shah, et al. 2015; Lim, et al. 2017). O CDT é geralmente um tumor indolente, usualmente diagnosticado em estágio inicial e com excelente prognóstico. Nesse contexto, o manejo do CDT tem mudado nos últimos anos, visto que as condutas preconizadas para tumores pequenos e de menor risco são mais conservadoras. (Haugen, et al. 2016; Ito, et al. 2018; Tuttle and Fagin, et al. 2017).

Recentemente, em janeiro de 2018 foi incorporado à prática clínica a oitava edição do *American Joint Committee on Cancer - Tumor, node, metastasis* (AJCC/TNM8) (Tuttle, et al. 2017). O TNM-8 traz duas principais mudanças com relação à antiga classificação do TNM-7 (sétima edição do tumor, linfonodos e metástases) (Edge, et al. 2010): o corte de idade ao diagnóstico passa dos 45 anos para os 55 anos e a classificação de T3 e T4. Pacientes que antes eram considerados de maior risco de morte relacionada ao CDT se estivessem acima de 45 anos e apresentassem metástases à distância, agora passam a ter o mesmo risco se estiverem acima de 55 anos. Tumores com extensão extratireoidiana mínima passam a ser incluídos nos tumores confinados a tireoide, fazendo com que tumores maiores que 4cm sejam considerados T3a e os maiores que 4cm e de extensão extratireoidiana grosseira sejam considerados T3b. Tumores classificados com T4a passaram a ser os que apresentam invasão grosseira de tecidos subcutâneos, laringe, traqueia, esôfago ou nervo laringeo recorrente e o T4b passa a ser para tumores com invasão grosseira da fáscia pré-vertebral ou vasos cervicais. (Tuttle and Fagin, et al. 2017) (**Tabela 1**). A atualização para o TNM-8 objetiva melhorar a predição de doença e sobrevida dos pacientes, separando os pacientes em risco de persistência e/ou recorrência de doença

em estágios mais avançados do TNM. Com objetivo de avaliar o impacto da mudança do TNM-7 para o TNM-8 em uma população brasileira de CDT fizemos uma análise comparativa das classificações e dos desfechos relacionados à doença representada pelo artigo integrante da parte I desta tese de mestrado.

Outra mudança foi o consenso Americano no Manejo de Pacientes Adultos com CDT em dezembro de 2015, que não mais indica radioiodoterapia (RIT) aos pacientes de baixo risco, em pacientes de intermediário risco ele pode ser considerado se associado a exames de imagem e laboratoriais e para os de alto risco é sempre indicado (Haugen, et al. 2016). Para os pacientes que recebem RIT, a realização do exame de rastreamento corporal total (RCT) após a dose de iodo é mandatória. Diante da incerteza dos benefícios da realização do RCT e dos potenciais risco do exame, faz-se necessária uma releitura da sua utilidade para os pacientes com CDT que receberam RIT. Exames com resultados falso-positivos são uma preocupação, visto que geram estresse psicológico ao paciente e custos financeiros na busca da confirmação da presença de metástase antes não suspeitada, com a realização de exames e consultas adicionais, com novos deslocamentos ao hospital que seriam desnecessários. Como limitação do método, há também que se considerar os exames falso-negativos, onde a metástase está presente e o RCT não foi capaz de diagnosticá-la.

Com o avanço das técnicas diagnósticas, faz-se necessária constante reavaliação dos métodos utilizados no manejo dos pacientes com CDT. Com objetivo de avaliar o desempenho do exame RCT após a primeira dose de RAI na população de CDT, concebemos estudo para avaliar a acurácia diagnóstica do exame no contexto atual, considerando como método padrão-ouro a presença de metástases confirmadas em exames de imagem (tomografia computadorizada, ressonância nuclear magnética, cintilografia óssea e ecografia cervical), biópsias ou exames anatomopatológicos pós-operatórios. Trabalho esse integrante da parte II desta tese de mestrado.

Tabela 1. Comparação entre o TNM-7 e TNM-8.

TNM-7		TNM-8	
Tx	Tamanho do tumor não pôde ser avaliado mas sem extensão extratireoidiana	**Tx**	Tamanho do tumor não pôde ser avaliado
T1	Tumor ≤ 2cm	**T0**	Sem evidência de tumor primário
		T1a	Tumor ≤ 1cm restrito à tireoide
		T1b	Tumor > 1cm ≤ 2cm restrito à tireoide
T2	Tumor > 2 cm ≤ 4cm	**T2**	Tumor > 2cm ≤ 4cm restrito à tireoide
T3	Tumor > 4cm restrito à tireoide ou EET mínima	**T3a**	Tumor > 4cm restrito à tireoide
		T3b	Tumor de qualquer tamanho com EET grosseira, invadindo músculos do pescoço (esterno hioide, esterno tireoide, tireo hioide ou omohioide)
T4a	Tumor de qualquer tamanho com EET grosseira - invadindo tecido subcutâneo, laringe, traquéia, esôfago ou n laringeo recorrente	**T4a**	Tumor de qualquer tamanho com EET grosseira - invadindo tecido subcutâneo, laringe, traquéia, esôfago ou n laringeo recorrente
T4b	Tumor de qualquer tamanho com EET grosseira - invadindo fáscia pré-vertebral, art carótida ou vasos mediastinais	**T4b**	Tumor de qualquer tamanho com EET grosseira - invadindo fáscia pré-vertebral, art carótida ou vasos mediastinais
Nx	LNFs não avaliados na cirurgia	**Nx**	Linfonodos (LNFs) não foram avaliados
N0	Sem evidência de metástases para LNFs	**N0**	Sem evidência de metástases para LNFs
		N0a*	Um ou mais LNFs com citologia ou histologia benigna
		N0b*	Sem evidência clínica ou radiológica de metástases para LNFs
N1a	Metástase para nível VI (pré-traqueal, paratraqueal ou perilaringeal/Delfiano)	**N1a**	Metástase para LNFs níveis VI ou VII (pré-traqueal, paratraqueal ou perilaringeal/Delfiano ou mediastino superior). Doença uni ou

			bilateral
N1b	Metástase para LNFs unilateral, bilateral, contralateral ou mediastino superior	**N1b**	Metástase unilateral, bilateral ou contralateral para LNFs níveis I, II, III, IV ou V ou retrofaríngeanos.
Mx	Metástase à distância desconhecida	**Mx**	Metástase à distância desconhecida
M0	Sem metástase à distância	**M0**	Sem metástase à distância
M1	Presença de metástase à distância	**M1**	Presença de metástase à distância

Estágio	**TNM-7**		**TNM-8**	
	< 45 anos	**≥ 45 anos**	**< 55 anos**	**≥ 55 anos**
I	Qualquer (QQ) T, N, M0	T1 N0 M0	QQ T, N, M0	T1-2 N0/Nx M0
II	QQ T, N, M1	T2 N0 M0	QQ T, N, M1	T1-2 N1 M0 T3a-b QQ N M0
III		T3 N0 M0 T1-3 N1a M0		T4a QQ N M0
IVA		T4a N0 M0 T4a N1a M0 T1-4a N1b M0		T4b QQ N M0
IVB		T4b QQ N M0		QQ T, QQ N, M1
IVC		QQ T, QQ N, M1		

CONCLUSÃO

Os trabalhos acima corroboram dados da literatura, sendo que o primeiro valida a nova classificação do TNM-8 na população brasileira de CDT e mostra que a atualização para a oitava edição do TNM classifica melhor os pacientes quanto ao risco de mortalidade relacionada à doença, fazendo com que a intensidade do tratamento adjuvante e o seguimento sejam mais adequados aos riscos da doença de base. O segundo trabalho levanta questão importante no manejo do CDT sobre exame obrigatório no acompanhamento dos pacientes que receberam RIT. Reavaliando a acurácia diagnóstica do RCT na primeira dose de RIT, confirmamos suspeita da prática clínica. Nos pacientes com risco baixo ou intermediário pela ATA o RCT deve ser reconsiderado e provavelmente contra-indicado. A baixa probabilidade pré-teste da presença de metástases à distância nesta população eleva o número de exames falsos positivos e não acrescenta dados ao seguimento. Já os pacientes de alto risco são beneficiados pelo exame, que pode ser uma ferramenta útil na detecção precoce de metástases à distância. Esses dados trazem grande potencial para mudar o consenso e o manejo do CDT no futuro.

REFERÊNCIAS BIBLIOGRÁFICAS

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A 2018 Global Cancer Statistics: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA CANCER J CLIN 2018;68:394–424.

2. Shah JP 2015 Thyroid Carcinoma: Epidemiology, Histology, and Diagnosis. Clin Adv Hematol Oncol. Apr; 13(4 Suppl 4): 3–6.

3. Lim H, Devesa SS, Sosa JA, Check D, Kitahara CM 2017 Trends in Thyroid Cancer Incidence and Mortality in the United States, 1974-2013. JAMA. Apr 4;317(13):1338-1348.

4. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, Pacini F, Randolph GW, Sawka AM, Schlumberger M, Schuff KG, Sherman SI, Sosa JA, Steward DL, Tuttle RM, Wartofsky L. 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer. Thyroid. 2016; 26:1-133.

5. Ito Y, Miyauchi A, Oda H. Low-risk papillary microcarcinoma of the thyroid: A review of active surveillance trials. Eur J Surg Oncol. 2018 Mar;44(3):307-315.

6. Tuttle RM, Fagin JA, Minkowitz G, Wong RJ, Roman B, Patel S, Untch B, Ganly I, Shaha AR, Shah JP, Pace M, Li D, Bach A, Lin O, Whiting A, Ghossein R, Landa I, Sabra M, Boucai L, Fish S, Morris LGT. Natural History and Tumor Volume Kinetics of Papillary Thyroid Cancers During Active Surveillance. JAMA Otolaryngol Head Neck Surg. 2017 Oct 1;143(10):1015-1020.

7. Tuttle M, Morris LF, Haugen B, Shah J, Sosa JA, Rohren E, Subramaniam RM, Hunt JL, Perrier ND. Amin MB, Edge SB, Greene F, Byrd D, Brookland RK, Washington MK, Gershenwald JE, Compton CC, Hess KR, Sullivan DC, Jessup JM, Brierley J, Gaspar LE, Schilsky RL, Balch CM, Winchester DP,

Asare EA, Madera M, Gress DM, Meyer LR, (eds) Thyroid-Differentiated and Anaplastic Carcinoma (Chapter 73) In: AJCC Cancer Staging Manual. 8th ed. Springer International Publishing. 2017; New York City.

8. Edge S, Byrd DR, Compton CC, Fritz AG, Greene FL, Trotti. AJCC cancer staging manual. 7th ed. Springer. 2010; New York.

PARTE I

IMPACT OF THE UPDATE TNM STAGING CRITERIA ON PREDICTION OF PERSISTENT DISEASE IN A DIFFERENTIATED THYROID CARCINOMA COHORT*

*Publicado na Archives of Endocrinology and Metabolism em fevereiro de 2019 e citado no Editorial da Revista.

Impact of the updated TNM staging criteria on prediction of persistent disease in a Differentiated Thyroid Carcinoma Cohort*

Carla Fernanda Nava M.D.; André B. Zanella, M.D.; Rafael Selbach Scheffel, M.D., Ph.D.; Ana Luiza Maia, M.D., Ph.D.; and Jose Miguel Dora, M.D., Ph.D.

Thyroid Unit, Endocrine Division, Hospital de Clínicas de Porto Alegre, Faculty of Medicine, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

Running title: *Impact of the new DTC TNM*

Key words: Differentiated thyroid carcinoma, TNM Staging, Prognosis.

Correspondence and Reprints:
Jose Miguel Dora, MD, PhD.
Grupo de Tireoide
Hospital de Clínicas de Porto Alegre
Rua Ramiro Barcelos 2350, 90035-003
Porto alegre, RS, Brazil
Phone: 55-51-3359.8733
Fax: 55-51-3332.5188
E-mail: jdora@hcpa.edu.br

Disclosure informations: All authors have nothing to declare.

* *Publicado na Archives of Endocrinology and Metabolism 2019 Feb;63(1):5-11.*

Abstract

Background: The 8th TNM system edition (TNM-8) released in 2018 presents significant changes when compared to the 7th edition (TNM-7). The aim of this study was to assess the impact of changing the TNM staging criteria on the outcomes in a Brazilian cohort of differentiated thyroid carcinoma (DTC).

Methods: DTC patients, attending a tertiary, University-based hospital, were classified by TNM-7 and TNM-8. Prediction of disease outcomes status of the two systems was compared in a retrospective cohort study design.

Results: 419 DTC patients were evaluated, comprised by 82%(345/419) women, with mean age at diagnosis of 46.4±15.6 years, 89%(372/419) papillary thyroid carcinoma, with a median tumor size of 2.3cm(P25-P75, 1.3-3.5). One hundred and sixty patients (38%) had lymph node metastases and 47(11%) distant metastases at diagnosis. Using the TNM-7 criteria, 236(56%) patients were classified as Stage I, 50(12%) as Stage II, 75(18%) as Stage III and 58(14%) as Stage IV. When evaluated by the TNM-8, 339(81%) patients were classified as Stage I, 64(15%) as Stage II, 2(0.5%) as Stage III and 14(3%) as Stage IV. After a median follow-up of 4.4years(P25-P75 2.6-6.6), the rate of incomplete biochemical and/or structural response was 54% vs 92%(P=0.004) and incomplete structural response was 42% vs 86%(P=0.009) for patients classified as stage IV by TNM-7 vs TNM-8, respectively. Only 4(1%) disease-related deaths were recorded.

Conclusions: In our cohort, 37% of DTC patients were down staged with the application of TNM-8 (vs TNM-7). Additionally, TNM-8 seems to better stratify the risk of structural incomplete response at follow-up.

Introduction

Differentiated thyroid carcinoma (DTC), comprising papillary (PTC) and follicular carcinoma (FTC), accounts for the majority (>90%) of all thyroid malignancies (1). In recent years, an increasing incidence of DTC diagnosis has been documented worldwide (2). The incremental use of imaging resources in medical practice, along with the continuous development of more sensitive techniques, led to the pandemic observation of "overdiagnosis" of DTC (2). Most of the increment in DTC diagnosis can be credited to increased detection of pre-clinical indolent tumors restricted to the thyroid gland, that present a more favorable disease profile (3,4).

The current 2015 ATA Management Guidelines for Adult Patients with DTC advise that "AJCC/UICC staging is recommended for all patients with DTC, based on its utility in predicting disease mortality, and its requirement for cancer registries" (5). In response to the change in DTC epidemiology, staging systems are being revised, to ensure a more balanced delivery of therapy for DTC, specially for those at low risk of disease-related morbidity/mortality.

The 8th edition (TNM-8) of the American Joint Committee on Cancer; Tumor, Lymph Nodes, Metastasis (AJCC/TNM) system, was incorporated into the management of DTC in January of 2018 (6). The most significant changes were the age at diagnosis cut-off (from 45 to 55 years) and the redefinition of the T3b and T4a classifications. In the TNM-7, microscopic extrathyroidal extension of the tumor warranted a T3 classification. In the TNM-8, T3 is now comprised of tumors greater than 4 cm and confined to the thyroid gland (T3a) or gross extrathyroidal extension of the strap muscles (T3b). T4a is now defined as gross invasion of the subcutaneous tissue, larynx, trachea, esophagus or recurrent laryngeal nerve and T4b as gross invasion of the prevertebral fascia or major vessels (6,7).

The update in the TNM staging criteria aims to improve prediction of disease specific and overall survival, allocating patients at risk of persistent/recurrent disease for more advanced TNM stages, qualifying diagnostic work up and interventions at follow-up (8).

In Brazil we have scarce data on the clinical impact of updating the TNM criteria. Thus, the objective of this study was to assess how changing

the TNM staging criteria affects the prediction of long term outcomes in a cohort of Brazilian DTC patients in tertiary, University-based hospital.

Subjects and Methods

Patients and study design

The patients included were followed in a retrospective cohort of DTC patients from the Thyroid Outpatient Clinic of the Thyroid Unit, Endocrine Division of Hospital de Clínicas de Porto Alegre (HCPA), a tertiary care, university teaching hospital in southern Brazil. From 2009 to 2015, all consecutive patients with a histological diagnosis of DTC and data available to be classified by TNM-7 and TNM-8 staging systems were included. The study was approved by the ethics committee of the institution (CAAE 68434617.5.0000.5327/GPPG 17-0482).

Treatment protocol and follow-up

Our DTC treatment protocol consists of performing total thyroidectomy and administration of radioactive iodine (RAI) remnant ablation activity according to the ATA risk assessment. Decisions regarding cervical lymph node dissection were made at the discretion of the surgical team from the center where the patients underwent the first surgery. Duration of follow-up was defined as the time between the first surgery and the last medical visit to the clinic.

Our ablation protocol used RAI activities prescribed at the attending physician's discretion. The RAI was administered in a stimulated TSH condition of endogenous hypothyroidism (TSH>30 mUI/L), after withdrawing levothyroxine (at least 3-4 weeks without thyroid hormone). A post-treatment whole body scan (post-treatment WBS) was performed seven to ten days after RAI administration.

In the first evaluation, the following data were recorded for each patient: demographics, tumor characteristics (e.g., histological features [papillary thyroid carcinoma or follicular thyroid carcinoma - including Hurthle cell carcinoma], extension and lymph node involvement) and treatment (e.g., surgery, RAI remnant ablation, and other interventions). Each patient was

classified using the 7th and 8th editions of the TNM/AJCC staging system (I, II, III, or IV) (6,9). N0 status was defined with postoperative neck ultrasound (US) imaging or pathological examination of patients with lymph node resection. Distant metastasis (M1) was considered present when there was a lesion outside the cervical bed on imaging (CT, MRI or scintigraphy) with: histological confirmation or post-treatment WBS uptake and/or elevated Tg. The risk of persistent/recurrent disease was assessed based on the proposed risk stratification system by the American Thyroid Association (ATA) guidelines, with patients classified into three risk groups: low, intermediate and high (5).

The follow-up protocol called for an initial assessment at 3 to 6 months after the initial treatment, which included a physical examination of the neck and measurements of the serum thyroglobulin levels under TSH suppression (Tg-T4) and anti-thyroglobulin antibody (TgAb). In a second evaluation, 6 to 12 months after the initial treatment, serum thyroglobulin (Tg) under stimulated TSH condition in endogenous hypothyroidism (TSH>30 mIU/L) (sTg) and TgAb were measured. Neck ultrasound was performed in this first year of follow-up. At this point, the patient was classified according to disease outcome to initial therapy (see below in the outcomes section). Patients classified as having an excellent response were scheduled for annual visits, during which a physical examination of the neck and measurements of Tg-T4 and TgAb were performed. Patients with indeterminate or incomplete responses were scheduled for the same examination at least twice a year. Additional imaging studies [e.g. x-ray, bone scintigraphy, computed tomography (CT), magnetic resonance (MR)] were performed, as needed, whenever the clinical or laboratory findings raised the suspicion of persistent or recurrent disease.

Outcomes

In the first year of follow-up, the disease outcome of initial therapy was assessed based on the serum Tg-T4 or sTg levels, neck US,and additional imaging exams, whenever indicated.

An excellent response was defined as having no clinical or imaging evidence of tumor (i.e., no imaging evidence of tumor on neck US or CT),

undetectable (<0.2 ng/mL) serum Tg-T4 levels and/or sTg levels <1 ng/mL. Indeterminate response was defined as nonspecific findings on imaging studies, Tg-T4 detectable but < 1 ng/mL, sTg detectable but < 10 ng/mL or TgAb stable or declining in the absence of structural or functional disease. Biochemical incomplete response was defined as negative imaging and Tg-T4 ≥ 1 ng/ml or sTg levels ≥ 10 ng/mL or rising TgAb levels. Structural incomplete response was defined as structural or functional evidence of disease at any Tg-T4 level with or without TgAb. Patients who were diagnosed with persistent disease were evaluated for additional treatment (e.g., surgery, RAI, external-beam radiation or multi-kinase inhibitors), depending on the site involved and disease progression.

All patients with biochemical incomplete and structural incomplete responses were considered to have persistent disease. Recurrence was defined as new biochemical or structural evidence of the disease detected in a patient who had been previously classified as having an excellent response. Disease specific survival (DSS) was considered the time from initial surgery to last follow-up or death related to DTC.

Laboratory analysis

Serum Tg measurements were conducted using immunoradiometric assays (from 2000 to 2002 radioimmunoassay; 2002 to 2010 electrochemiluminescence and 2010 until the present chemiluminescence) that indicated functional sensitivities of at least 0.2 ng/mL. Antithyroglobulin antibodies were measured using the passive agglutination method from 2000 to 2010 and by chemiluminescence from 2010 until the present. After each new technique had been implemented, the necessary procedures for standardization and validation were performed. TSH levels were measured by chemiluminescence assay from 2000 to 2006 (Immulite 2000 SIEMENS), electrochemiluminescence from 2006 to 2010 (Modular E ROCHE), chemiluminescence assay from 2010 to 2014 (Centaur XP SIEMENS) and electrochemiluminescence from 2014 until the present (Cobas E602 ROCHE). These tests were all conducted in the central laboratory of the HCPA.

Statistical analysis

The clinical and laboratory data are reported as the mean ± standard deviation (SD) values or as the median and percentiles 25 and 75 (P25-75) for continuous variables, or as absolute numbers and percentages for categorical variables. Comparative analyses of frequencies were performed using Pearson Chi-Square or Fisher's Exact Test, as appropriate. Agreement for qualitative (categorical) variables was accessed using the *Cohen's Kappa* coefficient.

All tests were two-tailed, and all analyses were performed using the Statistical Package for Social Science Professional software version 20.0 (IBM Corp., Armonk, NY). A two-tailed $P < 0.05$ was considered statistically significant.

Results

Patients

Four hundred and nineteen DTC patients were evaluated. Clinical and oncological characteristics of these patients are described in **Table 1**. The mean age at diagnosis was 46.4 ± 15.6 years, 345 (82%) were women and PTC was diagnosed in 372 (89%) patients. The median tumor size was 2.3 cm (P25-75 1.3-3.5cm). One hundred and sixty patients (38%) had lymph node metastases, and 47 (11%) patients had distant metastases. All patients underwent total thyroidectomy ± lymph node dissection and received RAI remnant ablation. The mean RAI activity was 112.3 ± 38.1 mCi.

TNM-7 vs TNM-8: staging comparison

Using the TNM-7 criteria, 236 (56%) patients were classified as stage I, 50 (12%) as stage II, 75 (18%) as stage III and 58 (14%) as stage IV. When evaluated by the TNM-8, 339 (81%) patients were classified as stage I, 64 (15%) as stage II, 2 (0.5%) as stage III and 14 (3%) as stage IV (**Table 2**).

The distribution of stage frequencies from TNM-7 to TNM-8 were statistically significant ($p \leq 0.0001$), and distributed as follows: 236 (100%) patients who were stage I remained the same; of the 50 previously in stage II, 36 (72%) were reclassified to stage I and only 14 (28%) maintained the classification. Of the 75 that were stage III, 49 (65%) were reclassified to

stage I, 26 (35%) to stage II and none remained in this classification; of the 58 patients in stage IV, 19 (33%) were reclassified to stage I, 24 (41%) to stage II, 2 (3%) to stage III, and only 14 (24%) remained in stage IV. One hundred and fifty-five (37%) patients with DTC were downstaged with the application of TNM-8 (vs TNM-7), 87 (56%) being due to the change in the age at diagnosis cut-off from 45 to 55 years and 68 (44%) to change in T3 definition **(Table 3)**.

TNM-7 vs TNM-8: ATA Risk

Using TNM-7, 163/233 (70%) stage I patients were classified as low risk, 63/233 (27%) as intermediate risk and 7/233 (3%) as high risk. Stage II patients were classified as follows: 31/50 (62%) as low risk, 6/50 (12%) as intermediate risk and 13/50 (26%) as high risk. Of those in stage III, 30/75 (40%) were classified as low risk, 44/75 (59%) as intermediate risk and 1/75 (1%) as high risk. Of those in stage IV, 8/58 (14%) were stratified as low risk, 23/58 (40%) as intermediate risk and 27/58 (47%) as high risk.

Applying the TNM-8, stage I patients were stratified as low risk in 212/335 (63%) cases, as intermediate risk in 113/335 (34%) and as high risk in 10/335 (3%). Those in stage II were classified as low risk in 20/65 (31%), as intermediate risk in 23/65 (35%) and as high risk in 22/65 (34%). All stage III (n=2) and stage IV (n=14) patients were classified as high risk (100%) (**Table 4**). The greater number of reclassifications in ATA low and intermediate risk patients led to a weak level of agreement between TNM-7 and TNM-8 (*Kappa* 0.14 and 0.13, respectively; $p \leq 0.0001$ for both comparisons). In ATA high-risk patients, the agreement between the two classifications was considered moderate (*Kappa* 0.54, $p \leq 0.0001$).

TNM-7 vs TNM-8: disease outcomes

Regarding disease outcomes after initial therapy, we verified an excellent response status in 47/223 (21%), 9/49 (18%), 21/71 (30%) and 6/58 (10%) for patients at stage I, II, III and IV of TNM-7, respectively. While using the TNM-8 at the same point in time, an excellent response status was as follows: 76/322 (24%), 7/63 (11%), 0/2 (0%) and 0/14 (0%) for patients at stage I, II, III and IV, respectively. Indeterminate status was observed in 99/223 (44%), 10/49 (20%), 29/71 (41%) and 18/58 (31%) for patients at

stage I, II, III and IV of TNM-7, respectively; and using the TNM-8 we found 135/322 (42%), 19/63 (30%), 1/2 (50%) and 1/14 (7%) for patients at stage I, II, III and IV, respectively. Biochemical incomplete response was the disease status in 40/223 (18%), 13/49 (27%), 11/71 (15%) and 7/58 (12%) for patients at stage I, II, III and IV of TNM-7, respectively; and while using TNM-8 62/322 (19%), 8/63 (13%), 1/2 (50%) and 0/14 (0%) for patients at stage I, II, III and IV, respectively. Structural incomplete response frequencies were 37/223 (17%), 17/49 (35%), 10/71 (14%) and 27/58 (47%) for patients at stage I, II, III and IV using TNM-7, respectively; and in 49/322 (15%), 29/63 (46%), 0/2 (0%) and 13/14 (93%) for patients at stage I, II, III and IV using TNM-8 **(Table 5)**.

After a median follow-up of 4.4 years (P25-P75 2.6-6.6), the majority of the patients maintained the same response to therapy status or were downgraded to a more favorable status, a finding observed with both TNM staging systems ($p \leq 0.001$). The rate of excellent response in stage I patients was 39% vs. 38% for TNM-7 and TNM-8; in stage II patients it was 26% vs 17%, in stage III patients it was 36% vs 0% and in stage IV patients 14% vs 0%. The rate of indeterminate responses was 40% vs 41% in stage I patients; 36% vs 34% in stage II patients; 45% vs 50% in stage III patients and 32% vs 7% in stage IV patients. The rate of biochemical incomplete response was 13% vs 12% in stage I patients; 14% vs 6% in stage II patients; 3% vs 50% in stage III patients and 12% vs 7% in stage IV patients. The rate of structural incomplete response was 8% vs 8% in stage I patients, 24% vs 42% in stage II patients; 16% vs 0% in stage III patients and 42% vs 86% ($p = 0.009$) in stage IV patients for TNM-7 and TNM-8, respectively **(Table 6, Figure 1)**. The agreement level was considered moderate only for patients with incomplete structural response (*Kappa* 0.51, $p \leq 0.0001$).

Recurrence was observed only on 3/419 (0.7%) patients during follow-up period with a median time of 2.6 years.

It was not possible to access the impact of the reclassification on DSS due to the low mortality rate observed in our cohort (n=4, 1%). The four deaths occurred in TNM-7 stage IV patients (n=4, 100%), and in TNM-8 stage II (n=2, 50%) and stage IV (n=2, 50%) patients.

Discussion

The classification of DTC patients using staging systems is a crucial step in the care of these patients, since it allowed the healthcare team to determine the best therapeutic and follow-up approach. This study examined the impact of updating the TNM staging system from TNM-7 to TNM-8 on the reclassification of DTC patients in a Brazilian cohort, showing that the TNM-8 reclassified 37% of patients to lower stages. We also compared the prognostic information from both systems, and observed that TNM-8 was superior in predicting disease status at follow-up.

In recent years the incidence of DTC has been increasing with a more frequent diagnosis of low-risk and indolent tumors (2). These patients do not benefit from an aggressive approach, but may be exposed to the risks of overtreatment. The impact of updating the TNM-7 staging system to TNM-8 has been explored by some studies (8,10,11). In our study the majority of reclassifications occurred due to age at the diagnosis cut-off change (n=87, 56%). The change in age at the diagnosis cut-off was proposed and validated by Nixon et al in a multicenter study of 9500 patients (12,13). Older patients with more advanced disease tend to have a worse prognosis and this is usually proportional to age at diagnosis. Nixon et al showed that increasing the age cut-off point from 45 to 55 years resulted in a better prediction of outcomes, and prevented low risk patients from being overstaged and, consequently, overtreated. The results of Kim et al (14) corroborate these findings, and showed that a cut-off of 55 years was better at predicting DSS. In our population age cut-off numbers were in accordance with other reports, and seem to better stratify low and high risk patients. Twenty percent of all patients in our study were downstaged because of the change in cut-off for age at diagnosis, which is in line with data from Kim et al (10) that found 27%, and Pontius et al results of 26% (SEER cohort) and 26% (NCDB cohort) (11).

Using TNM-8, we observed that stage I patients are mostly classified as ATA low risk group, stage II patients are evenly distributed between low, intermediate, and high risk and patients in stages III and IV are exclusively at ATA high risk. When we compared these figures with the classification obtained using TNM-7, we noticed that patients in stage I were also mostly in the low-risk group, while patients in stage II were distributed between low and

high risk, those in stage III between low and intermediate risk and those in stage IV distributed in all three groups. Taken together these findings suggest that the TNM-8 serves as a good predictor for relapses and complications related to the disease.

Regarding the prediction of disease outcomes of the TNM-8, initially and after a median of 4.4 years (P25-P75 2.6-6.6) follow-up, we observed that patients with an excellent response were only those initially classified as stage I and II, those with indeterminate response were mainly patients in stage I and II, also including some stage III and IV. The groups of patients with biochemical incomplete and structural incomplete responses were comprised by patients with TNM-8 stage I to IV. It is worth mentioning that 86 to 93% of TNM-8 stage IV patients were on incomplete structural response and none of them had an excellent response. While comparing the two staging systems, TNM-8 classifies patients with a more severe disease spectrum as stages III and IV, identifying groups that benefit from a more aggressive treatment and follow-up approaches.

Our study has some limitations. The fact that only two patients remained in stage III after reclassification to TNM-8, makes it difficult to evaluate disease outcome prediction for this subgroup. Also, all patients included in this study underwent total thyroidectomy and received radioiodine, which limits the extrapolation of our results to patients submitted to partial thyroidectomy and to those who did not receive radioiodine. Notwithstanding, one should note that our data reflect a real clinical practice scenario at a tertiary care center, where all patients have access to standardized treatment protocols. Additionally, our data comprise a considerable number of patients, and it is one of the first studies providing Brazilian information on the subject.

In summary, 37% of DTC patients were downstaged with the application of TNM-8 (vs TNM-7). The TNM-8 incorporates a large number of patients to lower stages, retaining the favorable prediction performance for stages I and II. It also selects more advanced DTC disease to stages III and IV. These findings confirm that the application of TNM-8 to our population is valid, and superior to TNM-7, regarding prediction of poor DTC outcomes. The use of TNM-8 has the potential to benefit a large number of patients,

helping deliver the appropriate intensity of care according to disease prognosis.

Table 1. Characteristics of 419 patients with differentiated thyroid carcinoma.

Age at diagnosis (years)	46.4 ± 15.6
Female sex – n (%)	345 (82.3)
Histology – n (%)	
Papillary	372 (88.8)
Follicular	47 (11.2)
Tumor size (cm)	2.3 (1.3-3.5)
Lymph node metastasis (N1) – n (%)	160 (38.2)
Distant metastasis – n (%)	47 (11.2)
RAI therapy – n (%)	419 (100)
RAI activity (mCi)	112.3 ± 38.1
Follow-up (years)	4.4 (2.6-6.6)

Data are expressed as the mean ± SD, median (percentiles 25-75) or frequencies.

RAI, Radioactive iodine.

Table 2. Modifications of differentiated thyroid carcinoma staging from the TNM-7 to the TNM-8 system.

DTC Staging		TNM-8			
		I (n=339)	II (n=64)	III (n=2)	IV (n=14)
TNM-7	I (n=236)	236 (100%)	0 (0%)	0 (0%)	0 (0%)
	II (n=50)	36 (72%)	14 (28%)	0 (0%)	0 (0%)
	III (n=75)	49 (65%)	26 (35%)	0 (0%)	0 (0%)
	IV (n=58)	19 (33%)	24 (41%)	2 (3%)	14 (24%)

DTC, differentiated thyroid carcinoma.

Chi-square of $p \leq 0.0001$ for TNM-7 vs TNM-8.

Table 3. Comparison of Tumor classification of TNM-7 and TNM-8 systems.

		TNM-8							
		T1a	T1b	T2	T3a	T3b	T4a	T4b	**Total**
TNM-7	T1	**69**	**85**	-	-	-	-	-	**154**
	T2	-	-	**111**	-	-	-	-	**111**
	T3	6	24	37	**63**	**2**	-	-	**132**
	T4a	-	-	-	-	-	**13**	-	**13**
	T4b	-	-	-	-	-	-	**4**	**4**
	Total	**75**	**109**	**148**	**63**	**2**	**13**	**4**	**414**

Table 4. ATA risk according to differentiated thyroid carcinoma staging using the TNM-7 and TNM-8 systems.

ATA risk	TNM-7 Stage n (%)					TNM-8 Stage n (%)			
	I	II	III	IV		I	II	III	IV
Low	163 (70)	31 (62)	30 (40)	8 (14)		212 (63)	20 (31)	0 (0)	0 (0)
Intermediate	63 (27)	6 (12)	44 (59)	23 (40)		113 (34)	23 (35)	0 (0)	0 (0)
High	7 (3)	13 (26)	1 (1)	27 (47)		10 (3)	22 (34)	2 (100)	14 (100)

Kappa for TNM-7 vs TNM-8 according to ATA risk: 0.14 for low risk ($p \leq 0.0001$); 0.13 for intermediate risk ($p \leq 0.0001$) and 0.54 for high risk ($p \leq 0.0001$).

Table 5. Disease outcomes after initial treatment according to differentiated thyroid carcinoma staging using the TNM-7 and TNM-8 systems.

Treatment response	TNM-7 Stage n (%)					TNM-8 Stage n (%)			
	I	II	III	IV		I	II	III	IV
Structural incomplete	37 17)	17 (35)	10 (14)	27 (47)		49 (15)	29 (46)	0 (0)	13 (93)
Biochemical incomplete	40 (18)	13 (27)	11 (15)	7 (12)		62 (19)	8 (13)	1 (50)	0 (0)
Indeterminate	99 (44)	10 (20)	29 (41)	18 (31)		135 (42)	19 (30)	1 (50)	1 (7)
Excellent	47 (21)	9 (18)	21 (30)	6 (10)		76 (24)	7 (11)	0 (0)	0 (0)

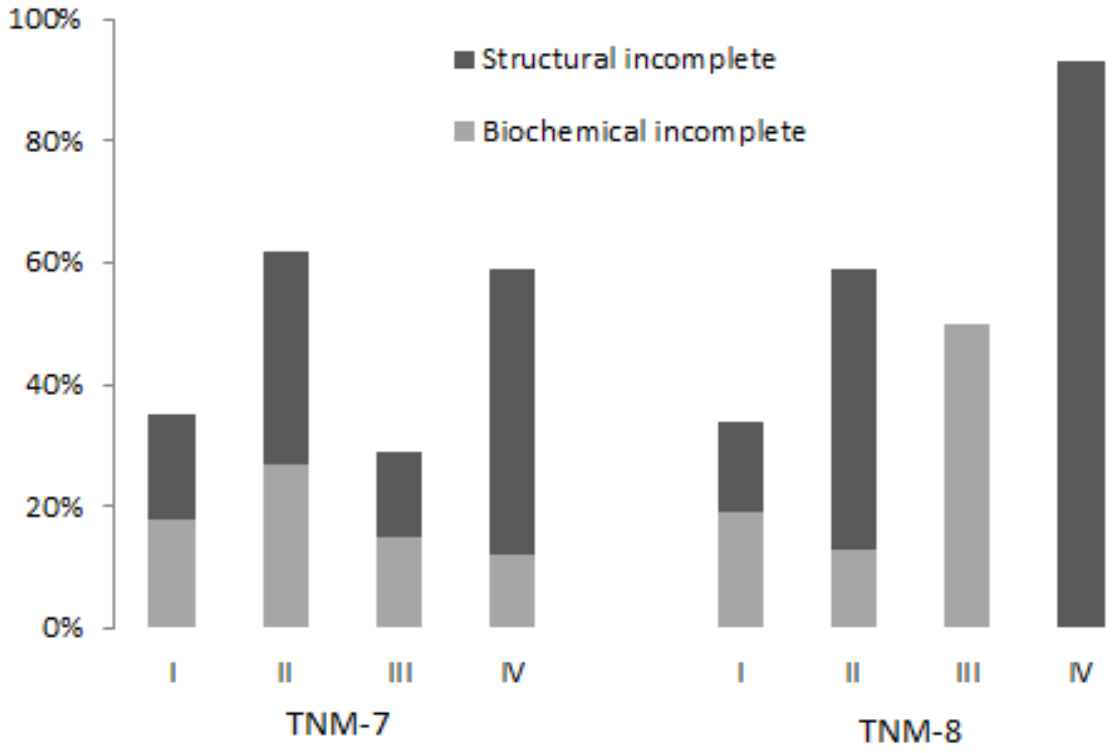
Kappa for TNM-7 vs TNM-8 according to initial treatment response: 0.13 for excellent response ($p = 0.006$); 0.19 for indeterminate response ($p \leq 0.0001$); 0.07 for biochemical incomplete response ($p = 0.26$) and 0.56 for structural incomplete response ($p \leq 0.0001$).

Table 6. Disease outcomes at long-term follow-up according to differentiated thyroid carcinoma staging using the TNM-7 and TNM-8 systems.

Treatment response	TNM-7 Stage n (%)					TNM-8 Stage n (%)			
	I	II	III	IV		I	II	III	IV
Structural incomplete	18 (8)	12 (24)	12 (16)	24 (42)		27 (8)	27 (42)	0 (0)	12 (86)
Biochemical incomplete	29 (13)	7 (14)	2 (3)	7 (12)		39 (12)	4 (6)	1 (50)	1 (7)
Indeterminate	87 (40)	18 (36)	34 (45)	18 (32)		133 (41)	22 (34)	1 (50)	1 (7)
Excellent	86 (39)	13 (26)	27 (36)	8 (14)		123 (38)	11 (17)	0 (0)	0 (0)

Kappa for TNM-7 vs TNM-8 according to treatment response at long-term follow-up: 0.13 for excellent response ($p \leq 0.001$); 0.13 for indeterminate response ($p \leq 0.001$); 0.26 for biochemical incomplete response ($p = 0.002$) and 0.51 for structural incomplete response ($p \leq 0.0001$).

Figure 1. Differentiated thyroid carcinoma biochemical incomplete and structural incomplete responses to therapy at long term follow-up, according to TNM-7 and TNM-8.



Acknowledgments: This work was a possible due to grants from CNPq, CAPES, FIPE and PRONEX / FAPERGS. We wish to thank the anonymous reviewers that contributed to improve our work, to the surgeons of our Hospital (Dr. Alceu Migliavacca, Dr. José Ricardo Guimarães and Dr. Diego Mossmann) for surgical management of our patients and Luciana Scotti and Rosana Scalco for helping with laboratory tests.

Author Disclosure Statement: C.F.N., R.S.S., A.B.Z, A.L.M. and J.M.D. have nothing to declare.

Corresponding Author

Jose Miguel Dora, M.D., Ph.D.

Thyroid Unit, Endocrine Division, Hospital de Clínicas de Porto Alegre

Rua Ramiro Barcelos, 2350. CEP 90035 –003

Porto Alegre, RS, Brazil

Phone: 55-51-3359-7858/ Fax: 55-51-3332-5188

E-mail: jdora@hcpa.edu.br

References

1. Sherman S. Thyroid Carcinoma. The Lancet. 2003; 361:501-511.

2. Salvatore Vaccarella, Silvia Franceschi, Freddie Bray, Christopher P. Wild, Martyn Plummer and Luigino Dal Maso. Worldwide Thyroid-Cancer Epidemic? The Increasing Impact of Overdiagnosis. NEJM. 2016; (375) 7: 614-617.

3. Durante C, Montesano T, Torlontano M, Attard M, Monzani F, Tumino S, Costante G, Meringolo D, Bruno R, Trulli F, Massa M, Maniglia A, D'Apollo R, Giacomelli L, Ronga G, Filetti S, Group PTCS. Papillary thyroid cancer: time course of recurrences during postsurgery surveillance. J Clin Endocrinol Metab. 2013; 98: 636–642.

4. Scheffel RS, Zanella AB, Antunes D, Dora JM, Maia AL. Low Recurrence Rates in a Cohort of Differentiated Thyroid Carcinoma Patients: A Referral Center Experience. Thyroid. 2015; 25: 883-889.

5. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, Pacini F, Randolph GW, Sawka AM, Schlumberger M, Schuff KG, Sherman SI, Sosa JA, Steward DL, Tuttle RM, Wartofsky L. 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer. Thyroid. 2016; 26:1‑133.

6. Tuttle M, Morris LF, Haugen B, Shah J, Sosa JA, Rohren E, Subramaniam RM, Hunt JL, Perrier ND. Amin MB, Edge SB, Greene F, Byrd D, Brookland RK, Washington MK, Gershenwald JE, Compton CC, Hess KR, Sullivan DC, Jessup JM, Brierley J, Gaspar LE, Schilsky RL, Balch CM, Winchester DP, Asare EA, Madera M, Gress DM, Meyer LR, (eds) Thyroid‑Differentiated and Anaplastic Carcinoma (Chapter 73) In: AJCC Cancer Staging Manual. 8th ed. Springer International Publishing. 2017; New York City.

7. Tuttle RM, Haugen B, Perrier ND. The Updated AJCC/TNM Staging System for Differentiated and Anaplastic Thyroid Cancer (8th edition): What changed and why? Thyroid. 2017; Jun;27(6):751-756.

8. Tae Hyuk Kim, Young Nam Kim, Hye In Kim, So Young Park, Jun-Ho Choe, Jung-Han Kim, Jee Soo Kim, Young Lyun Oh, Soo Yeon Hahn, Jung Hee Shin, Kyunga Kim, Jong Gill Jeong, Sun Wook Kim,, Jae Hoon Chung.

Prognostic value of the eighth edition AJCC TNM classification for differentiated thyroid carcinoma. Oral Oncology. 2017; 71: 81–86.

9. Edge S, Byrd DR, Compton CC, Fritz AG, Greene FL, Trotti. AJCC cancer staging manual. 7th ed. Springer. 2010; New York.

10. Mijin Kim, Won Gu Kim, Hye-Seon Oh, Suyeon Park, Hyemi Kwon, Dong Eun Song, Tae Yong Kim, Young Kee Shong, Won Bae Kim, Tae-Yon Sung, and Min Ji Jeon. Comparison of the 7th and 8th Editions of the AJCC/UICC TNM Staging System for Differentiated Thyroid Cancer. Thyroid. 2017; Sep;27(9):1149-1155.

11. Lauren N. Pontius, B.A.1, Taofik O. Oyekunle, M.S.5, Samantha M. Thomas, M.S.4,5, Michael T. Stang, M.D.2,5, Randall P. Scheri, M.D.2,5, Sanziana A. Roman, M.D.2,5, Julie A. Sosa, M.D., M.A..Projecting Survival in Papillary Thyroid Cancer: A Comparison of the 7th and 8th Editions of the AJCC/UICC Staging Systems in Two Contemporary National Patient Cohorts. Thyroid. 2017; Nov;27(11):1408-1416.

12. Nixon, I.J., Kuk, D., Wreesmann, V. Morris, L., Palmer F. I., Ganly I., Patel S.G., Singh B., Tuttle R.M., Shaha A.R., Gönen M. and Shah J.P.. Defining a valid age cut-off in staging of well-differentiated thyroid cancer. Annals of Surgical Oncology. 2016; 23, 410- 415.

13. Nixon IJ, Wang LY, Migliacci JC, Eskander A, Campbell MJ, Aniss A, Morris L, Vaisman F, Corbo R, Momesso D, Vaisman M, Carvalho A, Learoyd D, Leslie WD, Nason RW, Kuk D, Wreesmann V, Morris L, Palmer FL, Ganly I, Patel SG, Singh B, Tuttle RM, Shaha AR, Gonen M, Pathak KA, Shen WT, Sywak M, Kowalski L, Freeman J, Perrier N, Shah JP. An International multi-institutional validation of age 55 years as a cut-off for risk stratification in the AJCC/UICC staging system for well-differentiated thyroid cancer. Thyroid. 2016; 26:373-380.

14. Kim M, Kim YN, Kim WG, Park S, Kwon H, Jeon MJ, Ahn HS, Jung SH, Kim SW, Kim WB, Chung JH, Shong YK, Kim TH, Kim TY. Optimal cut-off age in the TNM Staging system of differentiated thyroid cancer: is 55 years better than 45 years? Clin Endocrinol (Oxf). 2016.

PARTE II

REAPPRAISING THE DIAGNOSTIC ACCURRACY OF POST-TREATMENT WHOLE-BODY SCAN FOR DIFFERENTIATED THYROID CARCINOMA

Reappraising the Diagnostic Accuracy of Post-Treatment Whole-Body Scan for Differentiated Thyroid Carcinoma

Carla Fernanda Nava M.D.[1]; Rafael Selbach Scheffel, M.D., Ph.D.[1]; André B. Zanella, M.D Ph.D.[1]; Flavio Zelmanovitz, M.D. Ph.D.[2], Ana Luiza Maia, M.D., Ph.D.[1]; and Jose Miguel Dora, M.D., Ph.D.[1]

1. Thyroid Unit, Endocrine Division, Hospital de Clínicas de Porto Alegre, Faculty of Medicine, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil
2. Nuclear Medicine, Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil

Running title: *When to perform ptWBS for DTC*

Key words: Differentiated thyroid carcinoma, Whole-Body Scan, Radioiodine, Prognosis.

Word Count: Text: 3363; Abstract: 242; Tables: 3; Figure: 1.

Type of article: Original article

Correspondence and Reprints:

Jose Miguel Dora, MD, PhD.

Unidade de Tireoide

Hospital de Clínicas de Porto Alegre

Rua Ramiro Barcelos 2350, 90035-003

Porto alegre, RS, Brazil

Phone: 55-51-3359.7858

Fax: 55-51-3332.5188

E-mail: jdora@hcpa.edu.br

Disclosure informations: All authors have nothing to declare.

Abstract

Background: Initial treatment for differentiated thyroid carcinoma (DTC) consists of surgery and for a subgroup of patients the administration of radioactive iodine (RAI). Within the context of patients receiving RAI, post-treatment whole-body scan (ptWBS) is considered a mandatory exam, but data on its contemporaneous diagnostic accuracy are scarce.

Objective: Evaluate the diagnostic value of ptWBS in a cohort of DTC patients.

Methods: Two hundred and sixty-eight DTC patients, who underwent total thyroidectomy, received adjuvant RAI and had ptWBS between 2009-2015 were included. Patients with distant uptake on ptWBS or documented distant metastases were independently reviewed by two thyroid cancer specialists.

Results: The mean age at diagnosis was 46±16 years, 220/268 (82%) were women, papillary thyroid carcinoma (PTC) was diagnosed in 234/268 (87%) patients. Median tumor size was 2.7cm (P25-75 1.3-3.5cm), 107/268 (40%) had lymph node metastases and 29/268 (11%) had distant metastases. Twenty-eight patients (10%) had distant uptake, 9 of them (32%) false-positive. Also, 9 false-negative ptWBS were identified. Stratifying ptWBS according to the American Thyroid Association (ATA) risk score, the sensitivity and specificity of the exam for the ATA low-intermediate risk were 17% and 96%, respectively. However, for ATA high-risk group, the performance of ptWBS was better, with a sensitivity of 82% and specificity of 100% ($p \leq 0.0001$).

Conclusions: ptWBS seems to be useful for a subgroup of ATA high-risk DTC patients. The poor overall performance of ptWBS suggests that we should reconsider its routine use for those at ATA low-intermediate risk.

Introduction

Differentiated thyroid carcinoma (DTC), comprising papillary (PTC) and follicular carcinoma (FTC), accounts for the majority of all thyroid malignancies and had an estimated global incidence of 3.1% in 2018 (1). Ten-year survival is high, exceeding 90%. Notwithstanding, 5-30% of patients will present persistent disease or recurrence and up to 8% will develop distant metastases during follow-up (2-4).

The current 2015 American Thyroid Association (ATA) Management Guidelines for Adult Patients with DTC (5) recommends initial treatment with surgery and a classification based on surgical and anatomopathological findings guided by the anatomical extent of disease classification (TNM, American Joint Committee on Cancer and Union for International Cancer Control (AJCC / UICC) (6). According to the risk classifications, the adjuvant therapies are defined as well as the frequency and intensity of follow-up.

Adjuvant radioactive iodine (RAI) is used as DTC treatment based on the fact that organification, trapping and storage of iodine is maintained in functioning thyroid tissues (7). DTC patients who might benefit from RAI therapy have tumors that concentrate iodine, and thus have positive images on post-treatment whole-body scan (ptWBS) (8,5).

ptWBS is considered a mandatory test within the context of patients receiving RAI. The goal of ptWBS is to: (i) document iodine avidity for structural disease, and (ii) refine disease staging by identifying patients with previously unidentified distant metastases (5). Previous studies have suggested that about 8% of DTC cases are reclassified through ptWBS examination as having distant metastases (9) and that ptWBS is more sensitive in detecting new lesions compared to diagnostic whole-body scan (DxWBS) performed before RAI administration (10). Notwithstanding, false-positive ptWBS are well known, which limits the diagnostic reliability of the test for management of DTC patients (11-17). False-positive ptWBS can occur due to functional expression of sodium/iodine symporter (NIS) in non-thyroid tissues, retention of RAI in body fluids, presence of benign cysts, infections and inflammatory processes (18,19).

Another point to be considered is that current DTC risk assessment strategies incorporate clinical and pathological criteria, along with

ultrasensitive thyroglobulin (Tg), a marker with high sensitivity and specificity for detection of persistent DTC (20,21). In this context, there is considerable uncertainty regarding the contemporaneous diagnostic usefulness of ptWBS as an adjunct tool for DTC staging.

Thus, the aim of this study was to evaluate the diagnostic value of ptWBS in a cohort of DTC patients.

Material and Methods

Patients and study design

The patients included were followed in a cohort of DTC patients from the Thyroid Outpatient Clinic of the Thyroid Unit, Endocrine Division of Hospital de Clínicas de Porto Alegre (HCPA), a tertiary care, university teaching hospital in southern Brazil. From 2009 to 2015, all consecutive patients submitted to total thyroidectomy, with a histological diagnosis of DTC, that received the first dose of RAI activity and had a ptWBS image available were included. The study was approved by the ethics committee of the institution (CAAE 68434617.5.0000.5327/GPPG 17-0482).

Treatment protocol and follow-up

Our DTC treatment protocol consists of performing thyroidectomy (total or subtotal), followed or not by administration of RAI activity as indicated and levothyroxine therapy (5,22,23). Decisions regarding cervical lymph node dissection were taken at the discretion of the surgical team at the center where the patients underwent the first surgery. Our institutional protocol does not contemplate prophylactic lymph node dissection. Duration of follow-up was defined as the time between the first surgery and the last medical visit to the clinic.

Our iodine administration protocol used RAI activities prescribed at the attending physician's discretion. RAI was administered in a stimulated thyrotropin (TSH) condition of endogenous hypothyroidism (TSH >30 mUI/L), after withdrawing levothyroxine (at least 3-4 weeks without thyroid hormone).

ptWBS was performed 7 to 10 days after RAI administration. At our service, DxWBS is not performed as a routine.

Stimulated postoperative thyroglobulin (sPOTg) was measured post-thyroidectomy and before RAI, under stimulated conditions and was considered appropriate if TSH was above 30 mIU/L (endogenous hypothyroidism). Serum levels of anti-thyroglobulin antibody (TgAb) were accessed in the same blood sample in which sPOTg was measured and patients with positive results were excluded from all analyses of sPOTg.

In the first evaluation, the following data were recorded for each patient: demographics, tumor characteristics (e.g., histological features [papillary thyroid carcinoma or follicular thyroid carcinoma - including Hurthle cell carcinoma], extension and lymph node involvement) and treatment (e.g., surgery, RAI, and other interventions). Each patient was classified using the 8th edition of the TNM/AJCC staging system (I, II, III, or IV) (6,24). N0 status was defined with postoperative neck ultrasound (US) imaging or pathological examination of patients with lymph node resection. Distant metastases (M1) were considered present when there was a lesion outside the cervical bed on imaging (computed tomography [CT], magnetic resonance [MRI] or bone scintigraphy) with histological confirmation or ptWBS uptake and/or elevated thyroglobulin (Tg). The risk of persistent/recurrent disease was assessed based on the proposed risk stratification system by the ATA 2015 guidelines, with patients classified into three risk groups: low, intermediate and high (5).

The follow-up protocol called for an initial assessment 3 to 6 months after the initial treatment, which included a physical examination of the neck and measurements of the serum Tg levels under TSH suppression (Tg-T4) and TgAb. In a second evaluation, 6 to 12 months after the initial treatment, Tg under stimulated TSH condition in endogenous hypothyroidism (TSH>30 mIU/L) (sTg) and TgAb were measured. Neck ultrasound was performed in this first year of follow-up. At this point, the patient was classified according to disease outcome to initial therapy (see below in the outcomes section). Patients classified as having an excellent response were scheduled for annual visits, during which a physical examination of the neck and measurements of Tg-T4 and TgAb were performed. Patients with indeterminate or incomplete responses were scheduled for the same examination at least twice a year.

Additional imaging studies (e.g. x-ray, bone scintigraphy, CT, MRI) were performed, as needed, whenever the clinical or laboratory findings raised the suspicion of persistent or recurrent disease.

Post-treatment Whole-Body Scan

Planar I-131 ptWBS was performed in both anterior and posterior projections using a dual-detector gamma camera (Infinia, GE Healthcare) with high-energy, parallel-hole collimators. Continuous acquisition mode was used at a table speed of 6 cm/min, 1024x256 matrix and automated body contour detection applied. The photopeak was centered at 364 keV with a ± 10% window. Additional spot views acquired for 5 min/view with 256x256 matrix were required in case of unexpected iodine uptakes and accumulations suspected of contamination (scanned after removing the contamination or the stained clothes).

All cases with distant uptake on ptWBS were reviewed by two independent physicians (CFN and JMD) and evaluated for the presence of distant metastases at the uptake sites through imaging tests (CT, MRI, bone scintigraphy) associated with postoperative laboratory tests (sPOTg, TSH and TgAb). Those with distant uptake on ptWBS and with images and laboratory tests confirming the presence of distant metastases in that same topography were classified as true-positive. If the distant uptake on ptWBS was not confirmed by other exams (CT, MRI, bone scintigraphy, biopsies) it was classified as a false-positive uptake. Cases with no distant uptake on ptWBS, but with proven distant metastases in other exams (CT, MRI, bone scintigraphy, biopsies), were considered false-negative uptakes.

Outcomes

In the first year of follow-up, the disease outcome of initial therapy was assessed based on the serum Tg-T4 or sTg levels, neck US, and additional imaging exams, whenever indicated.

An excellent response was defined as having no clinical or imaging evidence of tumor (i.e., no imaging evidence of tumor on neck US or CT), undetectable (<0.2 ng/mL) serum Tg-T4 levels and/or sTg levels <1 ng/mL. Indeterminate response was defined as nonspecific findings on imaging

studies, Tg-T4 detectable but < 1 ng/mL, sTg detectable but < 10 ng/mL or TgAb stable or declining in the absence of structural or functional disease. Biochemical incomplete response was defined as negative imaging and Tg-T4 ≥ 1 ng/ml or sTg levels ≥ 10 ng/mL or rising TgAb levels. Structural incomplete response was defined as structural or functional evidence of disease at any Tg-T4 level with or without TgAb (5). Patients who were diagnosed with persistent disease were evaluated for additional treatment (e.g., surgery, RAI, external-beam radiation or multi-kinase inhibitors), depending on the site involved and disease progression.

All patients with biochemical incomplete and structural incomplete responses were considered to have persistent disease. Recurrence was defined as new biochemical or structural evidence of the disease detected in a patient who had been previously classified as having an excellent response.

Laboratory analysis

Serum Tg measurements were conducted using immunoradiometric assays (from 2000 to 2002 radioimmunoassay; 2002 to 2010 electrochemiluminescence and 2010 until the present chemiluminescence) that indicated a functional sensitivity of 0.2 ng/mL. TgAb were measured using the passive agglutination method from 2000 to 2010 and by chemiluminescence from 2010 until the present. After each new assay had been implemented, the necessary procedures for standardization and validation were performed. TSH levels were measured by chemiluminescence assay from 2000 to 2006 (Immulite 2000 SIEMENS), electrochemiluminescence from 2006 to 2010 (Modular E ROCHE), chemiluminescence assay from 2010 to 2014 (Centaur XP SIEMENS) and electrochemiluminescence from 2014 until the present (Cobas E602 ROCHE). These tests were all conducted in the HCPA central laboratory.

Statistical analysis

The clinical and laboratory data are reported as the mean ± standard deviation (SD) values or as the median and percentiles 25 and 75 (P25-75) for continuous variables, or as absolute numbers and percentages for categorical variables. Comparative analyses of frequencies were performed

using Pearson Chi-Square or Fisher's Exact Test, as appropriate. These analyses were performed using the Statistical Package for Social Science Professional software version 20.0 (IBM Corp., Armonk, NY).

Sensitivity and specificity of ptWBS were calculated for the overall population, and for specific subgroups. Agreement of ptWBS with the gold standard (clinical classification) was accessed using the *Cohen's Kappa* coefficient for two-level mutually exclusive categories using VassarStats®. Comparisons of predictive value for ptWBS diagnostic performance between different DTC populations were calculated using the software WinPepi® version 11.65 with the Weighted Generalized Score (WGS) test of Kosinski. Youden's index was calculated using sensitivity and specificity for comparisons of the performance of the ptWBS between two different populations.

All tests were two-tailed, and a $p < 0.05$ was considered statistically significant.

Results

Clinical characteristics

Two hundred and sixty-eight DTC patients were evaluated. Clinical and oncological characteristics of these patients are described in **Table 1**. The mean age at diagnosis was 46 ± 16 years, 220 (82%) were women and PTC was diagnosed in 234 (87%) patients. The median tumor size was 2.7 cm (P25-75 1.3-3.5cm). One hundred and seven patients (40%) had lymph node metastases, and 29 (11%) patients had distant metastases. Two hundred and twelve patients (80%) were classified as stage I of TNM/AJCC, 41 (15%) as stage II, 2 (1%) as stage III and 10 (4%) as stage IV. One hundred and thirty-seven patients (52%) were ATA low risk, 96 (35%) ATA intermediate risk and 32 (12%) ATA high-risk. All patients underwent total thyroidectomy and received RAI**.** The mean administered RAI activity was 96 mCi (30-200mCi).

Post-treatment Whole-Body Scan

Two hundred and sixty-eight patients had a ptWBS uptake image available, 9 (4%) patients showed no uptake, 231 (86%) showed uptake on

the thyroid bed and 28 (10%) patients showed distant uptake on ptWBS. Twenty-nine patients had distant metastases: 14 patients with known metastases prior to RAI (11 with true-positive ptWBS uptake) and 15 patients had a diagnosis of metastases after RAI (9 with true-positive ptWBS uptake).

Of the 231 patients with uptake only on the thyroid bed, 9 (4%) had ptWBS considered false-negative. The patterns of metastases of the false-negative uptakes were: 5 patients with pulmonary metastases, 2 with bone metastases and 2 with metastases at multiple sites.

Of the 28 patients with distant uptake at ptWBS: 19 (68%) patients had ptWBS considered true-positive and 9 (32%) considered false-positive. The patterns of the false-positive uptake on ptWBS were: 6 on the thorax, 2 on the spine, and 1 on the pelvis. Five uptakes were considered focal, 3 diffuse and 1 diffuse with focal areas. The etiologies for false-positive uptakes were: 5 indeterminate, 2 osteoarthrosis, 1 renal cyst and 1 adrenal incidentaloma **(Table 2)**.

The overall sensitivity of ptWBS for distant metastases was 68%, with an specificity of 96%.

Post-treatment Whole-Body Scan and ATA Risk Score

One hundred thirty-seven patients were classified as ATA low risk, 6 (4%) had no ptWBS uptake, 128 (94%) had cervical and 3 (2%) distant uptake. Of the three ATA low risk patients with distant uptake, 2 (67%) were considered false-positive and 1 (33%) true-positive. In this group we had 2 (1%) false-negative. ptWBS sensitivity and specificity for distant metastases in this ATA low risk population were 33% and 98%, respectively.

Ninety-six patients were ATA intermediate risk, 2 (2%) without uptake on ptWBS, 88 (92%) with cervical uptake and 6 (6%) with distant uptake. Among those 6 ATA intermediate risk patients with distant uptake, all were considered false-positive. False-negative ptWBS occurred in 3 (3%) patients, rendering a ptWBS sensitivity of 0% and a specificity of 94% for ATA intermediate risk for distant metastases.

Thirty-two patients were ATA high-risk, 1 (3%) showed no uptake on ptWBS, 13 (41%) had cervical uptake and 18 (56%) distant uptake, all considered true-positive. For ATA high-risk patients we had 4 (13%) false-

negative exams. This figures resulted in a ptWBS senstivity and specificity of 82% and 100%, respectively, for the presence of distant metastases for ATA high-risk. **(Table 3/Figure 1**). Comparing the performance of ptWBS for the diagnosis of distant metastases according to the ATA risk score: for low-intermediate vs high-risk patients the sensitivity was 22% vs 83%, specificity was 96% vs 100% and the Youden index was 18% vs 83 % ($p \leq 0.0001$) respectively. Additionally, when we excluded ATA low risk patients and compared ATA intermediate risk vs high-risk patients, sensitivity was 82% vs 0% and specificity was 100% vs 94% and the Youden index was 83% vs -5% ($p \leq 0.0001$) respectively.

Post-treatment Whole-Body Scan and sPOTg

A hundred and seventy-four patients with negative TgAb and sPOTg values were included in these analyses, and 94/268 were excluded due to a positive TgAb result. We further stratified the DTC cohort according to sPOTg into three categories (<1.0 ng/mL, between 1-10 ng/mL and >10 ng/mL). The cutoff point of 10ng/mL were chosen based on a meta-analysis of 3947 patients (21).

Of the 28 patients with a sPOTg<1.0 ng/mL, 2 (7%) had no ptWBS uptake, 25 (89%) had cervical and 1 (4%) distant uptake. The uptake of the one patient with sPOTg<1.0 ng/mL was considered false-positive precluding the analysis of sensitivity, and rendering a ptWBS specificity of 96% for distant metastases in this population. No patient was considered false-negative in this group.

Of the 71 patients with sPOTg 1-10 ng/mL; 68 (96%) had cervical ptWBS uptake and 3 (4%) had distant uptake. Of the 3 with distant uptake, 1 (1%) was considered false-positive and 2 (3%) true-positive. One (1%) patient was considered false-negative. Sensitivity and specificity of ptWBS for this population was 67% and 99%, respectively.

Seventy-five patients had a sPOTg>10 ng/mL, 1 (1%) had no ptWBS uptake; 60 (80%) had cervical and 14 (19%) distant uptake. Of the 14 patients with sPOTg>10 ng/mL and distant uptake, 5 (7%) were considered false-positive and 9 (12%) true-positive. In this group 4 (5%) patients were considered false-negative. Sensitivity and specificity of ptWBS for distant

metastases for sPOTg>10 ng/mL patients were 69% and 92%, respectively **(Table 3)**.

For the comparisons of the performance of ptWBS for the diagnosis of distant metastases according to sPOTg we had to exclude the sPOTg<1 ng/mL category due to the impossibilitiy of calculating the sensitivity in this population. Comparing the performance of ptWBS for patients with sPOTg<10.0 ng/mL vs sPOTg>10 ng/mL the sensitivity was 75% vs 71%, specificity was 98% vs 92% and the Youden index 73% vs 63% (p=0.69), respectively. Comparing only patients with sPOTg 1-10 ng/mL vs sPOTg>10 ng/mL the sensitivity was 67% vs 71%, specificity was 99% vs 92% and the Youden index 65% vs 63% (p=0.95).

Discussion

ptWBS has pitfalls that limit its performance, mainly due to false-positive distant uptake on patients with a low pre-test probability of metastases. However, our findings suggest that for selected high-risk DTC patients, it seems to be a useful tool to identify those with distant metastases.

Management of DTC patients is under continuous review, and has changed substantially in recent years. Thus, we have questioned whether the mandatory ptWBS is plausible and which population would benefit from the test in the current context. False-positive ptWBS is a potential source of unintended psychological stress and wasteful financial health care expenditures, as a consequence of the uncertain possibility of distant metastases and the numerous exams needed to clarify the abnormal uptake.

False-positive ptWBS uptake has been documented in the literature since 1988 (11-15,19). Like any other exam, ptWBS has flaws, and is subject to diagnostic and interpretation problems. Any area that concentrates iodine-131 (I-131) may be interpreted as a region of thyroid cancer, but functional expression of NIS in other tissues, iodine retention in body fluids, technical contamination, cysts, bronchiectasis, arthrosis, bone fractures, inflammation, mammary glands and other tumors may exhibit the same pattern of uptake and can cause false-positive findings on ptWBS (16-19,25). False-negatives on ptWBS are well known. A negative ptWBS may misleadingly suggest DTC

is under control, which may happen when distant metastases have not been identified yet, whereas those patients with known distant metastases, the negative ptWBS uptake indicates that the tumor is undergoing a dedifferentiation process (26,27).

Oral A et al showed that the incidence of ptWBS false-positives is greater after the first dose of iodine, with predominance for focal uptake. The most frequent site for false-positive uptakes is the lung, followed by the pelvis. Besides that, in one third of patients the etiology of false-positives can not be determined (28) In our cohort we analysed only ptWBS after the first activity of RAI, and found that the main false-positive pattern was focal uptake, on the thorax. Of note in 5/9 (56%) patients we couldn't determine the etiology of the false positive uptake.

Given the limitations of ptWBS for the whole DTC population, we explored the performance of the exam according to DTC patient characteristics. Stratifying the DTC population according to the ATA risk score, the performance of ptWBS for ATA low-intermediate risk was poor, with a predominance of false-positive uptakes over true-positive, posing a greater risk of misdiagnosing patients. Considering that currently ATA low risk patients should not receive RAI, we analyzed the ATA intermediate risk patients separated from ATA low risk. Notwithstanding, for ATA intermediate risk patients the performance of ptWBS was even worse (sensitivity 0% and specificity 94%). On the other hand, for ATA high-risk patients, ptWBS displayed high sensitivity, specificity and level of agreement, contributing to the diagnosis of distant metastases in a substantial portion of this subgroup of patients and showing that if the pre-test probability was high for distant metastases, the exam has a positive predictive value of 100%.

We further stratified the DTC cohort according to sPOTg levels, using a cut-off point of 10 ng/mL, as suggested by a systematic review of sPOTg prognostic value (21). For patients with sPOTg<10ng/mL, the rates of distant uptakes were low, and false-positive uptakes were equal to true-positive. For those patients with sPOTg>10ng/mL, distant uptakes were more common, and distant true-positive exams exceeded false-positive. Notwithstanding, ptWBS performance was no different for patients with sPOTg<10ng/mL vs sPOTg>10ng/mL.

Our study has some limitations: it was conducted at a tertiary center, specialized in the treatment of DTC patients which can lead to sample bias. As a note of caution, to confirm the external validity of our findings, it will be of interest to replicate this study in cohorts from different contexts. Moreover, at our institution we do not perform SPECT/CT or FDG-PET, which could reduce the number of additional tests performed to confirm the nature of obscure ptWBS uptake.

In summary, our results suggest that given the poor overall performance of ptWBS, we should reconsider its routine use for all DTC patients. Notwithstanding, for ATA high-risk DTC patients, ptWBS seems to be a useful tool.

Acknowledgments: This work was a possible due to grants from CNPq, CAPES, FIPE and PRONEX / FAPERGS. We wish to thank Andreia C. F. da Silveira for helping with nuclear medicine settings and Aline Castello Branco Mancuso for clarifying the statistics pitfalls.

Author Disclosure Statement: C.F.N., R.S.S., A.B.Z, F.Z., A.L.M. and J.M.D. have nothing to declare.

Corresponding Author

Jose Miguel Dora, M.D., Ph.D.

Thyroid Unit, Endocrine Division, Hospital de Clínicas de Porto Alegre

Rua Ramiro Barcelos, 2350. CEP 90035 –003

Porto Alegre, RS, Brazil

Phone: 55-51-3359-7858/ Fax: 55-51-3332-5188

E-mail: jdora@hcpa.edu.br

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A 2018 Global Cancer Statistics: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA CANCER J CLIN 2018;68:394–424.

2. Gilliland, F. D., Hunt, W. C., Morris, D. M., & Key, C. R. 1997 Prognostic factors for thyroid carcinoma. Cancer,79(3), 564–573.

3. Mazzaferri, E. L., & Kloos, R. T. 2001 Clinical review 128: Current approaches to primary therapy for papillary and follicular thyroid cancer. The Journal of Clinical Endocrinology and Metabolism,86(4),1447–63.

4. Scheffel RS, Zanella AB, Antunes D, Dora JM, Maia AL 2015 Low Recurrence Rates in a Cohort of Differentiated Thyroid Carcinoma Patients: A Referral Center Experience. Thyroid,25(8), 883-9.

5. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, Pacini F, Randolph GW, Sawka AM, Schlumberger M, Schuff KG, Sherman SI, Sosa JA, Steward DL, Tuttle RM, Wartofsky L. 2016 2015 American Thyroid Association management guidelines for adult patients with thyroid nodules and differentiated thyroid cancer. Thyroid 26:1‑133.

6. Tuttle M, Morris LF, Haugen B, Shah J, Sosa JA, Rohren E, Subramaniam RM, Hunt JL, Perrier ND. Amin MB, Edge SB, Greene F, Byrd D, Brookland RK, Washington MK, Gershenwald JE, Compton CC, Hess KR, Sullivan DC, Jessup JM, Brierley J, Gaspar LE, Schilsky RL, Balch CM, Winchester DP, Asare EA, Madera M, Gress DM, Meyer LR, 2017 (eds)Thyroid‑Differentiated and Anaplastic Carcinoma (Chapter 73) In: AJCC Cancer Staging Manual. 8thed. Springer International Publishing. New York City.

7. Ahn BC 2012 Sodium iodide symporter for nuclear molecular imaging and gene therapy: from bedside to bench and back. Theranostics; 2: 392-402.

8. Chung JK 2002 Sodium iodide symporter: its role in nuclear medicine. J Nucl Med; 43: 1188- 1200.

9. Pacini, F., Agate, L., Elisei, R., Capezzone, M., Ceccarelli, C., Lippi, F., … Pinchera, A. 2001 Outcome of Differentiated Thyroid Cancer with Detectable Serum Tg and Negative Diagnostic 131 I Whole body scan: Comparison of

patients treated with high 131 I Activities versus untreated patients, J Clin Endocrinol Metab. Sep;86(9):4092-7.

10. Fatourechi, V., Hay, I. D., Mullan, B. P., Wiseman, G. a, Eghbali-Fatourechi, G. Z., Thorson, L. M., & Gorman, C. a. 2000 Are posttherapy radioiodine scans informative and do they influence subsequent therapy of patients with differentiated thyroid cancer? Thyroid : Official Journal of the American Thyroid Association, 10(7), 573–7.

11. Yildirim-Poyraz N, Ozdemir E, Amutkan C, Adiyaman N, Kilinc S, Kandemir Z, Saglam F, Turkolmez S, Cakir B 2013 False-positive iodine-131 whole body scan due to a benign dermal lesion; intradermal nevus ((131)I uptake in a benign nevus). Ann Nucl Med Oct;27(8):786-90.

12. Chaussé G, Kader T, Abikhzer G, Probst S 2018 Asthmatic Exacerbation as a Cause of False-Positive Whole-Body Iodine Scan in a Patient With Treated Papillary Thyroid Carcinoma. Clin Nucl Med Apr;43(4):256-257.

13. Ceccarelli C, Pacini F, Lippi F, Pinchera A 1988 An unusual case of a false-positive iodine-131 whole body scan in a patient with papillary thyroid cancer. Clin Nucl Med Mar;13(3):192-3.

14. Wang R, Zhou K, Fan Q, Chen H, Fan C 2018 A false-positive I-131 finding of duodenum diverticulum in thyroid cancer evaluation by SPECT/CT: A case report. Medicine (Baltimore) Feb;97(8):e9997.

15. Maciejewski A, Czepczyński R, Ruchala M 2018 False-positive radioiodine whole-body scan due to a renal cyst. Endokrynol Pol ;69(6):736-739.

16. Buton L, Morel O, Gault P, Illouz F, Rodien P, Rohmer V 2013 False-positive iodine-131 whole-body scan findings in patients with differentiated thyroid carcinoma: report of 11 cases and review of the literature. Ann Endocrinol (Paris) Jul;74(3):221-30.

17. Hannoush ZC, Palacios JD, Kuker RA, Casula S 2017 False Positive Findings on I-131 WBS and SPECT/CT in Patients with History of Thyroid Cancer: Case Series. Case Rep Endocrinol; 2017:8568347.

18. Oh JR, Ahn BC 2012 False-positive uptake on radioiodine whole-body scintigraphy: physiologic and pathologic variants unrelated to thyroid cancer. Am J Nucl Med Mol Imaging;2(3):362-85.

19. Shapiro B, Rufini V, Jarwan A, Geatti O, Kearfott KJ, Fig LM, Kirkwood ID, Gross MD 2000 Artifacts, anatomical and physiological variants, and unrelated diseases that might cause false-positive whole-body 131-I scans in patients with thyroid cancer. Semin Nucl Med Apr;30(2):115-32.

20. Baudin E, Do Cao C, Cailleux AF, Leboulleux S, Travagli JP, Schlumberger M 2003 Positive predictive value of serum thyroglobulin levels, measured during the first year of follow-up after thyroid hormone withdrawal, in thyroid cancer patients. J Clin Endocrinol Metab Mar;88(3):1107-11.

21. Webb, R. C., Howard, R. S., Stojadinovic, A., Gaitonde, D. Y., Wallace, M. K., Ahmed, J., & Burch, H. B. 2012 The utility of serum thyroglobulin measurement at the time of remnant ablation for predicting disease-free status in patients with differentiated thyroid cancer: A meta-analysis involving 3947 patients. Journal of Clinical Endocrinology and Metabolism, 97(8), 2754–2763.

22. Cooper DS, Doherty GM, Haugen BR, Kloos RT, Lee SL, Mandel SJ, Mazzaferri EL, McIver B, Pacini F, Schlumberger M, Sherman SI, Steward DL, Tuttle RM 2009 Revised american thyroid association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. Thyroid 19:1167–1214.

23. Rosario PW, Ward LS, Carvalho GA, Graf H, Maciel RM, Maia AL, Vaisman M 2013 Thyroid nodules and differentiated thyroid cancer: Update on Brazilian consensus. Arquivos Brasileiros de Endocrinologia e Metabologia 57: 205–232.

24. Nava CF, Zanella AB, Scheffel RS, Maia AL, Dora JM 2019 Impact of the updated TNM staging criteria on prediction of persistent disease in a differentiated thyroid carcinoma cohort. Arch Endocrinol Metab;63(1):5-11.

25. Garger YB, Winfeld M, Friedman K, Blum M 2016 In Thyroidectomized Thyroid Cancer Patients, False-Positive I-131 Whole Body Scans Are Often Caused by Inflammation Rather Than Thyroid Cancer. J Investig Med High Impact Case Rep. Feb 25;4(1):2324709616633715.

26. Durante C., Puxeddu E., Ferretti E., Morisi R., Moretti S., Bruno R., Barbi F., Avenia N., Scipioni A., Verrienti A., Tosi E., Cavaliere A., Gulino A., Filetti S. , and Russo D. 2007 BRAF mutations in papillary thyroid carcinomas inhibit

genes involved in iodine metabolism.*J. Clin. Endocrinol. Metab.* 92, 2840–2843.

27. Debyani Chakravarty, Elmer Santos, Mabel Ryder, Jeffrey A. Knauf, Xiao-Hui Liao, Brian L. West, Gideon Bollag, Richard Kolesnick, Tin Htwe Thin, Neal Rosen, Pat Zanzonico, Steven M. Larson, Samuel Refetoff,Ronald Ghossein, and James A. Fagin 2011 Small-molecule MAPK inhibitors restore radioiodine incorporation in mouse thyroid cancers with conditional BRAF activation. J Clin Invest; 121:4700–11.

28. Oral A, Yazici B, Eraslan C, Burak Z 2018 Unexpected False-positive I-131 Uptake in Patients with Differentiated Thyroid Carcinoma. Mol Imaging Radionucl Ther. Oct 9;27(3):99-106.

Table 1. Characteristics of 268 patients with differentiated thyroid carcinoma.

Age at diagnosis (years)	46 ± 16
Female sex – n (%)	220 (82)
Histology – n (%)	
Papillary	234 (87)
Follicular	34 (13)
Tumor size (cm)	2.7 (1.3-3.5)
Lymph node metastases (N1) – n (%)	107 (40)
Distant metastases – n (%)	29 (11)
8th TNM Stage – n (%)	
I	212 (80)
II	41 (15)
III	2 (1)
IV	10 (4)
ATA Risk Score - n (%)	
Low	137 (52)
Intermediate	96 (36)
High	32 (12)
Radioactive iodine activity (mCi)	96 ± 34
Follow-up (years)	4 (2-6)

Data are expressed as the mean ± SD, median (percentiles 25-75) or frequencies.

TNM Stage, tumor, node, metastases stage. ATA risk score, American Thyroid Association risk score.

Table 2. False-positive patients on post-treatment whole-body scan uptake.

Patient	TNM8	ATA risk	ptWBS Uptake	Confirmatory exam	Type of uptake	Etiology	Current status	sPOTg TgAb
1	I	Low	Toracic spine	CT and bone scintigraphy	Focal	Inflammatory	Biochemical incomplete	97,2 NR
2	I	Intermediate	Thorax	CT	Focal	Adrenal incidentaloma	Indeterminate	89,8 NR
3	II	Intermediate	Thorax	CT	Diffuse	Indeterminate	Structural incomplete	34,2 NR
4	NA	NA	Lumbar spine	CT and bone scintigraphy	Focal	Inflammatory	Excellent	4,91 NR
5	I	Intermediate	Thorax	CT	Focal and Diffuse	Indeterminate	Excellent	18,17 NR
6	I	Low	Thorax	CT	Focal	Indeterminate	Excellent	NA
7	I	Intermediate	Thorax	CT	Diffuse	Indeterminate	Indeterminate	26 NR
8	I	Intermediate	Pelvis	CT and bone scintigraphy	Focal	Cystic	Structural incomplete	NA
9	I	Intermediate	Thorax	CT	Diffuse	Indeterminate	Excellent	0.1 NR

TNM8, tumor, node, metastases system eighth edition. ATA risk score, American Thyroid Association risk score. ptWBS Uptake, post-treatment whole-body scan uptake. sPOTg, stimulated postoperative thyroglobulin. TgAb, anti-thyroglobulin antibody. CT, computed tomography. NR, non-reactive. RA, reactive. NA, not available.

Table 3. Performance of post-treatment whole-body scan between different groups of differentiated thyroid carcinoma.

	All (n=268)	American Thyroid Association risk score			Stimulated Postoperative Thyroglobulin		
		Low (n=137)	Intermediate (n=96)	High (n=32)	<1 ng/mL (n=28)	1-10ng/mL (n=71)	>10ng/mL (n=75)
Sensitivity	0.68	0.33	0.00	0.82	NC	0.67	0.69
Specificity	0.96	0.99	0.94	1.00	0.96	0.99	0.92
Kappa	0.66	0.27*	0.19*	0.74*	NC	0.23*	0.49*
FP - n (%)	9 (3%)	2 (2%)	6 (6%)	0 (0%)	1 (4%)	1 (1%)	5 (7%)
TP - n (%)	19 (7%)	1 (1%)	0 (0%)	18 (56%)	0 (0%)	2 (2%)	9 (12%)
FN - n (%)	9 (3%)	2 (1%)	3 (3%)	4 (13%)	0 (0%)	1 (1%)	4 (5%)

FP, false-positive. TP, true-positive. FN, false-negative. NC, not calculable.

** p≤0.0001 for ATA high-risk vs ATA low-intermediate risk, p=0.074 for ATA high-risk vs ATA low risk* and *p≤0.0001 for ATA high-risk vs ATA intermediate risk, p=0.69 for sPOTg <10ng/mL vs sPOTg >10ng/mL, p=0.95 for sPOTg 1-10ng/mL vs sPOTg > 10ng/mL.*

Figure 1. Post-treatment Whole-Body Scan uptake, stratified by the American Thyroid Association (ATA) risk score.