META-ANALYSIS

Diagnostic values of thyroglobulin measurement in fine-needle aspiration of lymph nodes in patients with thyroid cancer

Kyoungjune Pak · Sunghwan Suh · Hyunsook Hong · Gi Jeong Cheon · Seo Kyung Hahn · Keon Wook Kang · E. Edmund Kim · Dong Soo Lee · June-Key Chung

Received: 27 June 2014/Accepted: 27 August 2014/Published online: 4 September 2014 © Springer Science+Business Media New York 2014

Abstract We aimed to evaluate the diagnostic performance of measuring the concentration of thyroglobulin (Tg) in the washout fluid of the needle aspiration (FNA-Tg). We performed a systematic search of MEDLINE (inception to October 2013) and EMBASE (inception to October 2013) for English publications using keywords "thyroid", "aspiration", "washout", and "thyroglobulin". All searches were limited to human studies. We included studies of FNA-Tg measurement in the washout fluid rinsed with 1 ml of normal saline. Eight studies including 843 lymph nodes (LNs) were eligible for this study. The pooled sensitivity and specificity of preoperative studies are 0.89 [95 % CI 0.82–0.95], 0.60 [0.49–0.70], and those of postoperative studies are 1.0 [0.83–1.0], 1.0 [0.92–1.0]. To determine best cutoffs from each preoperative and

Kyoungjune Pak, Sunghwan Suh, and Hyunsook Hong have contributed equally to this work.

Electronic supplementary material The online version of this article (doi:10.1007/s12020-014-0410-z) contains supplementary material, which is available to authorized users.

K. Pak · G. J. Cheon (⊠) · K. W. Kang ·
D. S. Lee · J.-K. Chung
Department of Nuclear Medicine, Seoul National University
College of Medicine, 101 Daehak-ro, Jongno-gu, Seoul 110-744,
South Korea
e-mail: larrycheon@gmail.com

K. Pak e-mail: ilikechopin@daum.net

K. W. Kang e-mail: kangkw@snu.ac.kr

D. S. Lee e-mail: dsl@snu.ac.kr

J.-K. Chung e-mail: jkchung@snu.ac.kr postoperative study, the distance between the point (0, 1) and each observed cutoff values (1-specificity, sensitivity) was calculated, and the distance is minimal when the cutoff value of 32.04 for preoperative studies and of 0.9 for postoperative one are selected. FNA-Tg can be used for both preoperative and postoperative evaluation of LN metastasis. Although the cutoff values for the FNA-Tg has not been standardized, preoperative values of 32.04 ng/ml and postoperative values of 0.9 ng/ml are recommended for identifying neck LN metastasis

Keywords Thyroid cancer · Biopsy · Fine-needle · Thyroglobulin

Introduction

Differentiated thyroid cancer (DTC) is the most common endocrine cancer and its incidence has continuously increased in the last three decades all over the world including Korea [1, 2]. Although DTC is perceived as a cancer with a high survival

K. Pak

Department of Nuclear Medicine, Pusan National University Hospital, 179 Gudeok-ro, Seo-gu, Busan, 602-739, South Korea

S. Suh

Department of Internal Medicine, Dong-A University College of Medicine, 26 Daesingonwon-ro, Seo-gu, Busan, South Korea e-mail: taret@daum.net

H. Hong (⊠) · S. K. Hahn
Medical Research Collaborating Center, Seoul National
University College of Medicine, 101 Daehak-ro, Jongno-gu,
Seoul 110-744, South Korea
e-mail: hshong@snuh.org

rate, this sustained increase in DTC incidence worldwide has recently attracted considerable attention. DTC treatment employs a three-tiered approach including surgery, radioactive iodine treatment and long-term exogenous hormone replacement [3, 4]. Despite the above-mentioned excellent prognosis, a minority of these patients shows the recurrence [5]. Therefore, DTC care places an inevitable demand on social health care resources and economic burden.

During follow-up, serum markers (thyroglobulin; Tg) and imaging workup (neck ultrasonography and iodine whole body scan) are used to identify persistence or recurrence of disease. In case of suspicious ultrasonography findings, fine-needle aspiration (FNA) is generally required to cytologically confirm or exclude recurrence [6]. However, cytological analysis of FNA (FNA-C) samples may be inadequate or even false-negative, especially in lymph nodes (LN) with small metastases or partial involvement or cystic changes [7]. To improve the diagnostic yield of FNA, directly measuring the concentration of Tg in the washout fluid of the needle aspiration used in FNA (FNA-Tg) has been proposed by several authors since the first study conducted by Pacini et al. [8]. Although FNA-Tg serves as an important standard for suspicious recurrent or metastatic lesions [9], the cutoff value to define "high" is still controversial. Moreover, difference in sample treatments, Tg assays, and reference values may result in difficulty in comparing studies.

Therefore, we meta-analyze the previous studies regarding the diagnostic performance of FNA-Tg in patients with preoperative or postoperative thyroid cancers and suggest the optimal cutoff values of FNA-Tg.

Materials and methods

Data search and study selection

We performed a systematic search of MEDLINE (inception to October 2013) and EMBASE (inception to October 2013)

S. K. Hahn

E. E. Kim · D. S. Lee

WCU Graduate School of Concergence Science and Technology, Seoul National University College of Medicine, 101 Daehak-ro, Jongno-gu, Seoul, 110-744, South Korea e-mail: eedkim@yahoo.com

E. E. Kim

for English publications using keywords "thyroid," "aspiration," "washout," and "thyroglobulin." All searches were limited to human studies. Inclusion criteria were studies with thyroid FNA used in the initial workup (preoperative) or the follow-up (postoperative), FNA-Tg measured in the washout fluid rinsed with 1 ml of normal saline and the reference standards based on the pathologic reports. We excluded studies that used less or more than 1 ml of normal saline to wash the needle and syringe or other solution. Reviews, abstracts, and editorial materials were excluded. Two authors conducted the searches and screening independently. Any discrepancies were resolved by a consensus.

Quality assessment

Two reviewers independently assessed study quality using the Quality Assessment of Studies of Diagnostic Accuracy Studies (QUADAS-2) [10]. The QUADAS-2 tool assesses the quality of included studies in terms of the risk of bias and applicability to the clinical question for patient selection, index test, reference standard, and flow/timing. Studies were categorized into low/high/unclear about each domain.

Data synthesis and statistical analysis

For each study, we constructed a 2×2 contingency table with true positive, false-positive, false-negative, and true negative. Pooled sensitivity and specificity were estimated for each cutoff and heterogeneity was assessed by I^2 statistic, as described by Higgins et al. [11]. Summary receiver operating characteristic (SROC) curves with pooled sensitivity and specificity were estimated using Moses-Littenberg method [12]. To avoid calculation problems by having zero values, 0.5 was added to each cell of the respective contingency table. Threshold effect on heterogeneity was evaluated by Spearman correlation coefficient between sensitivity and 1-specificity [12]. To explore the influence of studies that reported multiple cutoffs, we conducted a sensitivity analysis using one cutoff for each study that is nearest the upper left corner of the ROC graph. Exact 95 % confidence intervals for the sensitivity and specificity were calculated based on a binomial distribution. Data were analyzed with SAS software (version 9.3).

Results

Study characteristics and quality assessment

440 articles were identified through database searching. After excluding non-English articles (n = 54) and

e-mail: hahns@snu.ac.kr

G. J. Cheon \cdot K. W. Kang \cdot D. S. Lee \cdot J.-K. Chung Cancer Research Institute, Seoul National University College of Medicine, Seoul, South Korea

Department of Radiological Science, University of California at Irvine, CA, USA



Fig. 1 Flowchart of study selection

conference abstracts (n = 121), 265 abstracts were assessed for eligibility. 8 studies including 843 LNs were eligible for this study [13–20]. The detailed procedure is drawn as a flowchart in Fig. 1. Studies were divided into 2 groups of preoperative and postoperative Tg measurement. The summary of studies included is presented in Table 1. Quality assessment was conducted on all 8 studies included in meta-analysis using QUADAS-2. Generally, studies met most of the quality criteria (Supplementary 1).

Diagnostic performance according to the cutoffs

Figure 2 shows the SROC curve of the 8 studies. The study with cutoff value using serum Tg is excluded in SROC curve [21]. The pooled sensitivity and specificity of preoperative studies are 0.89 [95 % CI 0.82–0.95], 0.60 [0.49–0.70], and those of postoperative studies are 1.0 [0.83–1.0], 1.0 [0.92–1.0].

Data from preoperative studies

Data from 3 preoperative studies [13, 15, 18] were included in calculation of sensitivity and specificity. Six different cutoff values (100, 39.3, 32.04, 10, 5, and 1 ng/ml) were adopted. The pooled sensitivity and specificity of each cutoff are summarized in Table 2. Figure 3a shows the trend of the specificity decrease according to the decrease of cutoffs; however, the sensitivity did not meet the trend. The study by Sohn et al. [15] showed relatively lower sensitivity. Four different methods were applied to determine cutoffs. 97.5 percentile of FNA-Tg in cytologically negative LNs was used as cutoff value in 1 study [13], arbitrarily decided cutoffs in 2 studies [15, 18], and mean + 2SD of FNA-Tg in cytologically negative LNs in

Author	Year	Country	State of operation	No. of patients	No. of LNs	Needle (gauge)	Tg measurement	Methods of determining cutoffs	Cutoffs (ng/ml)
rasoldati [13]	1999	Italy	Pre	23	23	22/23	IRMA	97.5 percentile of FNA-Tg in	39.3
			post	9	13	22/23	IRMA	cytologically negative LNs	1.1
Junha [14]	2007	Portugal	Post	51	65	22/25	CIA	Functional sensitivity	0.0
Cim [15]	2009	Korea	Pre	100	100	23	IRMA	Arbitrarily/mean + 2SD of FNA-	1/10/32.04/100/serum Tg
			Post	68	68		IRMA	Tg in negative LNs/serum Tg	1/10/32.04/100/serum Tg
anella [16]	2010	Brazil	Post	43	43	22-25	ECLA	Arbitrarily	10
iovanella [17]	2011	Switzerland	Post	108	126	21	IRMA	ROC	1.1
ohn [18]	2012	Korea	Pre	92	95	23	IRMA	Arbitrarily	1/5/10
eon [19]	2013	Korea	Post	207	263	23	IRMA	Arbitrarily	10
uh [20]	2013	Korea	Post	43	47	23	IRMA	Arbitrarily	1/10/serum Tg



Fig. 2 Summary receiver operating characteristic of preoperative and postoperative studies (*black circle* preoperative studies, *star* operator postoperative studies)

1 study [15]. Pooled sensitivity with each cutoff ranged from 0.69 to 0.95, and specificity between 0.53 and 1.0.

Data from postoperative studies

All 7 postoperative studies [13–17, 19, 20] were included in calculation of sensitivity and specificity. Six different cutoff values (100, 32.04, 10, 1.1, 1, and 0.9 ng/ml) were applied in postop studies. The pooled sensitivity and specificity of each cutoff are summarized in Table 3. The pooled sensitivity of each cutoff shows decrease according to the increase of cutoffs; however, the specificity did not meet the trend (Fig. 3b). Pooled sensitivity of each cutoff ranged between 0.79 and 1.0; specificity between 0.5 and 1.0.

Determination of best cutoff values

To determine best cutoffs from each preoperative and postoperative study, the distance between the point (0, 1) and each observed cutoff values (1-specificity, sensitivity) was calculated, and the closest to the (0, 1) was selected. The distance is minimal when the cutoff value of 32.04 for preoperative studies and of 0.9 for postoperative one are selected (Tables 2, 3).

Discussion

Collectively, FNA-Tg measurement was shown to be a useful tool to detect neck LN metastases from DTC with

Cutoff (ng/ml) Author Year TP FN FN TN Sample size Sensitivity [95 % CI] Specificity 100 Kim [15] 2009 43 18 0 39 100 0.70 [0.57, 0.81] 1.00 [0.91, 0.90] 39.3 Frasoldati [13] 1999 11 3 0 19 33 0.79 [0.49, 0.95] 1.00 [0.82, 0.81] 1.00 [0.82, 0.91] 1.00 [0.82, 0.91] 1.00 [0.82, 0.91] 1.00 [0.82, 0.91] 0.90 [0.76, 0.92] 1.00 [0.82, 0.92] 0.90 [0.76, 0.92] 0.90 [0.76, 0.92] 0.90 [0.76, 0.92] 0.90 [0.76, 0.92] 0.90 [0.76, 0.92] 0.90 [0.76, 0.92] 0.90 [0.76, 0.92] 0.90 [0.76, 0.92] 0.90 [0.76, 0.92] 0.87 [0.75, 0.92] 0.87 [0.75, 0.92] 0.87 [0.75, 0.92] 0.87 [0.75, 0.92] 0.87 [0.75, 0.92] 0.87 [0.75, 0.92] 0.87 [0.75, 0.92] 0.87 [0.75, 0.92] 0.87 [0.75, 0.92] 0.87 [0.75, 0.92] 0.87 [0.75, 0.92] 0.87 [0.75, 0.92] 0.88 [0.70, 0.92] 0.88 [0.70, 0.92] 0.88 [0.70, 0.92] 0.88 [0.70, 0.92] 0.88 [0.70, 0.92] 0.88 [0.76, 0.92] 0.88 [0.76, 0.92] 0.88 [0.76, 0.92] 0.88 [0.76, 0.92] <th></th>										
100Kim [15]200943180391000.70[0.57, 0.81]1.00[0.91,39.3Frasoldati [13]1999113019330.79[0.49, 0.95]1.00[0.82,32.04Kim [15]20095294351000.85[0.74, 0.93]0.90[0.76,32.04Kim [15]20095475341000.89[0.78, 0.95]0.87[0.75,10Kim [15]20122517746950.60[0.43, 0.74]0.87[0.75,5Sohn [18]20122913944950.60[0.43, 0.74]0.87[0.75,5Sohn [18]20122913944950.60[0.43, 0.74]0.87[0.75,1Kim [15]200958312271000.95[0.86, 0.99]0.69[0.52, 0.82]1Sohn [18]20123482528964950.66[0.53, 0.82]0.69[0.55, 0.52]1Sohn [18]20123482528950.95[0.66, 0.91]0.53[0.52, 0.55]2Sohn [18]20123482528950.56[0.56, 0.91]0.53[0.52, 0.55]1Sohn [18]20123482528950.55[0.56, 0.91]0.53	ml) Author	Year	TP	FN	FР	NT	Sample size	Sensitivity [95 % CI]	Specificity [95 % CI]	Distance from the point (0,1)
39.3Frasoldati [13]19991130000100.8332.04Kim [15]20095294351000.850.74, 0.930.900.76,10Kim [15]20095475341000.890.74, 0.930.900.76,10Kim [15]20122517746950.600.43, 0.740.870.75,5Sohn [18]20122913944950.690.53, 0.820.830.70,1Kim [15]200958312271000.950.86, 0.990.690.52, 0.52,1Sohn [18]20123482528950.610.53, 0.820.690.53, 0.932Sohn [18]20123482528950.610.56, 0.990.690.53, 0.93	Kim [15]	2009	43	18	0	39	100	0.70 [0.57, 0.81]	1.00 [0.91, 1.00]	0.295
32.04Kim [15]2009529435100 $0.85 [0.74, 0.93]$ $0.90 [0.76, 0.73]$ 10Kim [15]2009547534100 $0.89 [0.78, 0.95]$ $0.87 [0.73, 0.73]$ 5Sohn [18]2012251774695 $0.60 [0.43, 0.74]$ $0.87 [0.75, 0.75]$ 5Sohn [18]2012291394495 $0.69 [0.53, 0.82]$ $0.83 [0.70, 0.76]$ 1Kim [15]20095831227100 $0.95 [0.86, 0.99]$ $0.69 [0.52, 0.52]$ 5Sohn [18]2012348252895 $0.81 [0.66, 0.91]$ $0.53 [0.30, 0.52]$	Frasoldati [13]	1999	11	3	0	19	33	0.79 [0.49, 0.95]	1.00 [0.82, 1.00]	0.214
10 Kim [15] 2009 54 7 5 34 100 0.89 [0.78, 0.95] 0.87 [0.73, 0.74] 5 Sohn [18] 2012 25 17 7 46 95 0.60 [0.43, 0.74] 0.87 [0.75, 0.75] 5 Sohn [18] 2012 29 13 9 44 95 0.60 [0.43, 0.74] 0.87 [0.75, 0.75] 1 Kim [15] 2019 58 3 12 27 100 0.95 [0.86, 0.99] 0.69 [0.52, 0.82] 1 Sohn [18] 2012 34 8 25 28 95 0.81 [0.66, 0.91] 0.53 [0.30, 0.52]	Kim [15]	2009	52	6	4	35	100	0.85 [0.74, 0.93]	$0.90 \ [0.76, \ 0.97]$	0.180
Sohn [18] 2012 25 17 7 46 95 0.60 [0.43, 0.74] 0.87 [0.75, 0.87] 5 Sohn [18] 2012 29 13 9 44 95 0.69 [0.53, 0.82] 0.83 [0.70, 0.83] 1 Kim [15] 2009 58 3 12 27 100 0.95 [0.86, 0.99] 0.69 [0.52, 0.82] 0.69 [0.52, 0.82] 0.69 [0.52, 0.82] 0.69 [0.52, 0.82] 0.69 [0.52, 0.82] 0.53 [0.70, 0.92] <t< td=""><td>Kim [15]</td><td>2009</td><td>54</td><td>L</td><td>5</td><td>34</td><td>100</td><td>0.89 [0.78, 0.95]</td><td>0.87 [0.73, 0.96]</td><td>0.267</td></t<>	Kim [15]	2009	54	L	5	34	100	0.89 [0.78, 0.95]	0.87 [0.73, 0.96]	0.267
5 Sohn [18] 2012 29 13 9 44 95 0.69 [0.53, 0.82] 0.83 [0.70, 0.82] 1 Kim [15] 2009 58 3 12 27 100 0.95 [0.86, 0.99] 0.69 [0.52, 0.82] 0.53 [0.39, 0.82]	Sohn [18]	2012	25	17	7	46	95	0.60 [0.43, 0.74]	0.87 [0.75, 0.95]	
1 Kim [15] 2009 58 3 12 27 100 0.95 [0.86, 0.99] 0.69 [0.52, 0.90] Sohn [18] 2012 34 8 25 28 95 0.81 [0.66, 0.91] 0.53 [0.39, 0.39]	Sohn [18]	2012	29	13	6	44	95	$0.69 \ [0.53, 0.82]$	0.83 [0.70, 0.92]	0.353
Sohn [18] 2012 34 8 25 28 95 0.81 [0.66, 0.91] 0.53 [0.39, 0.30]	Kim [15]	2009	58	3	12	27	100	0.95 $[0.86, 0.99]$	$0.69 \ [0.52, \ 0.83]$	0.416
	Sohn [18]	2012	34	8	25	28	95	0.81 [0.66, 0.91]	0.53 $[0.39, 0.67]$	
Pooled 0.89 [0.82, 0.95] 0.60 [0.49,								0.89 $[0.82, 0.95]$	0.60 $[0.49, 0.70]$	

Deringer



Fig. 3 Forest plots of sensitivity and specificity of preoperative (a) and postoperative (b) studies according to cutoff values

high sensitivity and specificity in both preoperative and postoperative settings. This assay is now recommended by the both American [3] and European Guidelines [22]. In spite of those results, controversy remains on some issues.

Measuring Tg poses one of the major challenges to laboratories due to interfering factors that might alter the test results; the lack of methodological standards, inadequate functional sensitivity, and variability in the specificity of the commercially available antibody kits. All of these factors affect the accuracy of the results and are associated with large inter-assay variation resulting in difficulty of comparing between studies. Most of the studies included in this study used immunoradiometric technique except 2 studies (Table 1).

In this study, we selected studies with measuring FNA-Tg in the washout fluid rinsed with 1 ml of normal saline to minimize assay variables. Saline solution is the most widely accessible and least expensive option and is, therefore, the most commonly used [23]. The 1 ml volume of fluid to rinse FNA needles is recommended in the previous study [24]. The type and the volume of washout fluid in FNA-Tg measurement should be identical to decide LNs to be metastatic or not. As the methods used to measure FNA-Tg and serum Tg are the same, the possible interference of antithyroglobulin antibodies with FNA-Tg was taken into consideration until Boi et al. showed the absence of any significant interaction between the antithyroglobulin antibodies and FNA-Tg in patients with DTC [25]. The recent study evaluating 528 cases of FNA-Tg also showed

that antithyroglobulin antibodies was not correlated with FNA-Tg concentration, regardless of final diagnosis [26].

FNA-Tg may be higher if the patient has remnant thyroid gland because the thyroid is the main source of serum Tg [15]. Serum Tg is also a potential source of bias [25] because of blood contamination during aspiration. Our meta-analysis confirms that optimal cutoff value of FNA-Tg is much lower in postoperative than preoperative assessments as shown in the Fig. 2 (32.04 vs. 0.9 ng/ml). Absence of thyroid tissue after thyroidectomy and serum thyroid-stimulating hormone (TSH) suppression may play a role in lowering of both serum Tg and FNA-Tg. Recently, one study [26] suggested that serum TSH suppression and serum Tg presence should be considered in diagnosing LN malignancy with FNA-Tg in papillary thyroid cancer. Serum TSH suppression and serum Tg presence independently affected the diagnosis made by FNA-Tg [26]. Falsenegative FNA-Tg diagnosis can be done in the suppressed serum TSH level and false-positive FNA-Tg diagnosis in serum Tg presence [26]. Therefore, measuring FNA-Tg after TSH stimulation is recommended [26]. However, this has not been investigated in previous studies and needs further investigation. Moreover, the recent article by Torres et al. concluded the diagnostic performance of FNA-Tg is not affected by the presence of thyroid tissue [23]. In this study, the usefulness of FNA-Tg is confirmed in both preoperative and postoperative assessments, while the diagnostic performance showed difference between preoperative and postoperative studies.

Table 3 Pooled sensitivity and specificity according to cutoff values (Postop)

Cutoff (ng/ml)	Author	Year	TP	FP	FN	TN	Sample size	Sensitivity [95 % CI]	Specificity [95 % CI]	Distance from the point (0,1)
100	Kim [15]	2009	46	0	12	10	68	0.79 [0.67, 0.89]	1.00 [0.69, 1.00]	0.207
32.04	Kim [15]	2009	53	0	5	10	68	0.91 [0.81, 0.97]	1.00 [0.69, 1.00]	0.086
10	Kim [15]	2009	54	0	4	10	68	0.93 [0.83, 0.98]	1.00 [0.69, 1.00]	0.064
	Zanella [16]	2010	5	0	0	38	43	1.00 [0.48, 1.00]	1.00 [0.91, 1.00]	
	Suh [20]	2013	34	2	0	11	47	1.00 [0.90, 1.00]	0.85 [0.55, 0.98]	
	Jeon [19]	2013	99	7	6	151	263	0.94 [0.88, 0.98]	0.96 [0.91, 0.98]	
1.1	Frasoldati [13]	1999	10	1	1	1	13	0.91 [0.59, 1.00]	0.50 [0.01, 0.99]	0.026
	Giovanella [17]	2011	86	0	0	40	126	1.00 [0.96, 1.00]	1.00 [0.91, 1.00]	
1	Kim [15]	2009	56	1	2	9	68	0.97 [0.88, 1.00]	0.90 [0.55, 1.00]	0.218
	Suh [20]	2013	34	4	0	9	47	1.00 [0.90, 1.00]	0.69 [0.39, 0.91]	
0.9	Cunha [14]	2007	20	0	0	45	65	1.00 [0.83, 1.00]	1.00 [0.92, 1.00]	0
Pooled								1.00 [0.83, 1.00]	1.00 [0.92, 1.00]	

TP indicates true positive, FN false-negative, FP false-positive, TN true negative

A wide variety of diagnostic cutoff values ranging from 0.9 to 100.0 ng/ml have been described in the literature (Table 1). Our analysis investigated the best cutoff value in preoperative and postoperative studies, separately. The distance was minimal when cutoffs of 32.04 ng/ml for preoperative and of 0.9 ng/ml for postoperative studies (Tables 2, 3). Up to 30 % of the patients with papillary thyroid cancer exhibit recurrence or persistent metastasis to the neck LN [27]. The clinical evaluation of enlarged local LN is difficult at the beginning and throughout the followup, because inflammatory lymphadenopathies are extremely frequent. Furthermore, neck LN metastases from a multiplicity of extrathyroidal malignancies are a relatively common finding [25]. Although the specificity of serum Tg to detect recurrent disease is good, it is not the ideal measurement for metastases affecting small LN [28]. In addition, it is not able to establish the location of the neoplastic focus and is highly affected by the presence of serum antithyroglobulin antibodies, which are found in 25-30 % of patients with DTC [29].

Another meta-analysis to estimate the diagnostic accuracy of FNA-Tg technique was recently published [30]. Totally, 24 studies with diverse assays and methods were included, which resulted in significant heterogeneity. They did not state the clear cutoff values because their results are affected by a significant heterogeneity among the selected studies. In this study, we included the studies with FNA-Tg measured in the washout fluid rinsed with 1 ml of normal saline. Of course, our study has several limitations. First, as we excluded several studies with different methods and assays, a small number of studies were eligible for this study. In addition, some studies presented the results of several cutoff values of FNA-Tg [15, 16, 26]. Further

studies with individual patients' data will be needed to propose the standardized cutoff values. Second, although authors independently reviewed the primary studies, complete accuracy of data could not be ensured by the strategy.

FNA-C from LN has been widely used to confirm suspicious findings on ultrasonography, mainly in patients with undetectable serum Tg due to suppressed TSH or with a noniodine-concentrating metastasis [31, 32]. Nevertheless, the sensitivity of this method is far from ideal, as it varies from 75 to 85 %, with a rate of false-negative results of 6-8 % [8] and a rate of up to 20 % of non-representative samples or samples with inadequate cellularity dependent on cytopathologists' skill and experience [33]. Our cutoff value can be helpful for screening LN metastases in patients subjected to thyroidectomy, as well as to perform LN staging in patients with DTC who have not yet undergone initial surgery.

In conclusion, FNA-Tg can be used for screening LN metastases in patients subjected to thyroidectomy, as well as to perform LN staging in patients with DTC who have not yet undergone initial surgery. This measurement should be routinely performed, added to FNA-C, since the latter is often non-diagnostic, especially when LN are small-sized. Although the cutoff values for the FNA-Tg has not been standardized, preoperative values of 32.04 ng/ml and postoperative values of 0.9 ng/ml are recommended for identifying neck LN metastasis

Acknowledgments This work was supported by the National Research Foundation of Korea (NRF) Grant for the Global Core Research Center (GCRC) funded by the Korea government (Ministry of Science, ICT & Future Planning:MSIP) (No. 2011-0030680).

Conflict of interest Nothing to declare.

References

- G. Pellegriti, F. Frasca, C. Regalbuto, S. Squatrito, R. Vigneri, Worldwide increasing incidence of thyroid cancer: update on epidemiology and risk factors. J. Cancer Epidemiol. 2013, 965212 (2013). doi:10.1155/2013/965212
- S.S. Kweon, M.H. Shin, I.J. Chung, Y.J. Kim, J.S. Choi, Thyroid cancer is the most common cancer in women, based on the data from population-based cancer registries, South Korea. Jpn. J. Clin. Oncol. 43(10), 1039–1046 (2013). doi:10.1093/jjco/hyt102
- American Thyroid Association Guidelines Taskforce on Thyroid, Differentiated Thyroid, D.S. Cooper, G.M. Doherty, B.R. Haugen, R.T. Kloos, S.L. Lee, S.J. Mandel, E.L. Mazzaferri, B. McIver, F. Pacini, M. Schlumberger, S.I. Sherman, D.L. Steward, R.M. Tuttle, Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer. Thyroid 19(11), 1167–1214 (2009). doi:10.1089/ thy.2009.0110
- R. Gallicchio, S. Giacomobono, D. Capacchione, A. Nardelli, F. Barbato, A. Nappi, T. Pellegrino, G. Storto, Should patients with remnants from thyroid microcarcinoma really not be treated with iodine-131 ablation? Endocrine 44(2), 426–433 (2013). doi:10. 1007/s12020-013-9935-9
- A. Pisanu, A. Saba, M. Podda, I. Reccia, A. Uccheddu, Nodal metastasis and recurrence in papillary thyroid microcarcinoma. Endocrine (2014). doi:10.1007/s12020-014-0350-7
- A. Frasoldati, R. Valcavi, Challenges in neck ultrasonography: lymphadenopathy and parathyroid glands. Endocr. Pract. 10(3), 261–268 (2004). doi:10.4158/EP.10.3.261
- T.L. Hall, L.J. Layfield, A. Philippe, D.L. Rosenthal, Sources of diagnostic error in fine needle aspiration of the thyroid. Cancer 63(4), 718–725 (1989)
- F. Pacini, L. Fugazzola, F. Lippi, C. Ceccarelli, R. Centoni, P. Miccoli, R. Elisei, A. Pinchera, Detection of thyroglobulin in fine needle aspirates of nonthyroidal neck masses: a clue to the diagnosis of metastatic differentiated thyroid cancer. J. Clin. Endocrinol. metab. 74(6), 1401–1404 (1992)
- L. Giovanella, M. Bongiovanni, P. Trimboli, Diagnostic value of thyroglobulin assay in cervical lymph node fine-needle aspirations for metastatic differentiated thyroid cancer. Curr. Opin. Oncol. 25(1), 6–13 (2013). doi:10.1097/CCO.0b013e32835a9ab1
- P.F. Whiting, A.W. Rutjes, M.E. Westwood, S. Mallett, J.J. Deeks, J.B. Reitsma, M.M. Leeflang, J.A. Sterne, P.M. Bossuyt, Group, Q.-. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann. Intern. Med. 155(8), 529–536 (2011). doi:10.7326/0003-4819-155-8-201110180-00009
- J.P. Higgins, S.G. Thompson, J.J. Deeks, D.G. Altman, Measuring inconsistency in meta-analyses. BMJ **327**(7414), 557–560 (2003). doi:10.1136/bmj.327.7414.557
- L.E. Moses, D. Shapiro, B. Littenberg, Combining independent studies of a diagnostic test into a summary ROC curve: dataanalytic approaches and some additional considerations. Stat. Med. **12**(14), 1293–1316 (1993)
- A. Frasoldati, E. Toschi, M. Zini, M. Flora, A. Caroggio, C. Dotti, R. Valcavi, Role of thyroglobulin measurement in fine-needle aspiration biopsies of cervical lymph nodes in patients with differentiated thyroid cancer. Thyroid 9(2), 105–111 (1999)
- N. Cunha, F. Rodrigues, F. Curado, O. Ilheu, C. Cruz, P. Naidenov, M.J. Rascao, J. Ganho, I. Gomes, H. Pereira, O. Real, P. Figueiredo, B. Campos, F. Valido, Thyroglobulin detection in fine-needle aspirates of cervical lymph nodes: a technique for the diagnosis of metastatic differentiated thyroid cancer. Eur. J. Endocrinol. 157(1), 101–107 (2007). doi:10.1530/EJE-07-0088

- M.J. Kim, E.K. Kim, B.M. Kim, J.Y. Kwak, E.J. Lee, C.S. Park, W.Y. Cheong, K.H. Nam, Thyroglobulin measurement in fineneedle aspirate washouts: the criteria for neck node dissection for patients with thyroid cancer. Clin. Endocrinol. **70**(1), 145–151 (2009). doi:10.1111/j.1365-2265.2008.03297.x
- A.B. Zanella, E.L. Meyer, L. Balzan, A.C. Silva, J. Camargo, A. Migliavacca, J.R. Guimaraes, A.L. Maia, Thyroglobulin measurements in washout of fine needle aspirates in cervical lymph nodes for detection of papillary thyroid cancer metastases. Arq. Bras. Endocrinol. Metabol. 54(6), 550–554 (2010)
- L. Giovanella, L. Ceriani, S. Suriano, Lymph node thyroglobulin measurement in diagnosis of neck metastases of differentiated thyroid carcinoma. J. Thyroid Res. 2011, 621839 (2011). doi:10. 4061/2011/621839
- Y.M. Sohn, M.J. Kim, E.K. Kim, J.Y. Kwak, Diagnostic performance of thyroglobulin value in indeterminate range in fine needle aspiration washout fluid from lymph nodes of thyroid cancer. Yonsei Med. J. 53(1), 126–131 (2012). doi:10.3349/ymj. 2012.53.1.126
- M.J. Jeon, J.W. Park, J.M. Han, J.H. Yim, D.E. Song, G. Gong, T.Y. Kim, J.H. Baek, J.H. Lee, Y.K. Shong, W.B. Kim, Serum antithyroglobulin antibodies interfere with thyroglobulin detection in fine-needle aspirates of metastatic neck nodes in papillary thyroid carcinoma. J. Clin. Endocrinol. Metabol. 98(1), 153–160 (2013). doi:10.1210/jc.2012-2369
- Y.J. Suh, E.J. Son, H.J. Moon, E.K. Kim, K.H. Han, J.Y. Kwak, Utility of thyroglobulin measurements in fine-needle aspirates of space occupying lesions in the thyroid bed after thyroid cancer operations. Thyroid 23(3), 280–288 (2013). doi:10.1089/thy. 2011.0303
- E. Sigstad, A. Heilo, E. Paus, K. Holgersen, K.K. Groholt, L.H. Jorgensen, T.V. Bogsrud, A. Berner, T. Bjoro, The usefulness of detecting thyroglobulin in fine-needle aspirates from patients with neck lesions using a sensitive thyroglobulin assay. Diagn. Cytopathol. 35(12), 761–767 (2007). doi:10.1002/dc.20726
- 22. L. Leenhardt, M.F. Erdogan, L. Hegedus, S.J. Mandel, R. Paschke, T. Rago, G. Russ, 2013 European thyroid association guidelines for cervical ultrasound scan and ultrasound-guided techniques in the postoperative management of patients with thyroid cancer. Eur. Thyroid J. 2(3), 147–159 (2013). doi:10. 1159/000354537
- M.R. Torres, S.H. Nobrega Neto, R.J. Rosas, A.L. Martins, A.L. Ramos, T.R. da Cruz, Thyroglobulin in the washout fluid of lymph-node biopsy: what is its role in the follow-up of differentiated thyroid carcinoma? Thyroid 24(1), 7–18 (2014). doi:10. 1089/thy.2013.0244
- 24. Leenhardt, L., Borson-Chazot, F., Calzada, M., Carnaille, B., Charrie, A., Cochand-Priollet, B., Do Cao, C., Leboulleux, S., Le Clech, G., Mansour, G., Menegaux, F., Monpeyssen, H., Orgiazzi, J., Rouxel, A., Sadoul, J.L., Schlumberger, M., Tramalloni, J., Tranquart, F., Wemeau, J.L., Societe Francaise, d.E., Groupe de Recherche sur la, T., Societe Francaise de, R., Societe Francaise de Medecine, N., Association Francophone de Chirurgie, E., Societe Francaise de Cytologie, C., Groupe de Biologie Specialise de la Societe Francaise de Medecine, N., Reseau National des Tumeurs, T.: [Best practices guideline for the use of neck ultrasonography and echo-guided techniques in the management of differentiated thyroid cancers of the follicular type. Societe Francaise d'Endocrinologie.Groupe de Recherche sur la Thyroide]. Annales d'endocrinologie **72**(4 Suppl 1), H1-26 (2011). doi:10.1016/S0003-4266(11)70001-4
- 25. F. Boi, G. Baghino, F. Atzeni, M.L. Lai, G. Faa, S. Mariotti, The diagnostic value for differentiated thyroid carcinoma metastases of thyroglobulin (Tg) measurement in washout fluid from fine-needle aspiration biopsy of neck lymph nodes is maintained in the

presence of circulating anti-Tg antibodies. J. Clin. Endocrinol. Metabol. **91**(4), 1364–1369 (2006). doi:10.1210/jc.2005-1705

- 26. J.H. Moon, Y.I. Kim, J.A. Lim, H.S. Choi, S.W. Cho, K.W. Kim, H.J. Park, J.C. Paeng, Y.J. Park, K.H. Yi, Park do, J., Kim, S.E., Chung, J.K.: Thyroglobulin in washout fluid from lymph node fine-needle aspiration biopsy in papillary thyroid cancer: largescale validation of the cutoff value to determine malignancy and evaluation of discrepant results. J. Clin. Endocrinol. Metabol. **98**(3), 1061–1068 (2013). doi:10.1210/jc.2012-3291
- K.G. Schuff, Management of recurrent/persistent papillary thyroid carcinoma: efficacy of the surgical option. J. Clin. Endocrinol. Metabol. 96(7), 2038–2039 (2011). doi:10.1210/jc.2011-1663
- H. Kim, S.J. Kim, I.J. Kim, K. Kim, S. Kim, B.H. Kim, S.S. Kim, J.Y. Kyung, Limited clinical value of periablative changes of serum markers in the prediction of biochemical remission in patients with papillary thyroid cancer. Nucl. Med. Mol. Imaging 47(4), 268–272 (2013). doi:10.1007/s13139-013-0220-x
- C.A. Spencer, Challenges of serum thyroglobulin (Tg) measurement in the presence of Tg autoantibodies. J. Clin. Endocrinol. Metabol. 89(8), 3702–3704 (2004). doi:10.1210/jc.2004-0986

- G. Grani, A. Fumarola, Thyroglobulin in lymph node fine-needle aspiration washout: a systematic review and meta-analysis of diagnostic accuracy. J. Clin. Endocrinol. Metabol. 99(6), 1970–1982 (2014). doi:10.1210/jc.2014-1098
- G. Costante, S. Filetti, Diagnosis: thyroglobulin in fine-needle aspirates-a clue to metastasis? Nat Rev. Endocrinol. 5(5), 249–250 (2009). doi:10.1038/nrendo.2009.55
- 32. D.W. Kim, S.J. Jeon, C.G. Kim, Usefulness of thyroglobulin measurement in needle washouts of fine-needle aspiration biopsy for the diagnosis of cervical lymph node metastases from papillary thyroid cancer before thyroidectomy. Endocrine 42(2), 399–403 (2012). doi:10.1007/s12020-012-9636-9
- 33. B.D. Florentine, B. Staymates, M. Rabadi, J. Barstis, A. Black, Cancer Committee of the Henry Mayo Newhall Memorial, H, The reliability of fine-needle aspiration biopsy as the initial diagnostic procedure for palpable masses: a 4-year experience of 730 patients from a community hospital-based outpatient aspiration biopsy clinic. Cancer 107(2), 406–416 (2006). doi:10. 1002/cncr.21976