Cost-Effectiveness Analysis of Well-Differentiated Thyroid Carcinoma Surveillance Using Nuclear Medicine Procedures

John Kenneth V. Gacula, RN, MD[1], Jerry M. Obaldo, MD, MHA[1], Vio Jianu C. Mojica, MS[2]

[1]Division of Nuclear Medicine, Department of Medicine, Philippine General Hospital, University of the Philippines Manila
[2]Department of Physical Sciences and Mathematics, College of Arts and Sciences, University of the Philippines Manila
E-mail address: jvgacula@up.edu.ph, jerryobaldo@yahoo.com, vcmojica@up.edu.ph

ABSTRACT

Introduction:

Well-differentiated thyroid carcinoma (WDTC) is the most common type of thyroid cancer with a notable increasing incidence worldwide. It is prevalent among Filipino descent as compared to other nationalities. Its good prognosis and high survival rate predispose patients to lifetime surveillance with incomplete response, instead of death, as outcome measure. This eventually leads to increase in cost of care, utilization, and allocation of medical resources for the survivors of the disease. Thyroglobulin immunoradiometric assay (Tg IRMA) and I-131 diagnostic whole-body scan (dWBS) are two nuclear medicine procedures that are part of WDTC surveillance. Due to their varied availability in Asia-Pacific, most clinicians measure thyroglobulin (Tg) alone due to perceived cost-effectiveness.

Objective:

This study aims to analyze the cost-effectiveness of two nuclear medicine procedures used in WDTC surveillance, namely thyroglobulin immunoradiometric assay and I-131 diagnostic whole-body scan, in detecting incomplete response.

Methodology:

Three clinical guidelines on WDTC management were reviewed to identify frequency, total number and expenditure for surveillance, namely from the University of the Philippines-Philippine General Hospital in 2008 (PGH 2008), American Thyroid Association in 2015 (ATA 2015), and the Department of Health (DOH 2021). A Markov model was constructed to simulate a 36-month surveillance with complete and incomplete response to treatment as disease states. Parameter values like rate of incomplete response in WDTC patients, prognostic values per each surveillance test, and other relevant data were collected from literature search and established data. The cost of surveillance was based on the rates offered by Philippine General Hospital (PGH) Radioisotope Laboratory as of November 2022. One-way sensitivity was done to check robustness of results.

Results:

ATA 2015 incurs the most expenses, amounting to PHP 14,600.00 to 20,450.00 ($ 254.19 – 356.04) for three years of surveillance, followed by DOH 2021 (PHP 11,700.00 – 15,600.00 or $ 203.74 – 271.65), and PGH 2008 (PHP 3,900.00 – 6,825.00 or $ 67.91 – 118.85). The thyroglobulin IRMA arm costs lower (PHP 17,784.00 or $ 309.74) than I-131 dWBS (PHP 271,875.00 or $ 4,735.13) in detecting incomplete response. I-131 dWBS should cost around PHP 570.00 (or $ 9.92) to be as cost-effective as the thyroglobulin IRMA.

Conclusion:

This study has identified that thyroglobulin IRMA is more cost-effective than I-131 diagnostic whole-body scan in detecting incomplete response in WDTC patients. This supports the perceived cost-effectiveness of thyroglobulin measurement in surveillance, even without diagnostic whole body-scans. This study also identified that the new DOH 2021 guidelines will incur lesser expenditure in using nuclear medicine procedures for surveillance as compared to ATA 2015 guidelines. Local clinicians may also find it easier to follow as it is more suitable to the Philippine setting.

Keywords: *cost-effectiveness, well-differentiated thyroid carcinoma, surveillance*

INTRODUCTION

Among the thyroid cancers, well-differentiated thyroid carcinoma (WDTC) is the most common type and accounts for 90% of cases. Even with such high prevalence, it has better overall prognosis as compared to other types [1,2]. It can be further categorized into papillary thyroid carcinoma (PTC), follicular thyroid carcinoma (FTC), and Hurthle cell carcinoma (HTC). PTC accounts for 80-85% of cases, making it the most predominant and most diagnosed type [1–3]. Its incidence rose steadily with 5/100,000 in mid-1990s to 15.0 in 2014 in United States. This increasing trend is associated with more accessible healthcare services among developed nations that allows early and widespread detection [4]. The situation of South Korea was acclaimed to be remarkable, yet controversial, as they recorded a 15-fold increase in WDTC incidence rate from 1993 to 2011 [2]. This coincides with several literature identifying Asian populations being highly affected by WDTC.

Among Asians, Filipinos were identified to be more predisposed to have WDTC. The overall thyroid cancer incidence rate from 1990 to 2014 was 19.57/100,000 person years for Filipinos, compared to 10.45/100,000 for non-Filipino Asians and 13.94/100,000 for non-Hispanic Whites. Similarly, a notable increase from 16.27/100,000 person-years in 1990 to 20.18/100,000 person-years in 2014 for Filipinos living in the United States was seen [5]. Locally, the national incidence rate for WDTC has not been established yet. Several local studies like Lo et al., identified 723 patients with WDTC in UP-Philippine General Hospital; majority are papillary (649, 89.8%) while the rest are follicular (79, 10.2%) [6].

The goal for each patient with WDTC is to achieve excellent response or disease-free status: the patient must have no clinical or imaging evidence of tumor and maintains a low serum thyroglobulin (Tg) levels during thyroid-stimulating hormone (TSH) suppression or after stimulation without interfering antibodies [7]. Given that WDTC has good prognosis, it has been considered that survival may not be the appropriate outcome measure but rather incomplete response [8]. Recurrence is also considered as an outcome measure, however Bates et al., have identified that most patients that underwent re-operations for “recurrent disease” never actually achieved a disease-free state, and 71 out of 92 (77%) re-operations were categorized as persistence [9].

Despite the good prognosis, there is still the presence of persistent disease among patients who have undergone surgery and received radioiodine (RAI) therapy. Tuttle, et al., evaluated the patient’s response to therapy and formulated a category with dynamic risk estimates for long-term surveillance [10]. This was adapted by the American Thyroid Association in 2015, highlighting that it can be applied at any point during the patient’s follow-up [7].

Among 90 patients that underwent thyroidectomy and RAI therapy in Makati Medical Center, Santiago, et al., identified 12 (or 13.33%) with biochemical incomplete response and 23 (or 25.56%) with structural incomplete response. Among other factors considered, the presence of positive Tg and anti-Tg postoperatively were strongly associated with incomplete response [11]. With 225 WDTC patients in University of Santo Tomas Hospital, Mendoza, et al., noted 69 (or 30.67%) had incomplete response. Biochemical incomplete response was seen in six (or 8.7%) patients, while structural incomplete response was identified in 63 (or 25%) patients. They have identified gender, lymph node involvement and location, extent of malignancy, and multifocality as factors with significant association with incomplete response [12]. Both studies register higher occurrence of structural incomplete response as compared to biochemical incomplete response.

Detection of circulating Tg in patients that underwent total thyroidectomy and RAI therapy would signify presence of thyroid tissue. Through the years, thyroglobulin immunoradiometric assay (Tg IRMA) have greatly improved and an international calibration standard was applied. This shifted the importance of serum Tg monitoring from adjunctive to essential part of thyroid cancer surveillance [13]. It has been recommended that serum Tg and anti-Tg antibodies be done longitudinally in the same laboratory and same assay [7]. Absence of persistent disease is defined by Tg of less than or equal to 1 ng/mL and less than or equal to 2 ng/mL for basal and simulated, respectively [14]. For high-risk patients, Tg measurements may be done more frequent [7].

Diagnostic whole-body scan (dWBS) using iodine radioisotopes was once considered a central part of thyroid cancer surveillance. Patient preparations are similar to that of RAI therapy, which include low-iodine diet for two weeks and withdrawal from replacement thyroxine therapy for four to six weeks to achieve hypothyroidism and elevated TSH serum levels of greater than 30 mU/L. Scans were performed 48 to 72 hours after administering 2-5 millicurie (mCi) of iodine-131

(I-131) [13]. Majority of countries in the Asia-Pacific region use I-131, except for Australia, Hong Kong, Korea, and Taiwan where iodine-123 (I-123) is used [15].

Yang et al., have identified that there is a wide difference of medical resources available and economic capabilities among the Asia-Pacific countries and this hinders both physicians and patients to adhere to guideline recommendations, more specifically in surveillance [16]. Most countries in the region rely only on TSH-stimulated Tg measurements, even without dWBS, citing cost-effectiveness and convenience [15]. However, there are limited studies to prove this.

In the US, based on the increasing incidence rate, the cost of thyroid cancer care was estimated to be $18-21 billion dollars in 2019 [4]. The highest cost in the course of the disease was noted to be in the initial diagnosis and treatment, amounting to $658 million or 41% of total cost. Further expenditure during the continuing or monitoring phase amounts to $595 million or 37% of the total cost [17]. Surveillance-related costs are higher immediately post-operatively [18]. In Brazil, there was an observed 120% increase in treatment and follow-up related-procedures for thyroid cancer. Increasing trend in the procedures per 100,000 people for serum Tg and I-131 whole body scan from 2008 to 2015 was observed by Janovsky, et al., and is likely due to overdiagnosis of thyroid cancer cases [19]. The increasing incidence of patients with WDTC would further lead to the rise of patients for surveillance, with further increase in utilization and allocation of medical resources [18]. This trend has been visible not only in the Philippines, but also with other neighbor nations like Hong Kong and Korea, with having more than 10 cases on follow-up per year [16]. Survivors of the disease are expected to have repeated treatment, lifelong surveillance, and adjustments to thyroid hormone replacement that contribute to the physical, psychological, and financial costs of diagnostics and treatment [8].

Cost-effectiveness analysis (CEA) is used for program evaluation, specifically linking the costs to its benefits or effectiveness. It is used to compare set of programs and determine which provides greater outcome for the costs or costs achieved by the effectiveness unit [20,21]. Its advantage, as stated by Rudmik and Drummond, is for easier understanding and interpretation of clinicians as it uses familiar clinical endpoints or outcomes [22]. It could provide information for the nuclear medicine physicians on why a specific nuclear medicine technology is important in clinical management of patients, or how effective (or not) they are in achieving a certain benefit, and substantiating evidences to hospital administrators, insurance companies, and important bodies for allocating greater resources for certain study or procedure [21].

Markov model is one of the models used in cost-effectiveness studies in healthcare. It utilizes random processes and multiple possible consequences that occur over a long period of time, making highly suitable for chronic diseases. Distinct disease states and transition probabilities, cost estimates for resource use and health outcomes or events are identified and placed in a "Markov cycle" to run through several cycles to stimulate long periods of the disease progression [23]. As each cycle is completed, the cost and effect for each health state are identified and the cost-effectiveness ratio for a modality arm is calculated by adding all the weighted costs of each individual cycle [22]. A limitation of this model is its inherent "memorylessness" or disregard to the effects from the previous cycles, as each cycle is considered identical [23–25]. Nevertheless, Markov model has proven to be useful and has been applied in several healthcare studies concerning screening programs, therapeutic interventions, and diagnostic technologies [23].

Objectives

The main objective of this study was to analyze the cost-effectiveness of two nuclear medicine procedures used in surveillance of WDTC, namely Tg IRMA and I-131 dWBS, in detecting incomplete response.

Specifically, this paper aimed:

A. To identify the schedule and total number of Tg IRMA and I-131 dWBS requested for a WDTC patient based on the different clinical practice guidelines
B. To calculate the total cost per each surveillance arm in a 36-month time frame based on the different clinical practice guidelines
C. To determine the cost per detection event of Tg IRMA surveillance alone
D. To determine the cost per detection event of I-131 dWBS surveillance alone; and
E. To compare the difference of Tg IRMA and I-131 dWBS surveillance arms in terms of detection of incomplete response.

METHODOLOGY

This study was submitted to and was approved by the University of the Philippines Manila Research Ethics

Board (UPMREB) with UPMREB Code: 2022-0577-EX.

Selection and Review of Clinical Guidelines

To identify the total number and create a schedule scheme of the Tg IRMA and I-131 dWBS for WDTC patient surveillance, we reviewed three clinical guidelines that are highly utilized by clinicians locally. These were:

- 2008 Clinical Practice Guidelines of the Philippine General Hospital for the Management of Thyroid Nodules and Well-differentiated Thyroid Carcinoma (by the PGH Working Group on Thyroid Cancer; PGH 2008)
- 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer (by the American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer; ATA 2015)
- 2021 Philippine Interim Clinical Practice Guidelines for the Diagnosis and Management of Well-Differentiated Thyroid Cancer (released by the Department of Health, as commissioned to the Dr. Jose R. Reyes Memorial Medical Center; DOH 2021)

The frequency for each test, number of tests in the first year and subsequent two years, total incurred cost in the first year and subsequent two years, and actual recommendations were noted and tabulated. All of the guidelines are available in the internet and were accessed publicly.

Markov Model Structure

To identify and analyze the cost-effectiveness of Tg IRMA and I-131 dWBS in detecting incomplete response, a Markov model was constructed to simulate the surveillance of patients that have been diagnosed with WDTC through histopathology, underwent total thyroidectomy, and received RAI ablation therapy. After diagnosis and treatment, the patients can enter the model via two states: (1) complete response, and (2) incomplete response to treatment, with the latter being the outcome of interest. A diagram to visualize the model is shown on Figure 1.

The two states were based on the definition in the 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer. Complete response is defined as no clinical, biochemical, or structural evidence of disease. Meanwhile, incomplete response is defined as:

- biochemical: suppressed Tg > 1 ng/mL or stimulated Tg > 10 ng/mL or rising anti-Tg antibody levels; and
- structural: structural or functional evidence of disease, with any TG level, with or without anti-Tg antibodies [7].

A patient registering in as incomplete response leaves the model as it is assumed that they will undergo another bout of RAI therapy. On the other hand, patients who register as complete response will continuously undergo surveillance until an incomplete response is detected or until the end of the time horizon. Simulation was done utilizing a 36-month time frame as majority of the

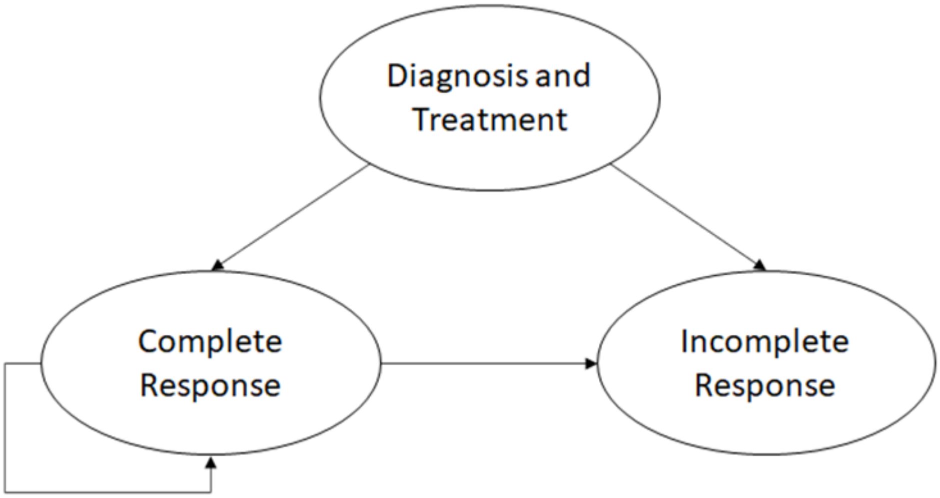

FIGURE 1. Diagram of the Markov model applied in this study

incomplete responses occur within the first two to three years of diagnosis, and majority of the costs allotted for surveillance are incurred immediately after surgery [18]. Additionally, this was also based on the most common recommended interval for the said tests across the three clinical guidelines reviewed.

Data Selection and Parameter Inputs

Parameter values like rate of incomplete response in WDTC patients, prognostic values per each surveillance tests, and other relevant data to create assumptions and transition probabilities were gathered from established data and literature search using several databases, primarily PubMed, Scopus, ScienceDirect, and JSTOR. For the cohort characteristics and epidemiological data, even with the vast presence of international data, the authors highly preferred local data (if available) to reflect adequately the local situation.

The subjects in the selected literature or sources should fulfill the criteria set by the authors. The inclusion and exclusion criteria are as follows:

Inclusion criteria was based on the definition of incomplete response, either it be:

- biochemical: suppressed Tg > 1 ng/mL or stimulated Tg > 10 ng/mL or rising anti-Tg antibody levels
- structural: structural or functional evidence of disease, with any Tg level, with or without anti-Tg antibodies

As for the cohort and literature to be used, the following were considered:

- patients diagnosed with WDTC based on histopathology
- patients underwent total thyroidectomy and RAI ablation therapy
- for Tg surveillance arm: immunoradiometric assay was specifically utilized
- for I-131 dWBS: radioiodine with low activity was utilized, with or without SPECT

Exclusion was based on the following:

- diagnosed with non-differentiated thyroid carcinoma or no clear histopathologic diagnosis
- did not undergo surgery or had less than total thyroidectomy for surgery
- did not receive RAI ablation therapy
- TG was measured using other technology (e.g., ECLIA)
- if with excellent response, had less than three years of surveillance

With the alarming findings in the study of Bates et al., “recurrence” was not considered in this study [9].

Transition probabilities for the Tg IRMA and I-131 dWBS surveillance arms were obtained from Giovanella et al. (2002) [26] and Schlumberger et al. (2007) [27] respectively. The data from the two studies reflect the inherent rates of registering incomplete responses for each arm. One main assumption of the model is that, at every test, the probability of registering an incomplete response remains constant.

Cost items consist solely of the unit cost per test. They were based on the current fee (as of November 2022) offered by the Radioisotope Laboratory of the UP-Philippine General Hospital, where the study was conducted. These rates mirror the healthcare expenditure of the hospital and patients. All input parameters are summarized in Table 1.

TABLE 1. Parameter values (Transition Probabilities and Cost Items) used in the Markov model and their respective values and sources

Parameter	Value	Source
Transition Probabilities		
Incomplete Response Rate (Tg IRMA)	40.5%	Baudin, et al. (2002) [28]
Positive Predictive Value (Tg IRMA)	53.3%	
Incomplete Response Rate (I-131 dWBS)	3.2%	Schlumberger, et al. (2007) [27]
Positive Predictive Value (I-131 dWBS)	35.7%	
Cost Items	**Value (PHP)**	**Source**
Tg IRMA (Service)	1560.00	UP-PGH Radioisotope Laboratory
Tg IRMA (Pay)	1870.00	
I-131 dWBS (Service)	8750.00	
I-131 dWBS (Pay)	9545.00	

The final costs derived from the Markov model were totaled up in Philippine Peso (PHP) and were converted to United States Dollars (USD) based on the currency conversion rate at the last month of the study (November 2022). A one-way sensitivity analysis was done to check the robustness of the results.

RESULTS

Costs

Cost-related parameters include the unit costs of the following nuclear medicine services: Tg IRMA, anti-Tg RIA, and I-131 dWBS. Costs were based from the service and pay rates offered by the UP-PGH Radioisotope Laboratory, as seen in Table 2.

Both rates offered for the service and pay outpatients were considered and applied in the model. Most service patients in UP-Philippine General Hospital receive full coverage for their diagnostic expenses. This is made possible with help of medical social services or other subsidized means. The service rate is a reflection of hospital expenditure on patients. The pay rate, on the other hand, reflects the out-of-pocket expenditure of patients as these diagnostic tests cannot be reimbursed through health insurance. The conversion rate applied is 1 USD = 57.43 PHP (as per November 2022).

Clinical Guidelines

First Year of Surveillance

The recommended schedule per each diagnostic test, total number of tests, total expenditure, and source per each clinical guideline for the first year of surveillance are tabulated in Table 3. With Tg IRMA and anti-Tg measurements, the DOH 2021 guidelines incurred greater than the costs of ATA 2015 and PGH 2008 guidelines. When totaled, ATA 2015 recommendations will incur the greatest cost as compared with the other two local guidelines.

Subsequent Years of Surveillance

The recommended schedule per each diagnostic test, total number of tests, total expenditure, and source per each clinical guideline for the subsequent two years of follow-up are tabulated in Table 4.

Using Tg and anti-Tg measurements, the DOH 2021 and ATA 2015 guidelines incurred similar total amount, while PGH 2008 garnered less. No recommendations were made on I-131 dWBS as part of surveillance for the subsequent years.

The total expenditure (in PHP and USD) for the three guidelines is shown in Table 5 and Figure 2. ATA 2015 incurs the most expenses for three years of surveillance. This is followed by DOH 2021 and PGH 2008.

In the first year of surveillance, percentage of expenditure is larger in ATA 2015 with 62%, followed by DOH 2021 (50%) and PGH 2008 (43%). As for the subsequent two-year surveillance, PGH 2008 incurred more expenses with 57%, followed by DOH 2021 (50%) and ATA 2015 (38%). This is shown in Figure 3.

Base-Case Results

Patients with histopathologically diagnosed WDTC, underwent total thyroidectomy, and received RAI ablation therapy were part of the cohort. The model ran for 6 cycles (1 cycle = 6 months), with a total of 36 months of surveillance.

The Tg IRMA arm has shown to dominate the I-131 dWBS arm as it costs lower while detecting incomplete response earlier and more accurately. The results of the base case are summarized in Table 6.

One-Way Sensitivity Analysis

Using Tg IRMA surveillance arm as reference, the parameters within the I-131 dWBS arm were varied separately .

TABLE 2. Service and Pay Rates of PGH Radioisotope Laboratory (as of November 2022)

Nuclear Medicine Test	Service Rate		Pay Rate	
	(in PhP)	(in USD)	(in PhP)	(in USD)
Thyroglobulin (Tg IRMA)	975.00	16.98	1,175.00	20.46
Anti-thyroglobulin (Anti-Tg)	975.00	16.98	1,165.00	20.28
I-131 Diagnostic Whole-Body Scan (I-131 dWBS)	8,750.00	152.36	9,545.00	166.19

TABLE 3. Interval, total number of tests, total cost, and source of thyroglobulin, anti-thyroglobulin, and I-131 diagnostic whole-body scan per clinical guideline for the first year of surveillance

FIRST YEAR OF SURVEILLANCE				
Nuclear Medicine Test	Interval	No of Tests	Cost (in PHP)	Source
2008 Clinical Practice Guidelines of the Philippine General Hospital for the Management of Thyroid Nodules and Well-differentiated Thyroid Carcinoma (PGH, 2008) [14]				
Thyroglobulin	Every 6-12 months	1-2	975.00-1,950.00	Section VII, Consensus A
Anti-thyroglobulin	At least once	1	975.00	Section VII, Consensus A
			1,950.00 - 2,925.00	
I-131 Diagnostic Whole-Body Scan	-	-	**-**	-
2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer (ATA, 2015) [7]				
Thyroglobulin	Every 6-12 months	1-2	975.00-1,950.00	Recommendation 62B
Anti-thyroglobulin	Every 6-12 months	1-2	975.00-1,950.00	Recommendation 62B
			1950.00-3,900.00	
I-131 Diagnostic Whole-Body Scan	At 6-12 months	1	**8,750.00**	Recommendation 67
2021 Philippine Interim Clinical Practice Guidelines for the Diagnosis and Management of Well-Differentiated Thyroid Cancer (DOH, 2021) [29]				
Thyroglobulin	Every 3-6 months	2-4	1,950.00-3,900.00	5.2C
Anti-thyroglobulin	Every 3-6 months	2-4	1,950.00-3,900.00	5.2C
			3,900.00-7,800.00	
I-131 Diagnostic Whole-Body Scan	-	-	**-**	-

TABLE 4. Interval, total number of tests, total cost, and source of thyroglobulin, anti-thyroglobulin, and I-131 diagnostic whole-body scan per clinical guideline for the subsequent two years of surveillance

SUBSEQUENT TWO YEARS OF SURVEILLANCE				
Nuclear Medicine Test	**Interval**	**No of Tests**	**Cost (in PHP)**	**Source**
2008 Clinical Practice Guidelines of the Philippine General Hospital for the Management of Thyroid Nodules and Well-differentiated Thyroid Carcinoma (PGH, 2008) [14]				
Thyroglobulin	Every 6-12 months	2-4	1,950.00-3,900.00	Section VII, Consensus A
Anti-thyroglobulin	-	-	-	-
I-131 Diagnostic Whole-Body Scan	-	-	**-**	-
2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer (ATA, 2015) [7]				
Thyroglobulin	Every 6-12 months	2-4	1,950.00-3,900.00	Recommendation 62E
Anti-thyroglobulin	Every 6-12 months	2-4	1,950.00-3,900.00	Recommendation 62E
			3,900.00-7,800.00	
I-131 Diagnostic Whole-Body Scan	-	-	**-**	-
2021 Philippine Interim Clinical Practice Guidelines for the Diagnosis and Management of Well-Differentiated Thyroid Cancer (DOH, 2021) [29]				
Thyroglobulin	Every 6 months	4	3,900.00	5.3C
Anti-thyroglobulin	Every 6 months	4	3,900.00	5.3C
			7,800.00	
I-131 Diagnostic Whole-Body Scan	-	-	**-**	-

TABLE 5. Total expenditure in the first and subsequent two years of surveillance per each guideline

	Cost in First Year	Cost in Subsequent Two Years	Total Cost
ATA 2015	PHP 10,700.00 - 12,650.00 ($ 186.34 - 222.30)	PHP 3,900.00 - 7,800.00 ($ 67.92 – 135.84)	PHP 14,600.00 - 20,450.00 ($ 254.27 – 356.15)
PGH 2008	PHP 1,950.00 - 2,925.00 ($ 33.96 – 50.94)	PHP 1,950.00 - 3,900.00 ($ 33.96 – 67.92)	PHP 3,900.00 – 6,825.00 ($ 67.91 – 118.86)
DOH 2021	PHP 3,900.00 - 7,800.00 ($ 67.92 – 135.84)	PHP 7,800.00 ($ 135.84)	PHP 11,700.00 – 15,600.00 ($ 203.78 – 271.70)

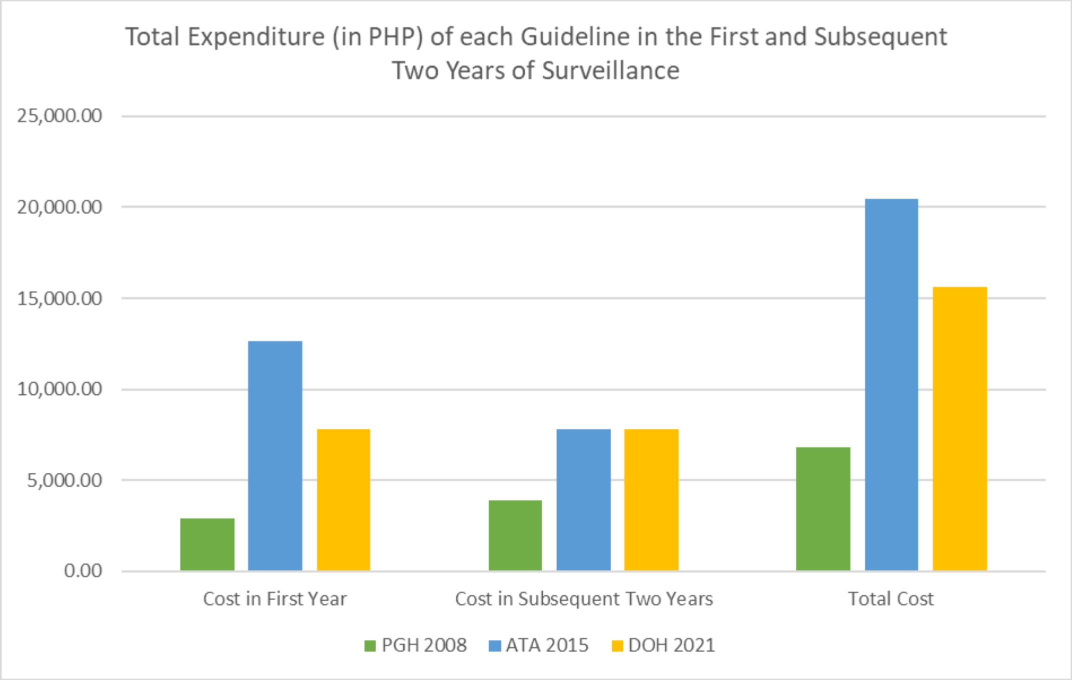

FIGURE 2. Graph comparing the total expenditure of each guideline in the first and subsequent two years of surveillance

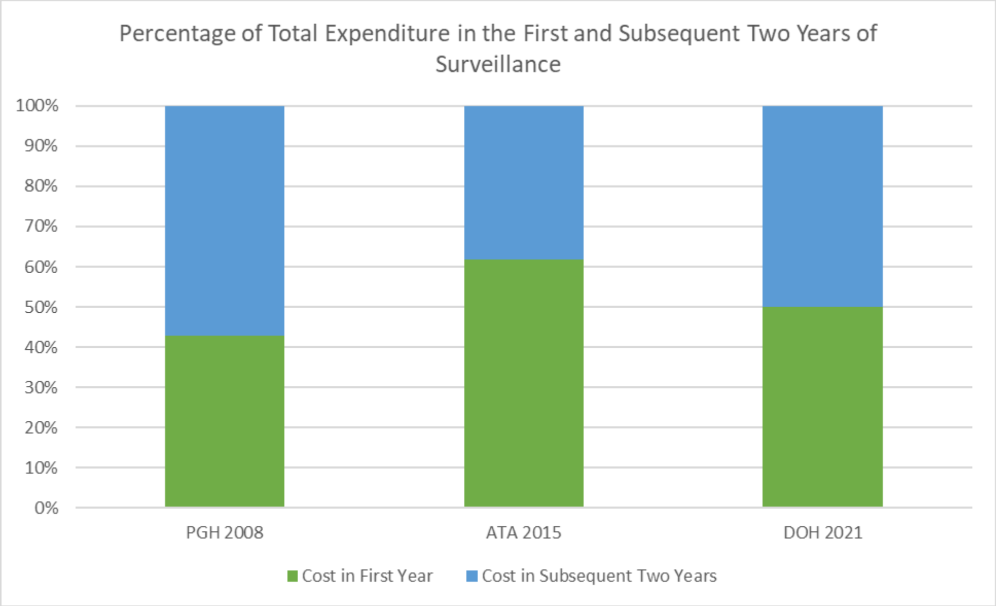

FIGURE 3. Graph comparing the division of total expenditure of each guideline in the first and subsequent two years of surveillance

TABLE 6. Base case analysis results.

	Thyroglobulin IRMA		I-131 dWBS	
	Service (in PHP)	**Pay (in PHP)**	**Service (in PHP)**	**Pay (in PHP)**
Cost per Incomplete Response Result	3,848.00	4,612.67	271,875.00	296,576.79
Cost per Correct Detection	7,215.00	8,648.75	761,250.00	830,415.00

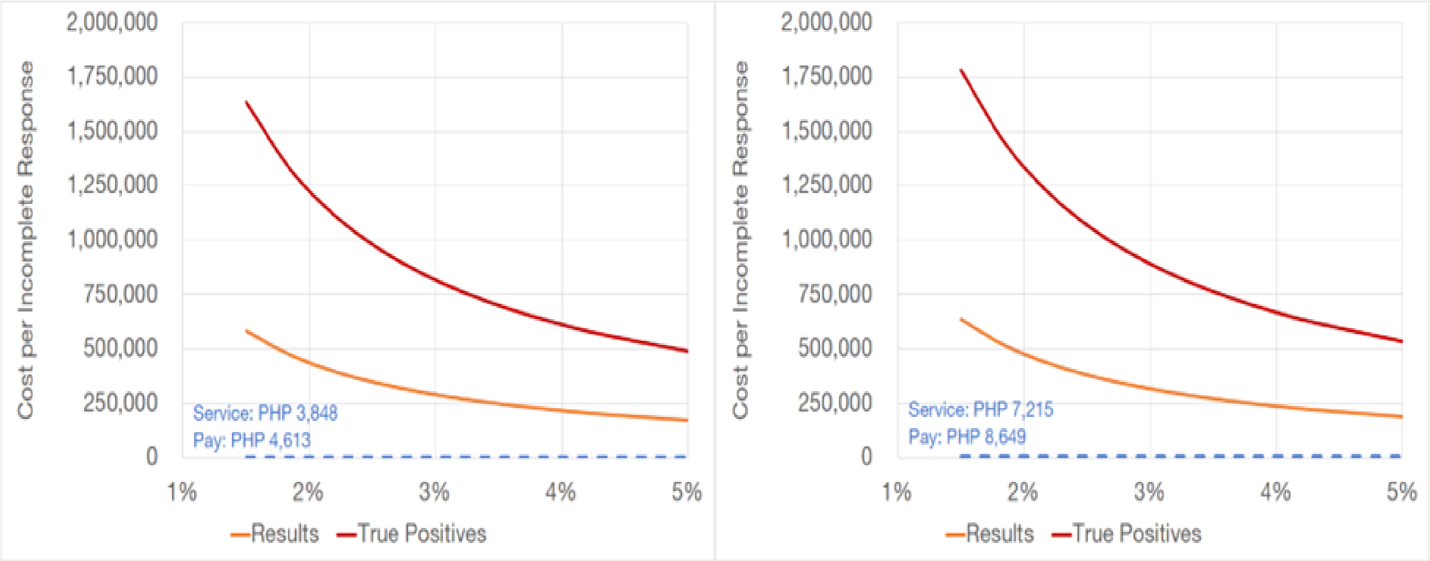

FIGURE 4. Costs per incomplete response result and successful detection for the I-131 dWBS arm over varying probabilities of positive result using service (left) and pay (right) rates.

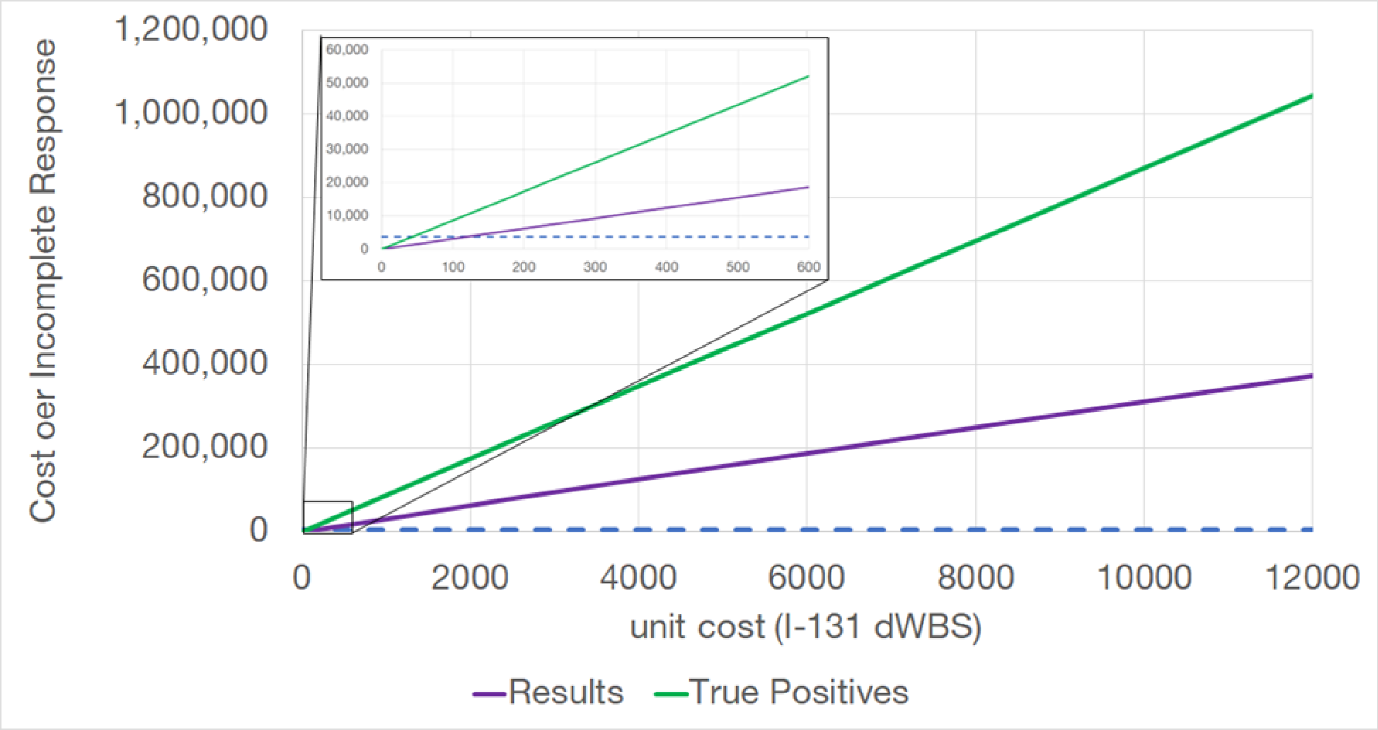

FIGURE 5. Costs per incomplete response result (green) and successful detection (purple) for the I-131 dWBS arm over varying unit cost per test; costs where I-131 dWBS will be as cost-effective as thyroglobulin IRMA is displayed (inset).

Probabilities of obtaining positive results with I-131 dWBS were varied throughout its 95% confidence interval (1.5% to 5.0%). Figure 4 shows that the cost per incomplete response result and successful detection is consistently higher than the base case for the Tg IRMA arm.

Unit cost of an I-131 dWBS was varied to determine how much should it cost in order to be as cost-effective as Tg IRMA (base case). It was found that an I-131 dWBS procedure should cost around PHP 125.00 ($ 2.18) to be as cost-effective as the Tg IRMA surveillance with a positive incomplete response result as outcome. This drops further to around PHP 85.00 (or $ 1.48) to be as cost-effective as Tg IRMA in successfully detecting incomplete response. This is shown in Figure 5.

DISCUSSION

Based on the total cost estimated from the Markov model, this study has identified that Tg IRMA is more cost-effective than I-131 dWBS in detecting incomplete response and correct (true positive) detection in WDTC patients. This could be associated with its low cost and high detection rate. Furthermore, I-131 dWBS should depreciate by 98.6% from its original cost to be equally cost-effective with Tg IRMA.

IRMA was already available in the 1980s and has then improved overtime with the current recorded functional sensitivity of 0.2 ng/mL [26,30]. Based on studies, it has a 99% negative predictive value of undetectable serum stimulated Tg level during the first year of surveillance [28]. It was identified to have 100% specificity, but with certain pitfalls [26,28]. Other than IRMA, radioimmunometric assay (RIA) is one of the nuclear medicine technologies utilized in serum Tg measurement. As IRMA is falsely lowered while RIA is falsely raised by anti-Tg antibodies in the sample [13]. RIA was identified to have lower sensitivity (87%), specificity (88.4%), and accuracy (86.1%) as compared to IRMA, which was utilized in this study and offered by the UP-PGH Radioisotope Laboratory [31]. IRMA is also used as reference in the development of other new, non-nuclear medicine technologies for serum Tg measurement like immunochemiluminometric, immunoenzymometric, and immunoluminometric assays [32–34].

I-131 dWBS, on the other hand, has been less utilized due to its low sensitivity and "stunning" [35]. According to literature, the frequency of stunning is highly variable from 5% to 40%. This makes a conundrum among clinicians because using higher doses of I-131 would provide greater sensitivity but would induce stunning [13]. Nevertheless, I-131 dWBS remains to be part of the ATA guidelines for WDTC surveillance.

The results of this study complement other literature indicating that I-131 dWBS may not be warranted in patients with undetectable Tg. Pacini et al., found that 71.4% of their WDTC patients had negative scans with Tg of <3 ng/mL, adding no relevant information or change in management [36]. More so, other studies have identified the absence of relation between the Tg levels and detectable uptake in thyroid bed [37].

Furthermore, this study estimated a cost as an outcome that is easy to understand. These results may help physicians and patients in coming up with a surveillance plan wherein the patient has to spend less. Laboratory managers, hospital administrators, and insurance companies can utilize these data in allocating and managing their resources and funding for the surveillance of WDTC patients.

Clinical Guidelines

This study has identified that the DOH 2021 guidelines in WDTC surveillance using nuclear medicine procedures will incur lesser expenditure as compared to the ATA 2015 guidelines. Even though the former recommends more frequent Tg and anti-Tg tests in the first and subsequent two years of surveillance, the total cost is still PHP 2,900.00 to PHP 4,850.00 less than the latter. It is expected that hospitals and patients will spend less on WDTC surveillance using nuclear medicine procedures as clinicians follow the recently released local guidelines.

With the increasing burden of WDTC in the Philippines, a national guideline on thyroid cancer care was commissioned and released in 2021 by the Department of Health. The "2021 Philippine Interim Clinical Practice Guidelines for the Diagnosis and Management of Well-Differentiated Thyroid Cancer" has been created to make recommendations that are applicable in the local setting due to cost and availability [29]. In terms of surveillance, there are only few deviations from the ATA 2015 guidelines. It is expected that the DOH 2021 guidelines would solve the encountered difficulties of local clinicians as this is more suitable to the Philippine setting. More so, the spread of new nuclear medicine facilities in the country will lead to increased referral and requests for these services.

Lubitz, et al., have identified that WDTC surveillance in 2013 incurred a total cost of $520,511,027.00 or 32%. This is only second to initial treatment which amounts to $623,367,851 or 39%. It is projected that by 2030, WDTC surveillance would balloon into $1,272,981,889, higher than the initial treatment ($907,578,188) [17]. With the worldwide increase in incidence rate of WDTC, it is expected to see this trend in the country .

The results of the cost-effective analysis through Markov model complements the lack of I-131 dWBS recommendations in the subsequent years of surveillance with the three clinical guidelines of thyroid cancer care. Furthermore, it justifies the more frequent measurement of Tg (with anti-Tg) in the DOH 2021 clinical guidelines. However, this does not negate the practice of combining the two nuclear medicine procedures as data on it remains limited. It is highly emphasized that patient care should be individualized and decisions made by the primary physician should be appropriate to setting, patient's needs and capabilities.

Limitations and Recommendations

This study utilized a simple Markov model since the scope is limited to two disease states and two nuclear medicine diagnostic procedures. This model could be replicated in other studies with varying parameters.

The incurred total expenditure estimates reflect only the direct costs from nuclear medicine procedures used in WDTC surveillance, specifically Tg IRMA and I-131 dWBS. Other sources of surveillance-related expenses like indirect costs (fare, salary from missed work days, etc.) and quality of life (QALY) could provide a wider analysis in the cost-effectiveness of the said nuclear medicine procedures. Further CEA studies on WDTC surveillance could include positron emission tomography (PET) to include dedifferentiated thyroid cancer. Other non-nuclear medicine diagnostic modalities like ultrasound and computed tomography (CT) scan could be incorporated to provide a comprehensive analysis as this study has proven that Markov analysis could be applied on diagnostic procedures

The parameters identified are assumed to be the best available fit for the set objectives of the study. However, local data remains to be limited and more recent studies are needed. These do not reflect the individual preferences of the clinicians, and the capability and/or willingness of patients to spend. Furthermore, the authors recognize that I-131 dWBS are usually done after Tg measurement as part of clinical practice guidelines.

The schedule of fees used in this study is based on what the UP-PGH Radioisotope Laboratory is offering as of November 2022. It is recognized that they are relatively cheaper compared to other institutions. To address this, service and pay rates were utilized to provide a range and to reflect the hospital expenditure and patient's out-of-pocket expenses. Replication of this study and further researches may cover the effect of inflation and other economic factors on the prices and costs .

A census-dependent approach in one institution may lead to underestimation of the actual incidence of incomplete response and true cost of surveillance as not all patients do their tests in the same institution or are compliant with their doctor's orders. This approach could be applied to write a descriptive study, and should involve several institutions catering to a large number of WDTC patients.

CONCLUSION

This study has identified that Tg IRMA is more cost-effective than I-131 dWBS in detecting incomplete response in patients with WDTC due to its low cost and high detection rate. This complements the lack of recommendations for I-131 dWBS as part of surveillance several clinical guidelines. In addition, the new DOH 2021 guidelines will incur less expenditure in WDTC surveillance using nuclear medicine procedures as compared to the ATA 2015 guidelines.

As this study provided a cost to WDTC surveillance, it could complement further researches on economic analysis and financial burden on WDTC patients

REFERENCES

1. Jameson J, Fauci A, Kasper D, Hauser S, Longo D, Loscalzo J, editors. Harrison's Principles of Internal Medicine. 20th ed. McGraw Hill; 2018.
2. Dunlap Q, Davies L. Differentiated Thyroid Cancer Incidence. Surgery of the Thyroid and Parathyroid Glands, Elsevier; 2021, p. 174-180.e2. https://doi.org/10.1016/B978-0-323-66127-0.00017-X.
3. Brown RL, de Souza JA, Cohen EE. Thyroid Cancer: Burden of Illness and Management of Disease. J Cancer 2011;2:193–9. https://doi.org/10.7150/jca.2.193.
4. Roman BR, Morris LG, Davies L. The Thyroid Cancer Epidemic, 2017 Perspective. Current Opinion in Endocrinology & Diabetes and Obesity 2017;24:332–6. https://doi.org/10.1097/MED.0000000000000359.
5. Megwalu UC, Osazuwa-Peters N, Moon P, Palaniappan LP. Thyroid Cancer Incidence Trends Among Filipinos in the

United States. Laryngoscope 2022; 132 : 1495 – 502. https://doi.org/10.1002/lary.29986.

6. Lo TEN, Uy AT, Maningat PDD. Well-Differentiated Thyroid Cancer: The Philippine General Hospital Experience. Endocrinology and Metabolism 2016; 31:72. https://doi.org/10.3803/EnM.2016.31.1.72.
7. Haugen BR, Alexander EK, Bible KC, Doherty GM, Mandel SJ, Nikiforov YE, et al. 2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer. Thyroid 2016;26:1–133. https://doi.org/10.1089/thy.2015.0020.
8. Lubitz CC, Sosa JA. The changing landscape of papillary thyroid cancer: Epidemiology, management, and the implications for patients. Cancer 2016;122:3754–9. https://doi.org/10.1002/cncr.30201.
9. Bates MF, Lamas MR, Randle RW, Long KL, Pitt SC, Schneider DF, et al. Back so soon? Is early recurrence of papillary thyroid cancer really just persistent disease? Surgery 2018;163:118–23. https://doi.org/10.1016/j.surg.2017.05.028.
10. Tuttle RM, Tala H, Shah J, Leboeuf R, Ghossein R, Gonen M, et al. Estimating Risk of Recurrence in Differentiated Thyroid Cancer After Total Thyroidectomy and Radioactive Iodine Remnant Ablation: Using Response to Therapy Variables to Modify the Initial Risk Estimates Predicted by the New American Thyroid Association Staging System. Thyroid 2010;20:1341–9. https://doi.org/10.1089/thy.2010.0178.
11. Santiago AG, Isidro MJ, Parra J. Predictors of Response to Therapy Among Post Thyroidectomy Adult Filipino Patients with Papillary Thyroid Carcinoma Based on the 2015 American Thyroid Association Guidelines. J ASEAN Fed Endocr Soc 2021;36. https://doi.org/10.15605/jafes.036.02.18.
12. Mendoza ES, Lopez AA, Valdez VAU, Cunanan EC, Matawaran BJ, Kho SA, et al. Predictors of incomplete response to therapy among Filipino patients with papillary thyroid cancer in a tertiary hospital. J Endocrinol Invest 2016;39:55–62. https://doi.org/10.1007/s40618-015-0319-2.
13. Ringel MD, Ladenson PW. Controversies in the follow-up and management of well-differentiated thyroid cancer. Endocr Relat Cancer 2004;11:97–116. https://doi.org/10.1677/erc.0.0110097.
14. Thyroid Cancer PWG. Clinical Practice Guidelines of the Philippine General Hospital for the Management of Thyroid Nodules and Well-Differentiated Thyroid Carcinoma (2008). Acta Med Philipp 2008;42.
15. Sundram F, Robinson BG, Kung A, Lim-Abrahan MA, Bay NQ, Chuan LK, et al. Well-Differentiated Epithelial Thyroid Cancer Management in the Asia Pacific Region: A Report and Clinical Practice Guideline. Thyroid 2006;16:461–9. https://doi.org/10.1089/thy.2006.16.461.
16. Yang SP, Ying LS, Saw S, Michael Tuttle R, Venkataraman K, Su-Ynn C. Practical barriers to implementation of thyroid cancer guidelines in the Asia-pacific region. Endocrine Practice 2015;21:1255 –68. https://doi.org/10.4158/EP15850.OR.
17. Lubitz CC, Kong CY, McMahon PM, Daniels GH, Chen Y, Economopoulos KP, et al. Annual financial impact of well-differentiated thyroid cancer care in the United States. Cancer 2014;120:1345–52. https://doi.org/10.1002/cncr.28562.
18. Wang LY, Roman BR, Migliacci JC, Palmer FL, Tuttle RM, Shaha AR, et al. Cost-effectiveness analysis of papillary thyroid cancer surveillance. Cancer 2015;121:4132–40. https://doi.org/10.1002/cncr.29633.
19. Janovsky CCPS, Bittencourt MS, Novais MAP de, Maciel RMB, Biscolla RPM, Zucchi P. Thyroid cancer burden and economic impact on the Brazilian public health system. Arch Endocrinol Metab 2018;62:537–44. https://doi.org/10.20945/2359-3997000000074.
20. Cellini SR, Kee JE. Cost-Effectiveness and Cost-Benefit Analysis. Handbook of Practical Program Evaluation, Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2015, p. 636–72. https://doi.org/10.1002/9781119171386.ch24.
21. Brockhuis B, Lass P, Popowski P, Scheffler J. An introduction to economic analysis in medicine - the basics of methodology and chosen terms. Examples of results of evaluation in nuclear medicine. Nuclear Medicine Review 2002;5:55–9.
22. Rudmik L, Drummond M. Health economic evaluation: Important principles and methodology. Laryngoscope 2013;123:1341–7. https://doi.org/10.1002/lary.23943.
23. Briggs AH, Claxton K, Sculpher MJ. Decision Modelling for Health Economic Evaluation. Oxford University Press; 2006.
24. Komorowski M, Raffa J. Markov Models and Cost Effectiveness Analysis: Applications in Medical Research. Secondary Analysis of Electronic Health Records, Cham: Springer International Publishing; 2016, p. 351–67. https://doi.org/10.1007/978-3-319-43742-2_24.
25. Sato RC, Zouain DM. Markov Models in health care. Einstein (São Paulo) 2010;8:376–9. https://doi.org/10.1590/s1679-45082010rb1567.

26. Giovanella L, Ceriani L. High-Sensitivity Human Thyroglobulin (hTG) Immunoradiometric Assay in the Follow-up of Patients with Differentiated Thyroid Cancer. Clin Chem Lab Med 2002;40. https://doi.org/10.1515/CCLM.2002.083 .
27. Schlumberger M, Hitzel A, Toubert ME, Corone C, Troalen F, Schlageter MH, et al. Comparison of Seven Serum Thyroglobulin Assays in the Follow-Up of Papillary and Follicular Thyroid Cancer Patients. J Clin Endocrinol Metab 2007;92:2487–95. https://doi.org/10.1210/jc.2006-0723.
28. Baudin E, Cao C do, Cailleux AF, Leboulleux S, Travagli JP, Schlumberger M. 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 2003;88:1107–11. https://doi.org/10.1210/jc.2002-021365.
29. Department of Health. The Philippine Interim Clinical Practice Guidelines for the Diagnosis and Management of Well-Differentiated Thyroid Cancer 2021. 2021.
30. Schlumberger M, Baudin E. Serum thyroglobulin determination in the follow-up of patients with differentiated thyroid carcinoma. Eur J Endocrinol 1998:249–52. https://doi.org/10.1530/eje.0.1380249.
31. Eustatia-Rutten CFA, Smit JWA, Romijn JA, van der Kleij-Corssmit EPM, Pereira AM, Stokkel MP, et al. Diagnostic value of serum thyroglobulin measurements in the follow-up of differentiated thyroid carcinoma, a structured meta-analysis. Clin Endocrinol (Oxf) 2004;61:61–74. https://doi.org/10.1111/j.1365-2265.2004.02060.x.
32. Cho YY, Chun S, Lee S-Y, Chung JH, Park H-D, Kim SW. Performance Evaluation of the Serum Thyroglobulin Assays With Immunochemiluminometric Assay and Immunoradiometric Assay for Differentiated Thyroid Cancer. Ann Lab Med 2016;36:413–9. https://doi.org/10.3343/alm.2016.36.5.413.
33. Broecker-Preuss M, Mehnert I, Gilman E, Herrmann K, Weber M, Görges R. Evaluation of a new automated assay for high-sensitivity thyroglobulin measurement and comparison with two established high-sensitivity thyroglobulin assays. Pract Lab Med 2021;26:e00250. https://doi.org/10.1016/j.plabm.2021.e00250.
34. Morgenthaler NG, Froehlich J, Rendl J, Willnich M, Alonso C, Bergmann A, et al. Technical evaluation of a new immunoradiometric and a new immunoluminometric assay for thyroglobulin. Clin Chem 2002;48:1077–83.
35. Klain M, Zampella E, Piscopo L, Volpe F, Manganelli M, Masone S, et al. Long-Term Prognostic Value of the Response to Therapy Assessed by Laboratory and Imaging Findings in Patients with Differentiated Thyroid Cancer. Cancers (Basel) 2021;13:4338. https://doi.org/10.3390/cancers13174338.
36. Pacini F, Capezzone M, Elisei R, Ceccarelli C, Taddei D, Pinchera A. Diagnostic 131-Iodine Whole-Body Scan May Be Avoided in Thyroid Cancer Patients Who Have Undetectable Stimulated Serum Tg Levels After Initial Treatment. J Clin Endocrinol Metab 2002;87:1499–501. https://doi.org/10.1210/jcem.87.4.8274.
37. Cailleux AF, Baudin E, Travagli JP, Ricard M, Schlumberger M. Is Diagnostic Iodine-131 Scanning Useful after Total Thyroid Ablation for Differentiated Thyroid Cancer? J Clin Endocrinol Metab 2000;85:175–8. https://doi.org/10.1210/jcem.85.1.6310.