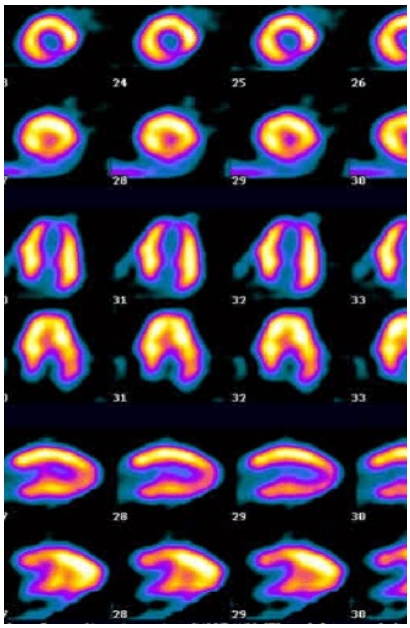
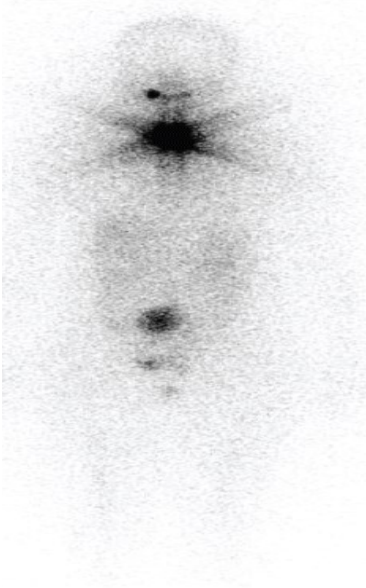
PJNM

Philippine Journal of Nuclear Medicine

Vol. 17 No. 2 July to December 2022





Contents:

8 Cost-Effectiveness Analysis of Well-Differentiated Thyroid Carcinoma Surveillance Using Nuclear Medicine Procedures
John Kenneth V. Gacula, RN, MD, Jerry M. Obaldo, MD, MHA, Vio Jianu C. Mojica, MS

22 Association of Osteoporosis with Radiologic Grading of the Hip Among Older Filipino Patients with Suspected Hip Osteoarthritis
Carl Johnry J. Santos, MD, Seth Gabriel F. Estanislao, MD, Irene S. Bandong, MD

36 ^{99m}Technetium Pertechnetate Thyroid Scan for Remnant Thyroid Tissue Detection Among Post-Thyroidectomy Patients at Jose R. Reyes Memorial Medical Center: A Retrospective Analysis
Athena Charisse S. Ong, MD, Marcelino A. Tanquilut, MD, Wenceslao S. Llauderes, MD, Emelito O. Valdez-Tan, MD, Ivan Ray F. David, MD

44 Role of Early Dynamic PET/CT Scan Imaging with ^{18}F-PSMA-1007 in Staging and Restaging Prostate Cancer in a Tertiary Private Hospital
Arrene Joy B. Baldonado, MD

54 Thyroid Disorder Classification using Machine Learning Models
Vincent Peter C. Magboo, MD, MS, Ma. Sheila A. Magboo, MS

Philippine Journal of Nuclear Medicine

Volume 17 No. 2
July to December 2022

Official Publication of the Philippine Society of Nuclear Medicine

The Philippine Journal of Nuclear Medicine is a peer-reviewed journal published by the Philippine Society of Nuclear Medicine. Subscription is free to all PSNM members in good standing as part of their membership privileges.

The Journal will be primarily of interest to medical and paramedical personnel working in nuclear medicine and related fields. Original works in clinical nuclear medicine and allied disciplines in physics, dosimetry, radiation biology, computer science, radiopharmacy, and radiochemistry are welcome. Review articles are usually solicited and published together with related reviews. Case reports of outstanding interest are likewise welcome. PSNM documents and position papers of interest to the reader will also be published as necessary.

Manuscripts for consideration should be sent to:

The Editor, Vincent Peter C. Magboo, MD
Philippine Journal of Nuclear Medicine
c/o Philippine Society of Nuclear Medicine
Unit 209 One Beatriz Tower Condo
4 Lauan St. cor. Aurora Blvd
Project 3, Quezon City 1102, Philippines
Contact. No.: +63 (966) 976 5676
Email: vcmagboo@up.edu.ph
philnucmed@gmail.com

All business communications and requests for complimentary copies should be addressed to the above.

ISSN 1655-9266

INFORMATION FOR AUTHORS

EDITORIAL POLICY

The Philippine Journal of Nuclear Medicine is the official peer-reviewed publication of the Philippine Society of Nuclear Medicine. The Journal accepts original articles pertinent to the field of nuclear medicine. Articles may be on any of the following: clinical and basic sciences, case reports, technical notes, special contributions, and editorials.

Submission of manuscripts

The submitted manuscript package should consist of: (1) the full text (including tables) in Microsoft Word, plain text or ConTeXt document format; and (2) high-resolution JPEG files of all images used in the manuscript. The complete manuscript package may be submitted as a compressed (.ZIP) file by email to philnucmed@gmail.com, or in an optical disc (CD/DVD) and mailed to

> The Editor: Philippine Journal of Nuclear Medicine, c/o Philippine Society of Nuclear Medicine, Unit 209 One Beatriz Tower Condo, 4 Lauan St. cor. Aurora Blvd., Project 3, Quezon City 1102, Philippines.

Manuscripts should be accompanied by a cover letter signed by the author responsible for correspondence regarding the manuscript. The cover letter should contain the following statement:

"All copyright ownership is transferred to the Philippine Journal of Nuclear Medicine upon acceptance of the article _____. This manuscript has been seen and approved by all the authors. The authors stipulate that the material submitted to the Philippine Journal of Nuclear Medicine is an original work and has not been submitted to another publication for concurrent consideration. Any human and/or animal studies undertaken as part of the research are in compliance with regulations of our in stitution(s) and with generally accepted guidelines governing such work."

The cover letter should also give any additional information that may be helpful to the Editor. Signed cover letters sent by email should be in PDF format.

MANUSCRIPT FORMAT

Manuscripts must be written in English, and printed on letter-sized white bond paper, 8.5 in x 11 in (21.6 cm x 27.9 cm). The text should be on one side of the paper only, single-spaced, with at least 1.5 in (4 cm) margins on all sides. Each of the following sections must begin on separate pages and in the following order: title page, abstract, text, acknowledgments, references, tables (each on a separate page), and legends. Pages should be consecutively numbered beginning with the title page. The first line of paragraphs should be indented by at least five spaces.

Title page

The title page should include: (1) a concise but informative title; (2) a short running head or footline of no more than 40 characters; (3) a complete byline, with first name, middle initial, and last name of each author and highest adacemic degrees; (4) the complete affiliation for each author, with the name of departments and institutions to which the work should be attributed; (5) disclaimers, if any; (6) the name, address, and telephone number of the author responsible for correspondence about the manuscript; and (7) the name and address of author to whom reprint requests should be directed.

Abstract and key words
An abstract of no more than 300 words should state the purpose of the study or investigation, summary of methodology, major findings, and principal conclusions. New and important aspects of the study or observations should be emphasized. No figures, abbreviations or reference citations are to be used in the abstract.

Text

The text of original scientific and technical articles is usually divided into the following sections: Introduction, Materials and Methods, Results, Discussion, and Summary or Conclusion.

Case reports are divided into the following sections: Introduction, Case Report, Discussion, and Conclusion. They should contain a concise description of one to three patients, emphasizing the nuclear medicine aspects and include methodology, data and correlative studies. Procedures should be described in sufficient detail to allow other investigators to reproduce the results.

Other articles, e.g. review articles, position papers, or editorials, should introduce a problem or question, present evidence, and conclude with an answer. Generally, review articles should have extensive documentation. Literature citations should represent the breadth and depth of the subjects being reviewed. The organization of review articles will depend greatly on the subject matter and material.

Generic names must be used throughout the text. Instruments and radiopharmaceuticals must be identified by manufacturer name and address in parentheses.

Acknowledgments

Persons or agencies contributing substantially to the work, including any grant support, must be acknowledged.

References

References must be cited in consecutive numerical order at first mention in the text and designated by the reference number in parentheses. References appearing in a table or figure should be numbered sequentially with those in the text.

The reference list must be numbered consecutively as in the text. The journal follows Index Medicus style for references and abbreviates journal names according to the List of Journals Indexed in Index Medicus. 'Unpublished observations' and 'personal communications' should not be used as references, although written—not verbal—communications may be noted as such in the text. The author is responsible for the accuracy of all references and must verify them against the original document.

For journal articles with six or less authors, all authors must be listed. For those with seven or more authors, only the first three are listed, and "et al." is added to the end of the list.

> Seabold JE, Conrad GR, Kimball DA, Ponto JA and Cricker JA. Pitfalls in establishing the diagnosis of deep venous thrombophlebitis by indium-111 platelet scintigraphy. J Nucl Med 1988;29:1169–1180.

For book and book chapters:

> Williams LJ. Evaluation of parathyroid function. In: Brock LJ, Stein JB, eds. The parathyroid and its diseases. 4th ed. New York: Wiley; 1985:196–248.
>
> Goodyear B. Bone marrow transplantation in severe combined immunodeficiency syndrome. In: Gree HJ, Blackscmith R, eds. Proceedings of the fourth biennial meeting of the International Society of Transplantation. Houston: International Society of Transplantation; 174: 44–46.

For journal article in electronic format:

> Author. Title. Journal name. Online publishing date. Available from: URL address.

Tables

Each table should be typed double-spaced on a separate page. Do not submit tables as photographs. Tables should be self-explanatory and should supplement, not duplicate, the text. Each table must be cited in consecutive numerical order in the text. Tables should be numbered consecutively with a Roman number following the word TABLE.

Illustrations

Illustrations should clarify and augment the text. Figures should be sharp and of high quality. Glossy photographs of line drawings rendered professionally on white drawing paper in black India ink, with template or typeset lettering, should be submitted. High quality computer-generated art is also acceptable. Letters, numbers, and symbols should be clear and of sufficient size to retain legibility after reduction.

Each illustration must be numbered and cited in consecutive order in the text. Illustrations should be identified on a gummed label. Legends should be typed double-spaced on a separate page. Figures should be numbered with an Arabic number following the word FIGURE.

Units of measurement

Use of the International System of Units (SI) is standard. Measurements of length, height, weight, and volume must be reported in metric units. Other measurements must be reported in the units in which they were made. Alternative units (non-SI units) should be added in parentheses by the author, if indicated.

Abbreviations and symbols

Only standard abbreviations and symbols should be used in the text. At first mention, the complete term, followed by its abbreviations in parentheses, must be used in the text. Standard units of measure should not be expanded at first mention. Consult a style manual, if necessary.

REVIEW PROCEDURE

Submitted manuscripts are peer-reviewed for originality, significance, adequacy of documentation, reader interest, composition, and adherence to the guidelines. Manuscripts are returned to the author for revision if suggestions and criticisms have been made. All accepted manuscripts are subject to editing for scientific accuracy, clarity, and style.

SAVING MORE LIVES

Providing access to innovative radiotracers and health solutions.

Health Assessment Cancer Diagnosis	Early Detection Cancer Staging	Monitoring of Treatment Response

FDG	FPSMA	FLORBETABEN	iFINDER

East Ave, Diliman, Quezon City, Metro Manila

NATIONAL KIDNEY AND TRANSPLANT INSTITUTE

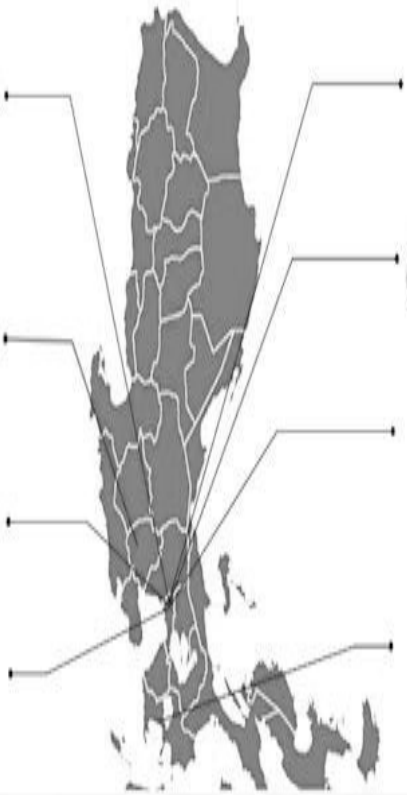
Angeles University Foundation Medical Center, Mc Arthur Highway, Angeles City, Pampanga,

2119 Dimasalang Rd, Santa Cruz, Manila, Metro Manila

i.Scan

850 United Nations Ave, Paco, Manila, Metro Manila

ManilaMed MEDICAL CENTER MANILA



CENTURIA MEDICAL MAKATI

Century City Kalayaan Ave. cor. Salamanca St., Brgy, Makati, Metro Manila

CARDINAL SANTOS MEDICAL CENTER

10 Wilson, Greenhills West, San Juan, 1502 Metro Manila

MMC MAKATI MEDICAL CENTER

2 Amorsolo Street, Legazpi Village, Makati, Kalakhang Maynila

MARY MEDIATRIX MEDICAL CENTER

Medical Center, JP Laurel Highway, Mary Mediatrix, Lipa, 4217 Batangas

www.khealthcorp.com

MedChoice

Thyroid & Diabetes

HEALTHCARE

Todo Aruga Sa Thyroid at Diabetes

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

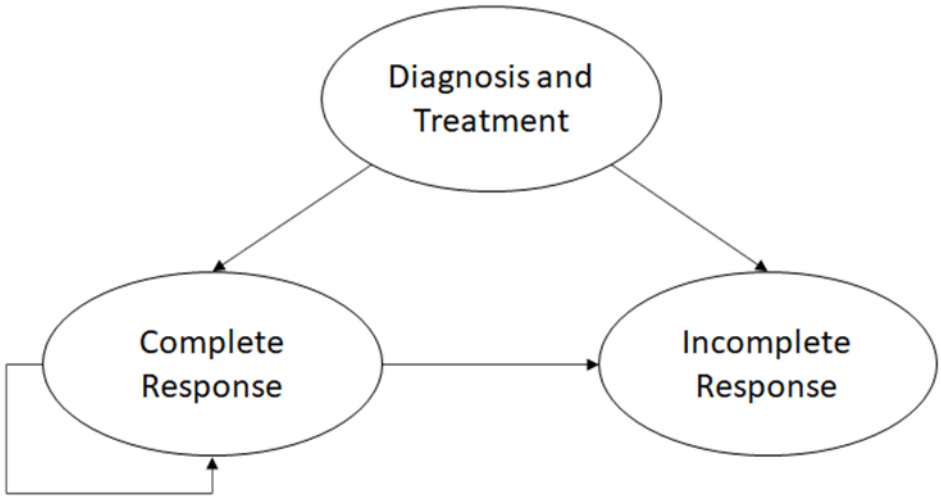


FIGURE 1. Diagram of the Markov model applied in this study

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	
Tg IRMA (Pay)	1870.00	UP-PGH
I-131 dWBS (Service)	8750.00	Radioisotope Laboratory
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)

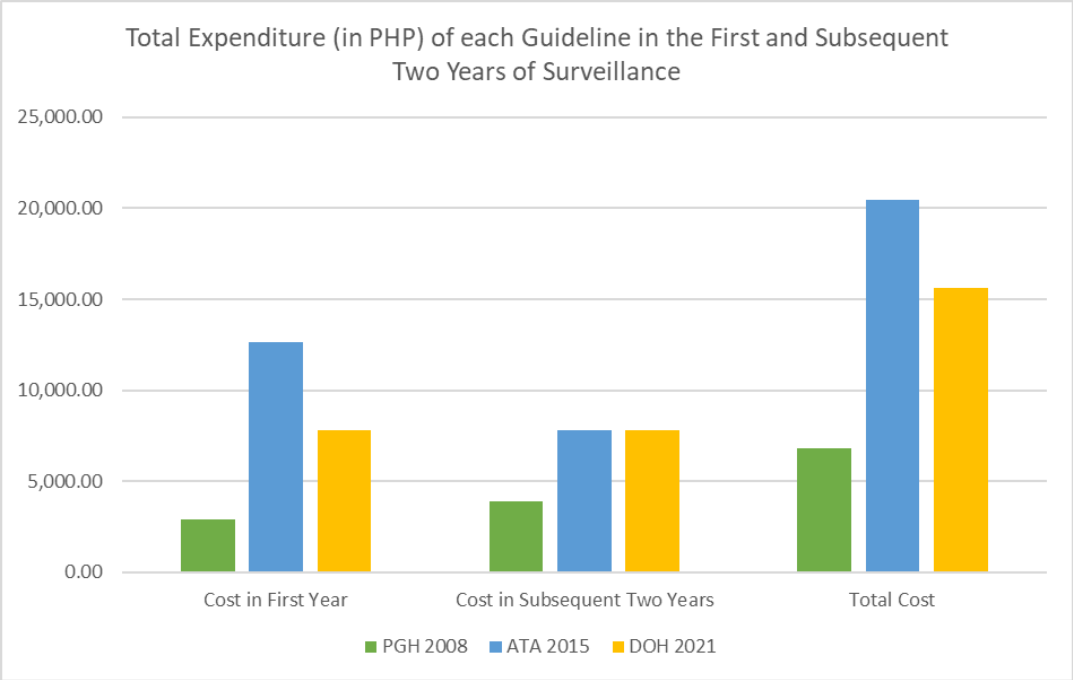


FIGURE 2. Graph comparing the total expenditure of each guideline in the first and subsequent two years of surveillance

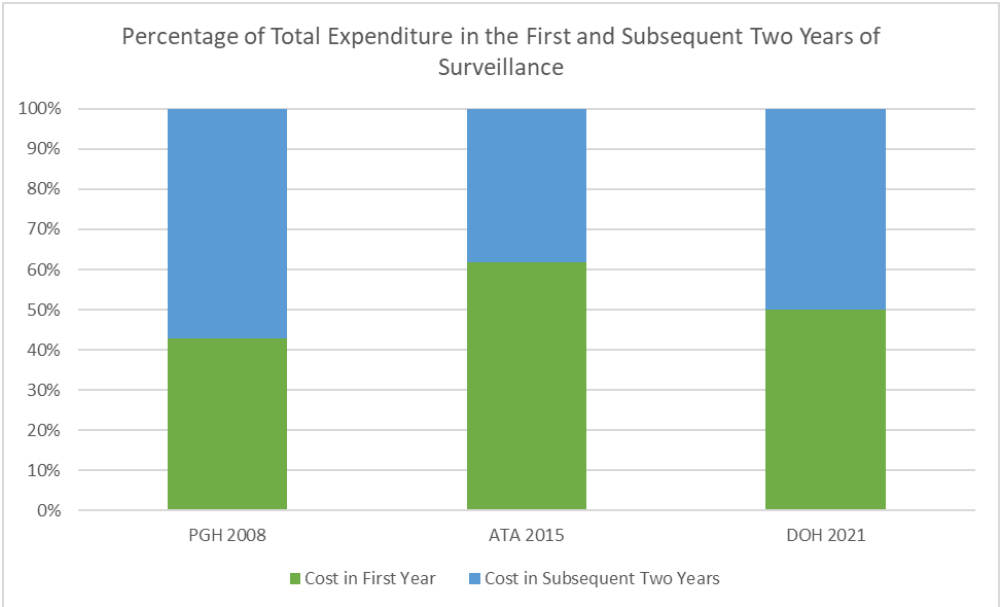


FIGURE 3. Graph comparing the division of total expenditure of each guideline in the first and subsequent two years of surveillance

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

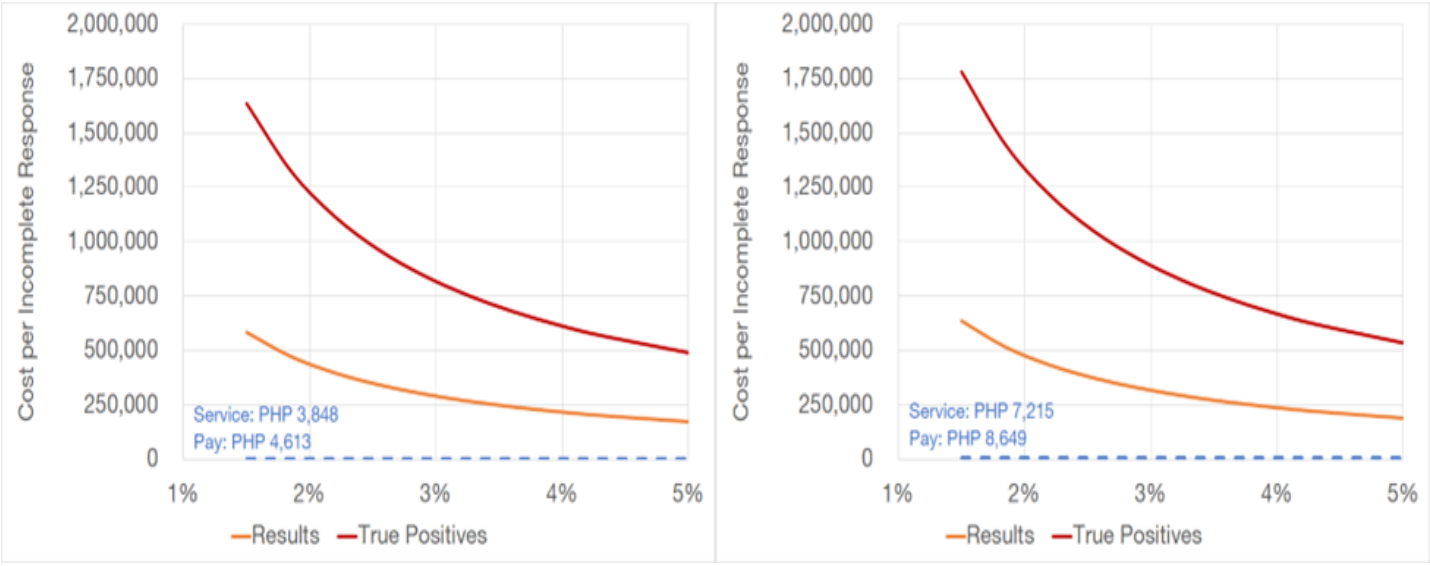


FIGURE 4. Costs per incomplete response result and successful detection for the I-131 dWBS arm over varying probabilities of positive result using service (left) and pay (right) rates.

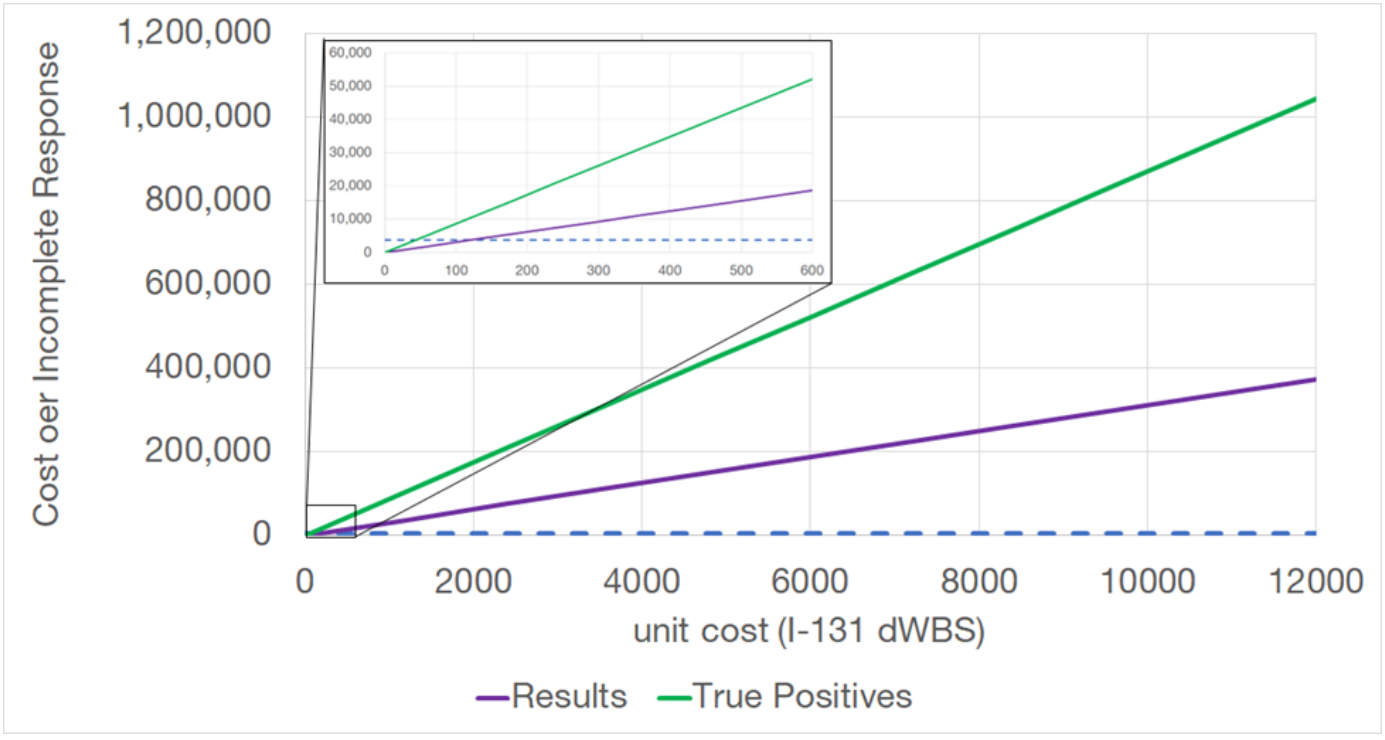


FIGURE 5. Costs per incomplete response result (green) and successful detection (purple) for the I-131 dWBS arm over varying unit cost per test; costs where I-131 dWBS will be as cost-effective as thyroglobulin IRMA is displayed (inset).

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.

PET-Positron Emission Production System

Radiochemistry System

Other RI Production System

NUCLEAR MEDICINE

Tc99m Generator

Radiopharmaceuticals

Sealed Sources

Hot laboratory set-up and accessories

RIA/IRMA Kits

Sumitomo Heavy Industries, Ltd.

PROTECH Medical

Assurance Controls Technologies Co., Inc.

Manila ~ Cebu ~ Davao +6327244149, +6327244150, +6327224580, +6327224588

Association of Osteoporosis with Radiologic Grading of the Hip Among Older Filipino Patients with Suspected Hip Osteoarthritis

Carl Johnry J. Santos, MD[1], Seth Gabriel F. Estanislao, MD[2], Irene S. Bandong, MD[1,2]

[1]*Department of Nuclear Medicine and Theranostics, St. Luke's Medical Center—Quezon City*
[2]*Institute of Radiology, St. Luke's Medical Center—Quezon City*
E-mail address: carljsantos@gmail.com, isbandong_md@yahoo.com

ABSTRACT

Introduction:

Among older populations, osteoarthritis (OA) is one of the most common chronic joint disorders and is a leading cause of disability, while osteoporosis is the most common metabolic bone disease, conferring fragility and significant risk of fracture. The relationship between OA and osteoporosis remains controversial. Although earlier studies reported an inverse association between the two diseases, more recent literature found a complex relationship mediated by various factors.

Objective:

The investigators sought to determine the association of osteoporosis with radiologic grading of the hip among older Filipino patients with suspected hip osteoarthritis.

Methodology:

A cross-sectional analytical study was conducted involving 256 patients with suspected hip OA who underwent radiography of the hips and central dual energy x-ray absorptiometry (DXA). Radiographs of the hips were evaluated by a radiologist using the Kellgren-Lawrence (KL) grading scale, while central DXA images were processed and evaluated by a nuclear medicine physician using the World Health Organization criteria for the diagnosis of osteoporosis and the 2019 International Society for Clinical Densitometry guidelines. The primary outcome measures were the prevalence of osteoporosis in patients with suspected hip OA, and the association of osteoporosis with radiologic KL grading of the hips. The secondary outcome measure was the association of osteoporosis with sex and BMI.

Results:

The study found that osteoporosis was present in 136 (53.1%) of the 256 patients who all presented with radiologic evidence of hip OA. There was a positive association between the presence of osteoporosis and the radiologic grade of hip OA (p-value: 0.006 on the right hip and 0.036 on the left). Osteoporosis was more prevalent in women compared to men (p-value: 0.031). Likewise, osteoporosis had a direct relationship with BMI (p-value: <0.001).

Conclusion:

Osteoporosis was prevalent in a significant proportion of older Filipino patients with clinical and radiologic evidence of hip OA, particularly among women, and was positively associated with increasing severity of OA. The study suggests that obesity may not necessarily protect against osteoporosis in this population, possibly relating to increased adiposity and decreased lean muscle mass.

Keywords: *osteoporosis, osteoarthritis, hip osteoarthritis, DXA scan*

INTRODUCTION

Osteoarthritis (OA) is one of the most common chronic joint disorders of the elderly and one of the leading causes of disability, affecting approximately 7% of the global population, disproportionately affecting women [1,2]. A similar demographic profile is noted in patients diagnosed with osteoporosis, which is regarded as the most common metabolic bone disease [3]. These two diseases constitute major health problems that confer substantial long-term economic burden on afflicted individuals [4].

Although traditional perspectives propose an inverse association between the two diseases with implications that OA and osteoporosis rarely coexist clinically, contemporary investigations now suggest that there exists a multifaceted relationship between the two musculoskeletal diseases that is influenced by various convergent and divergent factors [5,6,7,8]. Despite the continued global discourse on the relationship of OA and osteoporosis, there is a dearth of published literature concerning the two conditions in Filipino populations.

The present study investigated hip OA and its relation to osteoporosis using established measures for the two diseases – hip radiography with the Kellgren-Lawrence (KL) grading classification for OA and dual energy x-ray absorptiometry (DXA) using the 2019 guidelines from the International Society for Clinical Densitometry (ISCD) and the World Health Organization (WHO) criteria for osteoporosis. This investigation sought to supplement existing literature to further clarify the association between hip OA and osteoporosis among older Filipinos.

Relationship between osteoarthritis and osteoporosis

A negative association between OA and osteoporosis was first suggested half a century ago by Foss and Byers who found increased bone density in patients undergoing hip surgery due to OA [9]. This assertion was supported by cross-sectional studies which found that OA is associated with increased bone mass and density [6]. Sowers and colleagues found higher metacarpal bone mass in patients who were diagnosed with OA of the wrist and hands; however, bone mass was derived using radiographs as opposed to bone mineral densitometry. Cooper and colleagues likewise relied on radiographs of the hips in their study and reported a negative association between osteoporosis and OA [10,11]. As part of the large cross-sectional Study of Osteoporotic Fractures, Nevitt and colleagues used single photon absorptiometry and DXA to evaluate bone density at various sites, and found that elderly Caucasian women with moderate to severe radiographic hip OA had higher bone density in the hip, spine, and appendicular skeleton compared to women without hip OA [12]. Meanwhile, the Chingford study utilized central DXA to evaluate lumbar spine and femoral neck bone density, and reported small increases in the mean bone density of middle-aged women with early radiographic evidence of OA in the hands, knees, and lumbar spine; however, on follow-up, the same investigators interestingly found that low bone density at the femoral neck may be weakly associated with progression of the degree of OA [5,13]. The Rotterdam study also found that radiographic evidence of OA in the knees and hips was directly associated with higher femoral bone density, but follow-up DXA revealed a greater rate of bone loss over time among patients with OA compared to those without OA [14]. Chaganti and colleagues used both DXA and quantitative computed tomography (QCT) to assess for osteoporosis, and found higher bone density in older men with moderate to severe hip OA compared to those without hip OA [15]. More recently, Hardcastle and colleagues found that high bone mass, as defined by DXA bone density Z-scores, was directly associated to OA in the knees and osteophytosis, concluding that high bone mass confers a predisposition to a subtype of OA characterized by increased bone formation [16]. The Framingham study, on the other hand, found that while bone density at the proximal femur is higher among patients with grade 1 to 3 knee OA, no such association is seen in patients with grade 4 or more severe knee OA [17]. Data from the Baltimore Longitudinal Study of Aging showed that measures of appendicular bone mass using single photon absorptiometry had no significant association with the KL grade of hand OA, and on follow-up, observed that women with radiographic evidence of hand OA also experienced a greater rate of bone loss at the radius than women with normal hand radiographs similar to observations from the Rotterdam study [18]. This is supported by a recent study by Ding and colleagues, which found that older patients with hip and knee OA had a greater rate of total hip bone loss over time [19]. Small observational studies also found that a significant proportion of patients scheduled for total arthroplasty of the hip or knee also had osteoporosis [20, 21]. Histologic case reports likewise showed that sudden onset of knee

OA may be due to collapse of subchondral bone secondary to decreasing bone density [22]. The conflicting findings in literature may arise from the heterogeneity of OA across different joints, the variety of methods used for evaluating the diseases in question, as well as other mediating factors that may alter the relationship between osteoporosis and OA. As can be garnered from the abovementioned studies, OA may have different presentations in different joints with varying evidence for a negative association between OA and osteoporosis in the wrists, lumbar spine, hips, and knees. Sambrook and colleagues surmise that the association may even vary for bilateral hip OA and unilateral hip OA due to possible differences in their underlying causes [6]. The degree and severity of OA may also be a contributing factor, as data from the Framingham study suggests that the mere presence or absence of radiologic OA is insufficient to fully characterize its association with osteoporosis [17].

Measurements for osteoarthritis and osteoporosis

In interpreting existing literature on OA and osteoporosis, the tools being used for the assessment must also be considered. KL grading classification has largely been the method of choice to evaluate OA for the past six decades [23]. Bone density measurement, in contrast, has undergone various iterations over the years, as reflected in several of the studies that were discussed earlier. Nonetheless, the 2019 guidelines from the ISCD has reaffirmed the status of central DXA as the standard for the diagnosis of osteoporosis using the WHO criteria [24,25]. Reflective of these guidelines, current practice in clinical densitometry does not prescribe a cut-off bone density value, typically measured in g/cm^2, in the diagnosis of osteoporosis and instead utilizes T-scores for post-menopausal women and older men as the basis for bone density classifications as defined by the WHO.

Body mass index, obesity, and other mediating factors

Body mass index (BMI) and obesity are considered as important mediators between OA and osteoporosis [16]. Classic clinical experience characterized women with OA as obese with more fat, muscle mass, and strength, while women with osteoporosis were seen as generally slender with less fat, muscle girth, and strength [26].

Obesity is widely recognized as an important risk factor for the development and progression OA – initially attributed to biomechanical factors alone but in recent years has been found to involve complex mechanisms involving inflammatory and endocrine factors [27]. Obesity is also frequently linked to diabetes mellitus, which is understood to affect the risk of developing OA. Dubey and colleagues noted that the hyperglycemic state in diabetes can cause detrimental changes to the metabolism of normal articular cartilage, predisposing an individual to OA [28].

On the other hand, high BMI and obesity were traditionally associated with increased bone density and were largely believed to be protective against osteoporotic fractures [29]. Increased physical loading and strain were deemed favorable for bone geometry and modelling, while the adiposity associated with obesity was thought to preserve estrogen, which plays a key role in promoting bone formation while reducing bone resorption [30,31]. This was supported by literature showing that obese post-menopausal women have higher serum concentrations of estrogens compared to non-obese controls [32]. More recent studies, however, found that obesity may have adverse effects on bone mass and density due to its links with other metabolic changes, such as increased levels of pro-inflammatory cytokines TNF-α and IL-6 that are implicated in accelerated bone loss [31]. Moreover, obesity, diabetes mellitus, and insulin resistance are inversely associated with the concentration of adiponectin in the plasma, which is believed to have a favorable effect on bone mass and density [33,34]. Additionally, there is growing pre-clinical evidence in mice that obesity induced by a high-fat diet not only increases bone resorption but also facilitates fat infiltration of the bone marrow, which then facilitates osteoclastogenesis in the bone microenvironment [35,36]. These findings are supported by a recent cross-sectional study among elderly populations in Greece wherein osteoporosis was found to coexist with osteosarcopenic obesity, characterized by excess fat and low lean muscle mass [37].

Osteoarthritis and osteoporosis among Filipino patients

Considering the emergent literature for these complex relationships between OA, osteoporosis, obesity, and diabetes mellitus, further investigation appears to be warranted across different populations. Among Filipinos,

there is a paucity of data delving into such associations. A large cross-sectional study across the Philippines was done to identify risk factors for osteoporosis among Filipino adults and interestingly linked large body builds to increased prevalence of fractures; however, the data largely relied on self-reports rather than actual measurements of bone density [38]. Conversely, Miura and colleagues found low body weight as a predisposing factor for osteoporosis among post-menopausal Filipino women [39]. It must be noted, however, that this study used the non-conventional method of calcaneal measurements for the diagnosis of osteoporosis. Nevertheless, Mendoza and colleagues utilized DXA in their study involving adult Filipino males, and reached a similar conclusion regarding low BMI as a risk factor for osteoporosis [40]. To date, however, there have been no published studies investigating the relationship of OA and osteoporosis among Filipino populations.

OBJECTIVES:

General objective: To determine the association between osteoporosis and radiologic grading of the hip among older Filipino patients with suspected hip OA

Specific objectives:

1. To determine the prevalence of osteoporosis among older Filipino patients with suspected hip OA
2. To examine the relationship of osteoporosis with the radiologic grading of the hips among older Filipino patients with suspected hip OA

MATERIALS AND METHODS

Study design, population, and setting

This is a retrospective cross-sectional study involved adult patients, aged 50 and older, who presented with chronic hip pain, and had undergone central DXA and plain radiography of both hips in a period of six months, at St. Luke's Medical Center - Quezon City from January 1, 2018 to December 31, 2020 .

Exclusion criteria

1. Patients with previous hip injury or hip arthroplasty
2. Patients who were diagnosed with or suspected to have other rheumatologic diseases, and/or malignant lesions in the hips, apart from hip OA
3. Patients who have congenital abnormalities of the hips (e.g. developmental hip dysplasia)
4. Patients with incomplete data

Study procedure

Patients presenting with chronic hip pain and suspected to have hip OA had radiography of the hips using one of the three available stationary x-ray machines, namely: Siemens Aristos VX Plus S/N 1118, Siemens Multix Fusion VA20 S/N 1027, and Shimadzu Radspeed Pro DR S/N 3M5249A 64001. The radiographs were then reviewed and evaluated by a radiologist who was blinded from the bone mineral densitometry results. Grading of OA for each hip was done using the KL radiologic scale (see Table 1).

For central DXA, the patients were asked to avoid calcium-containing products (dairy products, calcium supplements, etc.) for 24 hours prior to the procedure. The patients were advised to avoid barium studies of the upper and lower gastrointestinal tract, intravenous pyelogram, or CT scan with contrast a week before the procedure. Patients were asked to change into a hospital gown and to remove their shoes and accessories. The patient's height and weight were measured and recorded prior to scanning. The DXA technologist assisted the patients in lying supine on the DXA machine. Scanning of the lumbar spine and hips was performed in Thick Mode, as determined by GE's Lunar software enCORE, lasting approximately 13 minutes. The

TABLE 1: Kellgren-Lawrence radiologic grading for hip osteoarthritis

Grade	**0**	**1**	**2**	**3**	**4**
Description	No joint space narrowing (JSN) or reactive changes	Doubtful JSN, possible osteophytic lipping	Definite osteophytes, possible JSN	Moderate osteophytes, definite JSN, some sclerosis, possible bone-end deformity	Large osteophytes, marked JSN, severe sclerosis, definite bone ends deformity

TABLE 2: WHO criteria for osteoporosis	
Normal	T-score at or above –1.0 SD
Low bone mineral density (osteopenia)	T-score between –1.0 and –2.5 SD
Osteoporosis	T-score at or below –2.5 SD
Severe osteoporosis	T-score at or below –2.5 SD and fragility fracture/s

machines were calibrated each day by using a standardized phantom to ensure consistency of the data collected. To evaluate for osteoporosis, the images were reviewed and evaluated by a nuclear medicine physician who was blinded from the hip radiography results. The images were processed and evaluated using the ISCD 2019 guidelines for adults based on the WHO criteria for the diagnosis of osteoporosis (see Table 2). Additionally, BMI was derived from the retrieved height and weight data that was routinely obtained as part of central DXA.

Outcome measures

1. Primary outcomes:

a) Prevalence of osteoporosis in patients with suspected hip OA
b) Association of osteoporosis with radiologic KL grading of the hips

2. Secondary outcome:

Association of osteoporosis with sex and BMI.

Sample Size

Based on the study of Ding and colleagues, a two-sided α of 90%, power of 10% was deemed significant [19]. Computing with a standard deviation of 0.34, the estimated sample size was 125 participants. Adjusting for 2 more variables (sex, BMI) in the analysis with an additional 20% for each control variable, the final sample size was 200 subjects.

Statistical Analysis

Data was encoded and tallied in SPSS version 10 for windows. Descriptive statistics were generated for all variables. For nominal data, frequencies and percentages were computed. For numerical data, mean ± SD were generated. Analysis of the different variables was done using the Chi-square test for nominal (categorical) data, while ANOVA was used to compare more than two groups with numerical data.

RESULTS

The study included a total of 256 patients of which 22 (9.0%) were male and 234 (91.0%) were female. Table 3 shows the quantitative characteristics of the patients in the study, detailing their mean ages, BMI, height, and weight. Between males and females, only height and weight were shown to have a significant difference (p-value <0.001).

Table 4 shows the clinical and radiologic profiles of the patients involved in the study. Osteoporosis was present in 136 (53.1%) patients, of which 15 (5.9%) were severe. All 256 (100%) patients had radiologic evidence of OA in both hips. Most patients were found to have KL grade 3 joint disease, as seen in the right hip of 174 (68%) patients and in the left hip of 160 (62.5%) patients.

Table 5 shows the association of sex to BMI classification, KL grade of hip OA, and bone mineral densitometry classifications. Females had significantly lower bone mineral densitometry classifications, i.e., more osteoporotic, compared to males (p-value of 0.031). No significant difference is seen between females and males in terms of BMI classification and KL grade of hip OA.

Table 6 shows that osteoporotic patients had significantly higher BMI than non-osteoporotic patients (p-value <0.001). It also shows that when grouped into BMI classifications, there was a significantly higher number of osteoporotic patients that were classified as obese compared to non-osteoporotic patients (p-value <0.001).

Meanwhile, Table 7 shows that that patients with osteoporosis had higher KL grade, i.e., more severe OA, in both hips compared to patients without osteoporosis (p-values of 0.006 for the right hip and 0.036 for the left hip). In a sub-group analysis of male patients, osteoporosis had no significant association with BMI. Likewise, no significant association was seen between osteoporosis and KL grading of hip OA among males (see Tables 8 and 9).

TABLE 3: Quantitative characteristics				
	Total **MEAN ± SD**	**Male** **MEAN ± SD**	**Female** **MEAN ± SD**	**p-value**
Age	67.9 ± 10.19	68.7 ± 10.15	67.8 ± 10.21	0.703
BMI	25.9 ± 4.83	26.1 ± 3.32	25.9 ± 4.95	0.799
Height (m)	1.5 ± 0.07	1.7 ± 0.06	1.5 ± 0.06	**<0.001**
Weight (kg)	61.1 ± 13.08	73.4 ± 11.50	60.0 ± 12.64	**<0.001**

TABLE 4: Clinical and radiologic profiles	
	n (%)
Prevalence of Osteoporosis	
Present (Osteoporosis + Severe osteoporosis)	136 (53.1)
Absent (Normal + Osteopenia)	120 (46.9)
Bone Mineral Densitometry Classification	
Normal	25 (9.8)
Osteopenia	95 (37.1)
Osteoporosis	121 (47.3)
Severe osteoporosis	15 (5.9)
Prevalence of Osteoarthritis	
Present (KL 1 + KL 2 + KL 3 + KL 4)	256 (100)
Absent (KL 0)	0 (0)
Kellgren-Lawrence Grade	
Right Hip	
KL 0	0 (0)
KL 1	5 (2.0)
KL 2	56 (21.9)
KL 3	174 (68.0)
KL 4	21 (8.2)
Left Hip	
KL 0	0 (0)
KL 1	6 (2.3)
KL 2	76 (29.7)
KL 3	160 (62.5)
KL 4	14 (5.5)

Meanwhile, among female patients, the data in Table 10 shows significantly higher BMI among osteoporotic patients versus non-osteoporotic patients (p-value of <0.001). It also demonstrates a significantly higher number of female patients that were classified as obese in the osteoporotic group than in the non-osteoporotic group (p-value <0.001). Table 11 shows that there was significantly higher KL grading of the right hip in osteoporotic patients compared to non-osteoporotic patients (p-value of 0.004). No such significance is observed for the KL grading of the left hip.

TABLE 5: Association of sex to body mass index classification, Kellgren-Lawrence grade and bone mineral					
			Male: n (%)	**Female: n (%)**	**p-value**
BMI Classification		Underweight	0 (0.0)	10 (4.3)	0.701
		Normal	8 (36.4)	92 (39.3)	
		Overweight	11 (50.0)	95 (40.6)	
		Obese I	3 (13.6)	29 (12.4)	
		Obese II	0 (0.0)	8 (3.4)	
Kellgren-Lawrence Grade	**Right Hip**	KL 0	0 (0.0)	0 (0.0)	0.247
		KL 1	0 (0.0)	5 (2.1)	
		KL 2	3 (13.6)	53 (22.6)	
		KL 3	15 (68.2)	159 (67.9)	
		KL 4	4 (18.2)	17 (7.3)	
	Left Hip	KL 0	0 (0.0)	0 (0.0)	0.518
		KL 1	1 (4.5)	5 (2.1)	
		KL 2	4 (18.2)	72 (30.8)	
		KL 3	15 (68.2)	145 (62.0)	
		KL 4	2 (9.1)	12 (5.1)	
Bone Mineral Densitometry Classification		Normal	6 (27.3)	19 (8.1)	**0.031**
		Osteopenia	8 (36.4)	87 (37.2)	
		Osteoporosis	7 (31.8)	114 (48.7)	
		Severe Osteoporosis	1 (4.5)	14 (6.0)	

TABLE 6: Association of osteoporosis to body mass index				
		OSTEOPOROSIS		**p-value**
		Present	**Absent**	
		MEAN ± SD	**MEAN ± SD**	
BMI		27.6 ± 4.98	24.4 ± 4.16	**<0.001**
		N (%)	**N (%)**	
BMI Classification	Underweight	0 (0.0)	10 (7.4)	**<0.001**
	Normal	36 (30.0)	64 (47.1)	
	Overweight	55 (45.8)	51 (37.5)	
	Obese I	21 (17.5)	11 (8.1)	
	Obese II	8 (6.7)	0 (0.0)	

DISCUSSION

Osteoporosis and OA are among the most prevalent musculoskeletal diseases across the globe, accounting for substantial fragility, disability, and healthcare utilization [1,2,3,4]. In the present study, females were found to be more osteoporotic compared to males, consistent with global trends of the disease [3].

Osteoporosis and OA were previously assumed to rarely

TABLE 7: Association of osteoporosis to Kellgren-Lawrence grade					
			OSTEOPOROSIS		**p-value**
			Present	**Absent**	
			n (%)	**n (%)**	
Kellgren-Lawrence Grade	**Right Hip**	KL 0	0 (0.0)	0 (0.0)	**0.006**
		KL 1	1 (0.8)	4 (2.9)	
		KL 2	16 (13.3)	40 (29.4)	
		KL 3	90 (75.0)	84 (61.8)	
		KL 4	13 (10.8)	8 (5.9)	
	Left Hip	KL 0	0 (0.0)	0 (0.0)	**0.036**
		KL 1	1 (0.8)	5 (3.7)	
		KL 2	29 (24.2)	47 (34.6)	
		KL 3	80 (66.7)	80 (58.8)	
		KL 4	10 (8.3)	4 (2.9)	

Table 8: Association of osteoporosis to body mass index among males				
		OSTEOPOROSIS		**p-value**
		Present	**Absent**	
		MEAN ± SD	**MEAN ± SD**	
BMI		26.5 ± 3.14	25.5 ± 3.74	0.536
		n (%)	**n (%)**	
BMI Classification	Underweight	0 (0.0)	0 (0.0)	0.592
	Normal	4 (50.0)	4 (50.0)	
	Overweight	8 (72.7)	3 (27.3)	
	Obese I	2 (66.7)	1 (33.3)	
	Obese II	0 (0.0)	0 (0.0)	

TABLE 9: Association of osteoporosis to Kellgren-Lawrence grade among males					
			OSTEOPOROSIS		**p-value**
			Present: n (%)	**Absent: n (%)**	
Kellgren-Lawrence Grade	**Right Hip**	KL 0	0 (0.0)	0 (0.0)	0.822
		KL 1	0 (0.0)	0 (0.0)	
		KL 2	2 (66.7)	1 (33.3)	
		KL 3	10 (66.7)	5 (33.3)	
		KL 4	2 (50.0)	2 (50.0)	
	Left Hip	KL 0	0 (0.0)	0 (0.0)	0.337
		KL 1	0 (0.0)	1 (100.0)	
		KL 2	2 (50.0)	2 (50.0)	
		KL 3	10 (66.7)	5 (33.3)	
		KL 4	2 (100.0)	0 (0.0)	

coexist clinically [5,6]. The present study showed that among older Filipino patients presenting with chronic hip pain – for which OA was the suspected cause – approximately half were found to have osteoporosis. Meanwhile, the hip radiographs show that all the patients in the study have varying degrees of OA in the hips. These findings appear to contradict earlier literature regarding osteoporosis and OA and support contemporary perspectives concerning the possibility of concomitant disease [7,8]. Ding and colleagues did not evaluate for the presence or absence of osteoporosis in their study but found an increased rate of total hip bone loss over time in patients with knee and hip OA [19]. The results of the current study are compatible with their observations and may imply that the loss of bone density experienced by such patients is clinically significant enough to warrant the diagnosis of osteoporosis based on the WHO criteria. These results are similar to the findings of small observational studies involving patients scheduled for total arthroplasty of the hip and knee secondary to OA; however, the current study osteoporosis at 53.1% as opposed to the previously reported 20–23%, probably owing to the differences in the populations involved [20,21]. Nevertheless, the diagnosis of osteoporosis is quite significant because it necessitates not only increased vigilance against potential fragility fractures but also the need for appropriate therapeutic interventions [24]. Additionally, in the context of providing surgical management for hip OA, the concurrent presence of osteoporosis may affect a higher proportion of individuals diagnosed with bone quality and may significantly compromise the stability of

TABLE 10: Association of osteoporosis to body mass index among females

		OSTEOPOROSIS		p-value
		Present	Absent	
		MEAN ± SD	MEAN ± SD	
BMI		27.8 ± 5.17	24.4 ± 4.19	**<0.001**
		n (%)	**n (%)**	
BMI Classification	Underweight	0 (0.0)	10 (100.0)	**<0.001**
	Normal	32 (34.8)	60 (65.2)	
	Overweight	47 (49.5)	48 (50.5)	
	Obese I	19 (65.5)	10 (34.5)	
	Obese II	8 (100.0)	0 (0.0)	

TABLE 11: Association of osteoporosis to Kellgren-Lawrence grade among females

			OSTEOPOROSIS		p-value
			Present: n (%)	Absent: n (%)	
Kellgren-Lawrence Grade	**Right Hip**	KL 0	0 (0.0)	0 (0.0)	**0.004**
		KL 1	1 (20.0)	4 (80.0)	
		KL 2	14 (26.4)	39 (73.6)	
		KL 3	80 (50.3)	79 (49.7)	
		KL 4	11 (64.7)	6 (35.3)	
	Left Hip	KL 0	0 (0.0)	0 (0.0)	0.122
		KL 1	1 (20.0)	4 (80.0)	
		KL 2	27 (37.5)	45 (62.5)	
		KL 3	70 (48.3)	75 (51.7)	
		KL 4	8 (66.7)	4 (33.3)	

the implant [21]. In light of these findings, it is thereby pertinent to screen for osteoporosis in older patients with suspected or confirmed hip OA, as the presence of osteoporosis may significantly alter their clinical outcomes.

Interestingly, the current investigation also showed that osteoporotic patients presented with higher KL grading in the hips compared to their non-osteoporotic counterparts, particularly among female patients. Among males alone, no such correlation was identified. There is the belief that lower bone density in the subchondral region may be seen in earlier OA prior to the onset of sclerotic changes that are observed in severe OA [1]. The results of the current study may seem to run contrary to this assertion, but given the nature of osteoporosis as a systemic condition as opposed to a largely localized pathology like OA, it is likely that while there is sclerosis in the subchondral region of the joint in severe OA as reflected in the KL grading scale, significant bone loss can still be observed in the bone regions being evaluated by central DXA (i.e., lumbar spine, femoral neck, total proximal femur); hence, the observations in the present study. These results are clinically pertinent given how patients with KL grades 3 and 4 – that is, more severe hip OA – are more likely to require surgical interventions such as total hip arthroplasty and, as mentioned earlier, may have worse clinical outcomes in the setting of concomitant osteoporosis.

Data from the current study suggests a direct association of high BMI to the presence of osteoporosis, which is again more significant among female patients. These results challenge the notion of obesity being protective against osteoporosis, providing further evidence for the growing literature concerning the unfavorable effects of obesity on bone density [29,31]. The data from the study appears to reflect the conclusions of Tanchoco and colleagues associating large body builds to osteoporosis among Filipino adults [38]. It must be emphasized, however, that BMI calculations only consider weight and height as factors, without characterizing distributions of body fat and muscle mass. There is growing pre-clinical and clinical evidence to suggest that excess adiposity and low lean muscle mass can contribute to decreased bone density and predispose an individual to osteoporosis [35,36,37]. The findings of the present study may be reflective of this process given how the patients involved are older with chronic hip pain secondary to OA, which may predispose them to sedentary lifestyles and osteosarcopenic obesity, characterized by high fat with concurrent loss of skeletal muscle [37]. This is further supported by sub-group analysis of the present data showing that significance is only observed among women who generally present with higher adiposity compared to men.

Certain limitations of the current study must be acknowledged. Whereas older studies have relied on measurements of bone density – quantified in g/cm^2 – to investigate the relationship of osteoporosis to OA, the investigators opted to dichotomize patients into either osteoporotic or non-osteoporotic categories in order to focus on clinically pertinent disease. Nevertheless, monitoring trends of bone loss over time is still best accomplished using changes in bone density values over the course of several years [24]. Such trends may provide further insight concerning the relationship of osteoporosis and OA; however, these lie beyond the scope of this cross-sectional study.

Another limitation is the relatively small sample population of male patients, which may have affected some of the results on sub-group analysis. Although both osteoporosis and OA disproportionately affect women, investigations involving older men may still provide insight regarding the two diseases.

Finally, the retrospective nature of the study is its most significant limitation, which restricted the extent of clinical data that could be examined. Apart from advanced age, there is limited information as to why central DXA was performed in this population. The hip radiographs would typically suffice for the primary complaint of chronic hip pain among the patients in the study, but the presence of other symptoms, if there are any, were not investigated. These unidentified factors may have contributed to the unexpectedly large number

of patients who had both osteoporosis and OA in this population. Although the study delved into the potential role of BMI in mediating the relationship between osteoporosis and OA, other conditions and metabolic states that could have affected both bone density and the hip joints were not investigated. Moreover, it must be acknowledged that the interval of several months between the hip radiographs and central DXA studies of some patients may have allowed extraneous variables to affect the findings. A prospective study, in contrast, may allow for a shorter time interval between the diagnostic studies, leading to stronger conclusions regarding the relationship of the two bone-related diseases.

CONCLUSIONS

In summary, osteoporosis may be found in a sizable proportion of older Filipino patients with hip OA, particularly among women. Furthermore, the presence of osteoporosis was positively associated with higher radiologic KL grading of the hips. The coexistence of both diseases may be mediated by obesity, a risk factor for OA that was previously thought to be protective against osteoporosis. The study found a positive association between osteoporosis and BMI, which suggests that obesity may not necessarily protect against osteoporosis in this particular population, possibly due to increased adiposity and decreased lean muscle mass.

RECOMMENDATIONS

Screening for osteoporosis using bone mineral densitometry may be warranted in older Filipinos with suspected or known hip OA given the substantial likelihood of an individual having both conditions. Further studies regarding the relationship of osteoporosis and hip OA are recommended, involving a larger sample population that would ideally include more male participants. Moreover, longitudinal studies involving serial central DXA scanning and hip radiographs can provide insights concerning the changes in radiologic KL grading and bone density over the course of several years. In view of the current study's findings regarding the relationship of obesity and osteoporosis, evaluating fat and lean muscle mass distribution among patients with concomitant OA and osteoporosis is recommended. Conveniently, the same DXA scanners used for bone mineral densitometry are capable of total body composition studies, which would provide a wealth of information for future investigators.

CONFLICTS OF INTEREST

The authors have no conflicts of interest to declare.

REFERENCES

1. Egloff C, Hügle T, Valderrabano V. Biomechanics and pathomechanisms of osteoarthritis [Internet]. Vol. 142, Swiss Medical Weekly. Swiss Med Wkly; 2012 [cited 2021 Jun 19]. Available from: https://pubmed.ncbi.nlm.nih.gov/22815119/
2. Hunter DJ, March L, Chew M. Osteoarthritis in 2020 and beyond: a Lancet Commission [Internet]. Vol. 396, The Lancet. Lancet Publishing Group; 2020 [cited 2021 Jun 19]. p. 1711–2. Available from: https://doi.org/10.1016/10.1001/jamapediatrics.2020.4573.
3. Sozen T, Ozisik L, Calik Basaran N. An overview and management of osteoporosis. Eur J Rheumatol [Internet]. 2017 Mar 1 [cited 2021 Jun 19];4(1):46–56. Available from: /pmc/articles/PMC5335887/
4. Sezer I, Illeez OG, Tuna SD, Balci N. The Relationship Between Knee Osteoarthritis and Osteoporosis. Eurasian J Med [Internet]. 2010 Dec 1 [cited 2021 Jun 20];42(3):124–7. Available from: https://pubmed.ncbi.nlm.nih.gov/25610141/
5. Hart DJ, Mootoosamy I, Doyle D V., Spector TD. The relationship between osteoarthritis and osteoporosis in the general population: the Chingford Study. Ann Rheum Dis [Internet]. 1994 [cited 2022 Nov 19];53(3):158. Available from: /pmc/articles/PMC1005278/?report=abstract
6. Sambrook P, Naganathan V. What is the relationship between osteoarthritis and osteoporosis? Baillieres Clin Rheumatol [Internet]. 1997 [cited 2022 Nov 18];11(4):695–710. Available from: https://pubmed.ncbi.nlm.nih.gov/9429732/
7. Geusens PP, van den Bergh JP. Osteoporosis and osteoarthritis: shared mechanisms and epidemiology. Curr Opin Rheumatol. 2016 Mar;28(2):97–103.
8. Hartley A, Gregson CL, Paternoster L, Tobias JH. Osteoarthritis: Insights Offered by the Study of Bone Mass Genetics [Internet]. Vol. 19, Current Osteoporosis Reports. Springer; 2021 [cited 2021 Jun 20]. p. 115–22. Available from: /pmc/articles/PMC8016765/
9. Foss M V., Byers PD. Bone density, osteoarthrosis of the hip, and fracture of the upper end of the femur. Ann Rheum Dis [Internet]. 1972 [cited 2021 Jun 19];31(4):259–

64. https://pubmed.ncbi.nlm.nih.gov/5045904/
10. Sowers M, Zobel D, Hawthorne VM, Carman W, Weissfeld L. Progression of osteoarthritis of the hand and metacarpal bone loss. A twenty-year followup of incident cases. Arthritis Rheum [Internet]. 1991 Jan 1 [cited 2022 Nov 21];34(1):36–42. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/art.1780340106
11. Cooper C, Cook PL, Osmond C, Fisher L, Cawley MID. Osteoarthritis of the hip and osteoporosis of the proximal femur. Ann Rheum Dis [Internet]. 1991 Aug 1 [cited 2022 Nov 21];50(8):540–2. Available from: https://ard.bmj.com/content/50/8/540
12. Nevitt MC, Lane NE, Scott JC, Hochberg MC, Pressman AR, Genant HK, et al. Radiographic osteoarthritis of the hip and bone mineral density. Arthritis Rheum [Internet]. 1995 Jul 1 [cited 2022 Nov 21];38(7):907–16. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/art.1780380706
13. Hart DJ, Cronin C, Daniels M, Worthy T, Doyle D V, Spector TD, et al. The Relationship of Bone Density and Fracture to Incident and Progressive Radiographic Osteoarthritis of the Knee The Chingford Study. ARTHRITIS Rheum. 2002;46(1):92–9.
14. Bergink AP, Uitterlinden AG, Van Leeuwen JPTM, Hofman A, Verhaar JAN, Pols HAP. Bone mineral density and vertebral fracture history are associated with incident and progressive radiographic knee osteoarthritis in elderly men and women: The Rotterdam Study. Bone [Internet]. 2005 Oct [cited 2021 Jun 20];37(4):446–56. Available from: https://pubmed.ncbi.nlm.nih.gov/16027057/
15. Chaganti RK, Parimi N, Lang T, Orwoll E, Stefanick ML, Nevitt M, et al. Bone mineral density and prevalent osteoarthritis of the hip in older men for the Osteoporotic Fractures in Men (MrOS) Study Group. Osteoporos Int [Internet]. 2010 Aug [cited 2021 Jun 19];21(8):1307–16. Available from: /pmc/articles/PMC3354730/
16. Hardcastle SA, Dieppe P, Gregson CL, Arden NK, Spector TD, Hart DJ, et al. Individuals with high bone mass have an increased prevalence of radiographic knee osteoarthritis. Bone [Internet]. 2015 Feb 1 [cited 2021 Jun 19];71:171–9. Available from: https://pubmed.ncbi.nlm.nih.gov/25445455/
17. Hannan MT, Anderson JJ, Zhang Y, Levy D, Felson DT. Bone mineral density and knee osteoarthritis in elderly men and women. The Framingham Study. Arthritis Rheum [Internet]. 1993 Dec 1 [cited 2022 Nov 21];36(12):1671–80. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/art.1780361205.
18. Hochberg MC, Lethbridge-Cejku M, Tobin JD. Bone mineral density and osteoarthritis: Data from the Baltimore Longitudinal Study of Aging. Osteoarthr Cartil [Internet]. 2004 [cited 2021 Jun 19];12(SUPLL.):45–8. Available from: https://pubmed.ncbi.nlm.nih.gov/14698641/
19. Ding C, Cicuttini F, Boon C, Boon P, Srikanth V, Cooley H, et al. Knee and hip radiographic osteoarthritis predict total hip bone loss in older adults: A prospective study. J Bone Miner Res [Internet]. 2010 Apr [cited 2021 Jun 19];25(4):858–65. Available from: https://pubmed.ncbi.nlm.nih.gov/19821767/
20. Lingard EA, Mitchell SY, Francis RM, Rawlings D, Peaston R, Birrell FN, et al. The prevalence of osteoporosis in patients with severe hip and knee osteoarthritis awaiting joint arthroplasty. Age Ageing. 2010 Mar;39(2):234–9.
21. Domingues VR, de Campos GC, Plapler PG, de Rezende MU. Prevalence of osteoporosis in patients awaiting total hip arthroplasty. Acta Ortop Bras. 2015;23(1):34–7.
22. Horikawa A, Miyakoshi N, Shimada Y, Kodama H. The Relationship between Osteoporosis and Osteoarthritis of the Knee: A Report of 2 Cases with Suspected Osteonecrosis. Case Rep Orthop [Internet]. 2014 [cited 2021 Jun 19];2014:1–6. Available from: /pmc/articles/PMC4090430/
23. Kohn MD, Sassoon AA, Fernando ND. Classifications in Brief: Kellgren-Lawrence Classification of Osteoarthritis. Clin Orthop Relat Res [Internet]. 2016 Aug 1 [cited 2022 Nov 19];474(8):1886. Available from: /pmc/articles/PMC4925407/
24. Shuhart CR, Yeap SS, Anderson PA, Jankowski LG, Lewiecki EM, Morse LR, et al. Executive Summary of the 2019 ISCD Position Development Conference on Monitoring Treatment, DXA Cross-calibration and Least Significant Change, Spinal Cord Injury, Peri-prosthetic and Orthopedic Bone Health, Transgender Medicine, and Pediatrics. J Clin Densitom [Internet]. 2019 Oct 1 [cited 2022 Nov 19];22(4):453–71. Available from: https://pubmed.ncbi.nlm.nih.gov/31400968/
25. Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. Report of a WHO Study Group. World Health Organ Tech Rep Ser. 1994;843:1–129.
26. Dequeker J, Goris P, Uytterhoeven R. Osteoporosis and Osteoarthritis (Osteoarthrosis): Anthropometric Distinctions. JAMA [Internet]. 1983 Mar 18 [cited 2022 Nov 21];249(11):1448–51. Available from: https://jamanetwork.com/journals/jama/fullarticle/384873 .
27. Bliddal H, Leeds AR, Christensen R. Osteoarthritis, obesity and weight loss: evidence, hypotheses and horizons - a scoping review. Obes Rev an Off J Int Assoc Study Obes. 2014 Jul;15(7):578–86.
28. Dubey NK, Ningrum DNA, Dubey R, Deng YH, Li YC, Wang PD, et al. Correlation between diabetes mellitus and knee osteoarthritis: A dry-to-wet lab approach. Int J Mol Sci [Internet]. 2018 [cited 2021 Jun 20];19(10). Available from: https://pubmed.ncbi.nlm.nih.gov/30282957/
29. Premaor MO, Comim FV, Compston JE. Obesity and fractures. Arq Bras Endocrinol Metabol. 2014 Jul;58(5):470–7.
30. Addison O, Marcus RL, Lastayo PC, Ryan AS. Intermuscular fat: a review of the consequences and causes. Int J Endocrinol. 2014;2014:309570.

31. Gkastaris K, Goulis DG, Potoupnis M, Anastasilakis AD, Kapetanos G. Obesity, osteoporosis and bone metabolism. J Musculoskelet Neuronal Interact. 2020 Sep;20(3):372–81.
32. Leeners B, Geary N, Tobler PN, Asarian L. Ovarian hormones and obesity. Hum Reprod Update. 2017 May;23 (3):300–21.
33. Yamauchi T, Kamon J, Waki H, Terauchi Y, Kubota N, Hara K, et al. The fat-derived hormone adiponectin reverses insulin resistance associated with both lipoatrophy and obesity. Nat Med [Internet]. 2001 [cited 2022 Nov 22];7 (8):941–6. Available from: https:// pubmed.ncbi.nlm.nih.gov/11479627/
34. Jürimäe J, Rembel K, Jürimäe T, Rehand M. Adiponectin is associated with bone mineral density in perimenopausal women. Horm Metab Res [Internet]. 2005 May [cited 2022 Nov 22];37(5):297–302. Available from: https:// pubmed.ncbi.nlm.nih.gov/15971153/
35. Patsch JM, Kiefer FW, Varga P, Pail P, Rauner M, Stupphann D, et al. Increased bone resorption and impaired bone microarchitecture in short-term and extended high-fat diet-induced obesity. Metabolism. 2011 Feb;60(2):243–9.
36. Halade G V, El Jamali A, Williams PJ, Fajardo RJ, Fernandes G. Obesity-mediated inflammatory microenvironment stimulates osteoclastogenesis and bone loss in mice. Exp Gerontol. 2011 Jan;46(1):43–52.
37. Keramidaki K, Tsagari A, Hiona M, Risvas G. Osteosarcopenic obesity, the coexistence of osteoporosis, sarcopenia and obesity and consequences in the quality of life in older adults ≥365 years-old in Greece. J Frailty, Sarcopenia Falls [Internet]. 2019 Dec 1 [cited 2021 Jun 20];4(4):91–101. Available from: /pmc/articles/ PMC7155308/
38. Tanchoco C, Villadolid MF, Duante CA, Limbaga MLS, Yee GA. Risk factors associated with osteoporosis among Filipino adults. J Philipp Med Assoc [Internet]. 2004 [cited 2022 Nov 23];9–29. Available from: https:// www.herdin.ph/index.php?view=research&cid=37347.
39. Miura S, Saavedra OL, Yamamoto S. Osteoporosis in urban post-menopausal women of the Philippines: prevalence and risk factors. Vol. 3, Archives of Osteoporosis. 2008. p. 17–24.
40. Mendoza ES, Lopez AA, Valdez VAU, Mercado-Asis LB. Osteoporosis and Prevalent Fractures among Adult Filipino Men Screened for Bone Mineral Density in a Tertiary Hospital. Endocrinol Metab (Seoul, Korea). 2016 Sep;31(3):433–8.

MERCK

ALWAYS CURIOUS

IMAGINE

THE NEXT 350 YEARS

^{99m}Technetium Pertechnetate Thyroid Scan for Remnant Thyroid Tissue Detection Among Post-Thyroidectomy Patients at Jose R. Reyes Memorial Medical Center: A Retrospective Analysis

Athena Charisse S. Ong, MD, Marcelino A. Tanquilut, MD, Wenceslao S. Llauderes, MD, Emelito O. Valdez-Tan, MD, Ivan Ray F. David, MD

Department of Nuclear Medicine, Jose R. Reyes Memorial Medical Center
E-mail address: athenacharisse@gmail.com

ABSTRACT

This study aims to determine the diagnostic value of a ^{99m}Tc-pertechnetate (^{99m} TcO-4) thyroid scan among patients with DTC who underwent thyroidectomy to assess functioning thyroid remnants before radioactive iodine therapy. A retrospective non-experimental cross-sectional design was done to compare the results of the ^{99m} TcO-4 thyroid scan with the patient's post-RAI scan. A review of all our patients' charts was done for eight years, and after excluding those that did not fit the criteria, 70 patients were included in the study. Data collected was analyzed on a "per patient" basis– where patients either had a "positive scan" or "negative scan", and on a "per lesion" basis – where every lesion's presence and size were compared on both modalities. ^{99m} TcO-4 thyroid scan in the "per patient" analysis showed a sensitivity of 73.91%, specificity of 100%, positive predictive value (PPV) or 100%, and accuracy of 74.29%, however, negative predictive value was determined to be 5.26%. In the "per lesion" analysis, the scan had a less favorable performance with the computed sensitivity of 61.69%, PPV of 94.93%, and accuracy at 59.41%. It was then concluded that 99mTc-pertechnetate scan may be useful in determining functioning remnant thyroid tissue and subsequent management of DTC patients after thyroidectomy, but must take note of its low negative predictive value.

***Keywords**: Thyroid cancer, ^{99m}Tc-pertechnetate scan, ^{131}I post-therapy scan, Thyroid remnants*

INTRODUCTION

Thyroid cancer cases have been on the rise and is estimated globally to have 586,202 new cases in 2020, where 43,464 have succumbed to the disease [1]. The increase in the trend of diagnosed thyroid cancer patients was suggested to have been contributed by increasing early detection with the use of evolving technology and surveillance [2]. In the Philippine Interim Clinical Practice Guidelines for the Diagnosis and Management of Well-Differentiated Thyroid Cancer released in 2021, thyroid cancer in the Philippines ranked as the 6th most common cancer, ranking 21st as the cause of cancer-related mortality [3].

Currently, the management for well-differentiated thyroid cancer is mainly surgical. This may be followed with a conservative approach (monitoring and surveillance) or adjunctive/ablative radioactive iodine therapy. The Philippine Interim CPG recommends the consideration of the post-operative disease status of the patient as the attending physician decides on the next steps in the management of the patient [3]. The 2015 ATA guidelines report that post-operative imaging can change clinical management as it can modify the status assessment based on the results of the scan. Should RAI be considered, a pre-ablation low-dose ^{131}Iodine whole body scan is ideal to guide the clinician in prescribing the dose, along with other diagnostics such as neck ultrasound, serum tumor markers, and thyroid stimulating hormone [4].

Unfortunately, the low-dose ^{131}Iodine scan comes with its issues. Multiple studies have shown that diagnostic radioactive iodine is associated with an increased risk of ablation failure or "stunning" as mentioned in the 2015 ATA Guidelines [4]. A study by Park and colleagues showed that in patients who underwent diagnostic 131Iodine scan (dose of 3 to 10 mCi), there were 20 out

26 patients that showed impairment of radioiodine uptake on post-therapy scans compared to the group that used ^{123}Iodine [5]. To address this, the guidelines recommend using a lower dose of ^{131}Iodine (1-3 mCi) or an alternative radioisotope such as ^{123}Iodine to minimize the risk of ablation failure [4]. At present, ^{123}Iodine is not locally available in the Philippines. In our institution, as an alternative to ^{123}Iodine, the ^{99m} TcO-4 thyroid scan was used to assess for functioning tissue remnants, although not particularly mentioned in the existing clinical guidelines. Given the pathophysiology of the well -differentiated tumor, its use appears plausible, with the additional benefit of its favorable cost and availability and the absence of possible "stunning." On literature review, other institutions have also employed ^{99m} TcO-4 thyroid scan with the same objectives and have also attempted to determine the accuracy of this scan in determining thyroid tissue remnants. Their data showed promising results, but due to their limited sample size, there is still a lack of robust evidence to support its reliability. Several of them have yielded values for sensitivity of 77-81% (patient-based) and 59-61% (lesion-based), with PPV of 100% (patient-based) and 81-99% (lesion-based) [6]. Other studies reported less impressive data, with a sensitivity of only 13% for ^{99m} TcO-4 thyroid scan – significantly inferior to a diagnostic radioactive iodine scan having 67% sensitivity [7].

Locally, there is no established data on the clinical utility of ^{99m} TcO-4 scan. As mentioned above, it is still being utilized as an alternative to an ^{131}Iodine diagnostic scan despite the scarcity of information concerning its accuracy. This study then aims to compare ^{99m}TcO-4 scan with the post-RAI scans and determine its sensitivity, specificity, positive predictive value, negative predictive value, and accuracy in detecting remnant thyroid tissue in well-differentiated thyroid cancer patients. In doing this, we hope to contribute relevant information concerning the utility of this easily accessible and inexpensive diagnostic test to guide our Nuclear Medicine physicians in their practice.

MATERIALS AND METHODS

Patients and Data Collection

A retrospective, non-experimental, cross-sectional study design was used for this study. A review of records of thyroidectomized patients who have undergone their first RAI remnant ablation or adjunctive therapy for the past eight years was done (January 2015 - August 2022). Included in the data pool are patients aged 19-90 years old with histopathologically proven well-differentiated thyroid cancer (Papillary Thyroid Cancer or Follicular Thyroid Cancer), excluding patients with known distant metastases. We excluded those with distant metastases since most of these patients almost always warranted subsequent RAI therapy, with their post-operative stratification being high risk [3,4]. This means that the presence or absence of residual disease in the neck area may have less gravity in the decision-making process in this population versus those who are at low to intermediate risk. Those included in the study must also have undergone one of the following operations: 1) Near -Total, 2) Total Thyroidectomy, and/or 3) Completion Thyroidectomy, since patients who have undergone partial thyroidectomy or lobectomy are not recommended to undergo RAI [3,4]. A total of 256 patient records were initially reviewed. After eliminating pediatric patients, those with high-risk stratification, and those with no existing record of ^{99m} TcO-4 thyroid scan or post-therapy scans, a total of 70 patients remained to be part of the analysis. Figure 1 illustrates the selection process of eligible records.

Patient Preparation and Tc-99m Pertechnetate Scan

Our patients had their post—thyroidectomy ^{99m}TcO-4

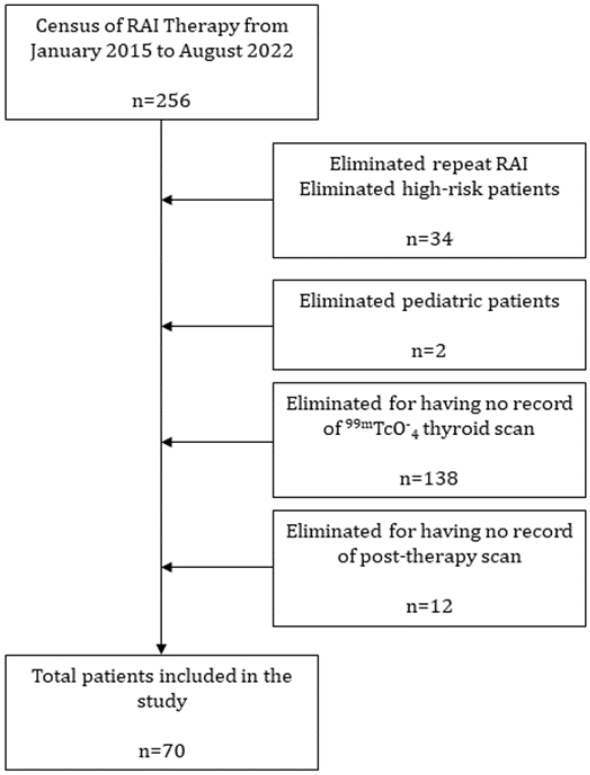


FIGURE 1. . Patient Selection Flow Chart

thyroid scan acquired at least 4-6 weeks after the operation. For patients who were already on thyroid hormone replacement therapy, the thyroid scan was further delayed for 4-6 weeks after cessation of the medication. The patients were also put on a low-iodine diet at this time. A dose of 259 MBq (7 mCi) of ^{99m}TcO-4 was given intravenously and was imaged 20 minutes after. Imaging was done using a Mediso Anyscan SPECT gamma-camera using a low energy high-resolution parallel-hole collimator at 20% energy window centered at 140 keV (frame size 256 X 256). Anterior, right anterior oblique, and left anterior oblique static views with markers on surgical scars and the sternal notch was acquired. The images were processed using the InterViewXP Clinical Processing System.

Post-Radioactive Iodine Therapy Whole Body Scan

Prior to therapy, TSH values were measured. At serum TSH of 30 mIU/L, they then proceeded to RAI therapy. A post-therapy whole body scan was done after 72-168 hours post oral administration of RAI. Our patients received doses ranging from 1850 to 5550 MBq (50-150 mCi) of ^{131}I. Whole-body scan acquisition was done using the Mediso Anyscan SPECT gamma-camera using a high energy parallel-hole collimator at 20% energy window centered at 364 keV (frame size 256 X 256). An additional view of the thyroid bed anteriorly using the high-energy parallel-hole collimator was also acquired.

Data Analysis

The images were again reviewed and blindly interpreted by us. The lesions were also individually remeasured blindly by 1) an experienced nuclear medicine technologist and 2) repeated by us. The average of the two values were used for data analysis.

The data gathered was analyzed on a "Per-Patient" and "Per-Lesion" basis. For the purposes of this study, the ^{131}I post-therapy scans of the patients were regarded as the "gold standard" to compute for the sensitivity, specificity, PPV, NPV, and accuracy of ^{99m}TcO-4 thyroid scan.

For the "Per-Patient" analysis, A ^{99m}TcO-4thyroid scan was deemed "positive" if at least one lesion was present in the scan. A lesion is any abnormal uptake in the neck area. The same method was used to identify a positive scan for a post-^{131}I post-therapy scan. The results of the ^{99m}TcO-4 thyroid scan were compared with the patient's radioactive iodine (RAI) post-therapy scan using a paired T-test. For the "Per-Lesion" analysis, lesion presence and size were compared in both scans. Lengths and widths (in centimeters) of the lesions were measured. Lesions that were not present in either scan had measurements of 0 cm. A paired T-test was applied to compare the length and width of ^{99m}TcO-4 and post-therapy scan results. Analysis was done using Medcalc Statistical software with a significance level set at 0.05

RESULTS

Seventy patients fulfilled the inclusion and exclusion criteria and were included in the study. The majority of the patients were females (77.1%). Most of our patients were post-operatively stratified as low risk, and RAI therapy doses ranged from 1,850 MBq to 5,550 MBq, as seen in Table 1.

Table 2 shows the 2 x 2 contingency table of the patients' thyroid and post—therapy scans. A total of 69 patients presented with positive post - therapy scans. Among these, 51 patients concordantly had positive thyroid

TABLE 1. Demographical and Clinical Characteristics

Age (years), mean ± sd	46.5 ± 13.7	
Sex, n, %		
Male	16 (22.9%)	
Female	54 (77.1%)	
Risk Stratification	**RAI Dose in MBq (mCi)**	**n, %**
Low-risk	1,850 (50)	10 (14.3%)
	3,700 (100)	46 (65.7%)
Intermediate-risk	5,180 (140)	2 (2.86%)
	5,550 (150)	12 (17.14%)

scans (true positive), and 18 had negative thyroid scans (false negative). One patient demonstrated both negative ^{99m}TcO-4 and ^{131}I scans (true negative).

The sensitivity, specificity, PPV, NPV, and accuracy of the ^{99m}TcO-4 thyroid scan in detecting the presence of thyroid remnants are shown in Table 3. The data reveals that the sensitivity of the 99mTcO-4 thyroid scan in predicting a positive 131I Scan is 73.91%, specificity of 100%, positive predictive value of 100%, negative predictive value of 5.26%, and accuracy of 74.29%.

Table 4 shows the 2 x 2 contingency table for the lesions that were detected in both ^{99m}TcO-4 and ^{131}I post-therapy scans. The ^{99m}TcO-4 thyroid scan detected 107 lesions, while 164 were noted in the ^{131}I post-therapy scans. One hundred-one lesions were detected in both scans (true positives), with a congruence of 101/170 (59.4%).

Table 5 shows the diagnostic accuracy parameters analyzed on a per-lesion basis. The sensitivity and specificity of ^{99m}TcO-4 thyroid scan on a per lesion basis are 61.69% and 0%, respectively. A high PPV was noted at 94.39% and an NPV of 0%. Of note are the six lesions that were present in the ^{99m}TcO-4 thyroid scan but not in the post-therapy scan. The computed accuracy was 54.91%. The specificity and NPV of 0 in this instance are irrelevant as it is impossible to quantify lesions that truly do not exist.

To measure if there is a significant difference in the dimensions of lesions detected in both scans, only lesions that were present in both the ^{99m}TcO-4 thyroid scan and ^{131}I Post-Therapy Scan were analyzed. Table 6 reveals that there was a significant difference in the lengths obtained by ^{99m}TcO-4 thyroid scan (M=1.98, SD=0.98) and lengths measured in the post-therapy scans (M=2.70, SD=2.49, $p < 0.05$). Similarly, there is a significant difference in the width of lesions measured in ^{99m}TcO-4 thyroid scan (M=1.48, SD=0.69) compared to the widths measured in the post-therapy scans (M=2.26, SD=0.88, $p < 0.05$). These results suggest that the dimensions of the lesions measured are significantly larger in the post-therapy scans compared to the ^{99m}TcO-4 thyroid scan.

DISCUSSION

Per Patient Analysis

Our data shows that the ^{99m}TcO-4 thyroid scan is 74.29% accurate in determining thyroid remnants in post-thyroidectomy patients. Its computed sensitivity is at

TABLE 2. Contingency Table for Per Patient Analysis

	Positive ^{131}I Scan	*Negative ^{131}I Scan*	*Total*
Positive $^{99m}TcO^-_4$ Scan	51 (TP)	0 (FP)	51
Negative $^{99m}TcO^-_4$ Scan	18 (FN)	1 (TN)	19
Total	69	1	70

TABLE 3. Diagnostic accuracy of ^{99m}TcO-4 thyroid scan (Per Patient)

		95% CI
Sensitivity	51/69 (73.91%)	(61.94 to 83.75%)
Specificity	1/1 (100%)	(2.50 to 100%)
PPV	51/51 (100%)	
NPV	1/19 (5.26%)	(3.60 to 7.63%)
Accuracy	74.29%	(62.55 to 83.99%)

TABLE 4. Contingency table for the Per Lesion analysis

	Positive in ^{131}I Scan	*Negative in ^{131}I Scan*	Total
Positive in $^{99m}TcO^-_4$ Scan	101 (TP)	6 (FP)	107
Negative in $^{99m}TcO^-_4$ Scan	63 (FN)	0 (TN)	63
Total	164	6	170

TABLE 5. Diagnostic accuracy of ^{99m}TcO-4 thyroid scan (Per Lesion)

		95% CI
Sensitivity	61.69%	53.68 to 69.09%
Specificity	0	0 to 45.93%
PPV	94.39%	93.72 to 95%
NPV	0	
Accuracy	59.41%	51.63 to 66.86%

TABLE 6. Comparison of lengths and widths of lesions present in both ^{99m}TcO-4 thyroid scan and ^{131}I post-therapy scan

	n	*Tc-99m Thyroid Scan*			*Post-Therapy Scan Result*			*p value*
		Mean	SD	Median	Mean	SD	Median	
Overall Length	101	1.98	0.97	1.72	2.70	0.95	2.49	0.0001
Overall Width	101	1.48	0.69	1.37	2.26	0.88	2.13	0.0001

73.91%, which does not fare differently from the results of similar studies, which showed sensitivities of 72.2%-81% [6,8,9]. The positive predictive value of 100% is likewise similar or identical to the results of the same publications. Our data suggests that a fair number of patients with positive ^{131}I post-therapy scans can be predicted by the ^{99m}TcO-4 thyroid scan among patients with low to intermediate-risk DTC.

The specificity of ^{99m}TcO-4 thyroid scan was computed to be 100%. Our data was significantly greater compared to another study, which showed specificity estimation of 70.5% [8]. However, it is important to consider that only one was truly negative, and was the only value used to compute for the specificity. Similar studies conducted by Kueh [6] and Tsai [9] could not produce estimates for specificity since their study samples all had positive post-therapy scans. NPV was likewise not possible to compute in their studies.

Our negative predictive value was computed to be 5.26%. This was expected given that there were 18 (26.1%) patients with negative ^{99m}TcO-4 thyroid scans that were positive in their post-thyroidectomy scans. This shows then that a negative ^{99m}TcO-4 thyroid scan is unlikely to indicate the absence of a thyroid remnant tissue.

Per Lesion Analysis

It was found that the ^{99m}TcO-4 thyroid scan had a sensitivity of 61.69% and a positive predictive value of 94.39%. This indicates that only a little more than half of the lesions can appear in the ^{99m}TcO-4 thyroid scans, but its presence in the ^{99m}TcO-4 thyroid scan almost guarantees its uptake in the subsequent post-therapy scan. A representative image from a patient in Figure 2 demonstrates lesions not detected in the ^{99m}TcO-4 thyroid scan, which then appeared in the post-therapy scan.

This was almost similar to the results of Tsai [9] which were 59% (sensitivity) and 100% (PPV). The specificity and negative predictive value were both 0, as expected. A limitation in this analysis was the absence of lesions in our gold standard that can only be quantified as 0 and would mathematically give us no value. Interestingly, six false-positive lesions were identified in the ^{99m}TcO-4 thyroid scan, which was not identified in the post—thyroidectomy scan. A study by Long similarly encountered false positive lesions in their study, which also compared a ^{99m}TcO-4 scan with post-therapy scans.

One focus of ^{99m}TcO-4 uptake was seen in the axillary lymph nodes and was speculated to be lymphatic drainage of the radiotracer from the injection site [10]. This might not be a plausible explanation for our case since we only assessed the neck area, and it is unlikely that the tracer will concentrate in the cervical nodes. It may be possible that interpretation error may be involved, such as physiologic ^{99m}TcO-4 uptake in the salivary glands, or may also be explained by increased tracer accumulation secondary to an inflammatory process [9].

The lesion dimensions measured in the 99mTcO-4 scan and post-therapy scans were also compared, showing a

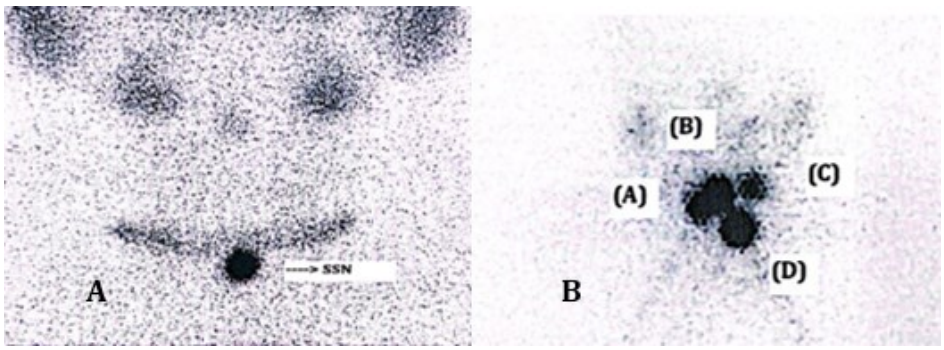


FIGURE 2. False negative lesions. **A.** $^{99m}TcO^{-}_{4}$ thyroid scan shows a mild uptake in the midline; **B** Post-RAI scan in the same patient showing four foci of intense iodine uptake in the neck

significant difference in the lesion length and width. Although statistically different, scintigraphic images are unreliable in measuring organ volume. A communication by Tanahill revealed that thyroid scintigraphy thyroid measurement using ^{99m}TcO-4 did not correlate to ultrasound, surgical specimens, or clinical palpation [11]. This disagreement on size measurement may be operator-dependent but may be addressed by concomitant anatomic imaging such as SPECT/CT. Size measurement is vital as it can impact subsequent management. To illustrate, a study comparing thyroid volume measurement using radioiodine versus ultrasonography found that scintigraphic volume estimation using the Himanka formula overestimated the thyroid volume by 53% in patients with diffuse goiter dose. The variability in computed volume using scintigraphy also resulted in differences in therapeutic radioiodine dose [12].

Limitations and Pitfalls

The ^{99m}TcO-4 scintigraphy may show uptake of remnants due to its ability to be taken up by functioning thyroid tissue through the sodium iodide transporters and trapped without proceeding to organification. It is also more accessible and has more favorable imaging characteristics (i.e., half-life of 6 hours and a photopeak of 140 keV) compared to the ^{131}Iodine [13]. However, our data has shown its limitations in its ability to demonstrate all functioning thyroid tissue, and while focal uptake in the thyroid bed is considered a remnant, certainty of its histology is not certain due to the lack of corresponding anatomical imaging. In this case, the addition of SPECT/CT may give us more valuable information in the same way it has in other literature. A study conducted by Chantadisai, which investigated the usefulness of a whole body ^{99m}TcO-4 imaging with SPECT/CT to detect remnants and metastasis, has shown that out of the 111 positive foci of ^{99m}Tc-pertechnetate, 106 were seen in the thyroid bed, two foci in the lymph nodes and one bone lesion [14]. Another study was likewise able to identify extrathyroidal ^{99m}Tc-pertechnetate uptake, with high sensitivity for regional nodal metastasis but low sensitivity for distant metastasis. SPECT/CT for patients with equivocal ^{99m}Tc-pertechnetate scans was done and ultimately confirmed mediastinal uptake in 2 patients and only physiologic esophageal activity in 2 other patients [15]. Utilization of SPECT/CT is recommended if ^{99m}Tc-pertechnetate thyroid scan is contemplated, and especially if a ^{99m}Tc-pertechnetate whole body scan being considered.

We focused mainly on thyroid remnant tissue and did not attempt to evaluate distant metastases, so we only included patients with low to intermediate-risk stratification. Our investigation is also limited to well-differentiated thyroid carcinoma, and patients demonstrating dedifferentiation will entail an entirely different study.

Impact on Management

The ^{99m}Tc-pertechnetate scan may help determine the following steps after confirming functioning thyroid remnant post-thyroidectomy. Although remnant ablation in low-risk patients is not always recommended, it was reported that more than 90% of post-thyroidectomy patients across the Asia-Pacific region are treated with radioactive iodine ablation therapy [16]. This may be due to the higher risk of recurrence in this population, resulting in a more aggressive approach. It was also reported that Filipinos are more at risk for disease recurrence [3]. A retrospective study by Lo concluded that Filipinos are associated with a more aggressive papillary thyroid cancer disease course. However, follicular thyroid cancer patients did not have any significant risk compared to other populations. It was also found that their patients who underwent RAI

therapy, along with surgical management and TSH suppression therapy achieved a disease-free status on long-term follow-up.[17] With this knowledge, identifying the presence of thyroid remnants can aid the physician in weighing the risk and benefits of an ablative RAI .

CONCLUSION

^{99m}Tc pertechnetate thyroid scan has a high sensitivity and predictive value for determining thyroid remnants on per-patient and per-lesion analyses. However, its low negative predictive value warrants cautious utilization of the ^{99m}Tc pertechnetate thyroid scan, and must not be complacent about the absence of remnants in a negative thyroid scan. The significant difference in lesion size measurements between the ^{99m}Tc pertechnetate thyroid and post-therapy scans suggests that this may not accurately measure the remnants. To address the limitations of ^{99m}Tc pertechnetate thyroid scan, we recommend the addition of SPECT/CT and further exploring its utility in determining regional and distant metastases. Overall, the ^{99m}Tc-pertechnetate scan is useful in determining functioning remnant thyroid tissue and subsequent management of DTC patients after thyroidectomy.

FUNDING

The authors received no financial support for this article's research, authorship, and/or publication.

REFERENCES

1. International Agency for Research on Cancer. (2020). Thyroid. Fact Sheet, World Health Organization.
2. American Cancer Society. (2020). Cancer Statistics Center. Retrieved July 26, 2020, from https:// cancerstatisticscenter.cancer.org/
3. DOH. (2021). The Philipine Interim Clinical Practice Guidelines for the Diagnosis and Management of Well-Differentiated Thyroid Cancer 2021. Department of Health. DOI: https://doi.org/10.47895/amp.vi0.6535.
4. Haugen, B. R. (2016). 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 : official journal of the American Thyroid Association, 1-133. PMID: 26462967.
5. Park, H. P. (1994). Influence of diagnostic radioiodines on the uptake of ablative dose of iodine-131. Thyroid, 49-54. PMID: 8054861.
6. Kueh, S. S. (2010). Role of Tc-99m pertechnetate for remnant scintigraphy post-thyroidectomy. Clinical nuclear medicine, 671-674. PMID: 26462671.
7. Liu, M. C. (2017). 99mTc-pertechnetate-avid metastases from differentiated thyroid cancer are prone to benefit from 131I therapy: A prospective observational study. Medicine, 1-7. PMID: 28816945 PMCID: PMC5571682.
8. Ozdemir, D. C. (2016). The role of post-operative Tc-99m pertechnetate scintigraphy in estimation of remnant mass and prediction of successful ablation in patients with differentiated thyroid cancer. Nuclear medicine communications, 640-645. PMID: 26895488.
9. Tsai, C. J. (2016). Tc-99m imaging in thyroidectomized differentiated thyroid cancer patients immediately before I-131 treatment. Nuclear medicine communications, 182-187. PMID: 26626550.
10. Long, B.L.F. (2021). Clinical significance of extra-thyroid 99mTc pertechnetate uptake before initial radioiodine therapy for differentiated thyroid carcinoma. Journal of International Medical Research, 1-10. PMID: 34024177.
11. Tanahill, A. H. (1978). Measurement of thyroid size by ultrasound, palpation, and scintiscan. Clincal Endocrinology, 483-486. PMID: 668154.
12. Wesche, M. T. (1998). Ultrasonographic versus scintigraphic measurement of thyroid volume in patients referred for I-131 therapy. Nuclear Medicine Communication, 341-346. PMID: 9853324.
13. Mettler, F.A. & Guiberteau, M.J., (2019). Thyroid, Parathyroid, and Salivary Glands, Essentials of Nuclear Medicine and Molecular Imaging (Seventh Edition), Elsevier, 2019, 85-115.
14. Chantadisai, M., & Kingpetch, K. (2014). Usefulness of (99m)Tc-pertechnetate whole body scan with neck and chest SPECT/CT for detection of post-surgical thyroid remnant and metastasis in differentiated thyroid cancer patients. Annals of nuclear medicine, 28(7), 674–682. PMID: 24889127.
15. Lou, K., Gu, Y., Hu, Y., Wang, S., & Shi, H. (2018). Technetium-99m-pertechnetate whole-body SPET/CT scan in thyroidectomized differentiated thyroid cancer patients is a useful imaging modality in detecting remnant thyroid tissue, nodal and distant metastases before 131I therapy. A study of 416 patients. Hellenic journal of nuclear medicine, 21(2), 121–124. PMID: 30089313.
16. Sundram, F., Robinson, B.G., Kung, A., Lim-Abrahan, M.A., Bay, N.Q., Chuan, L.K., Chung, J.H., Huang, S., Hsu, L., Kamaruddin, N., Cheah, W.K., Kim, W.B., Koong, S., Lin, H.D.,Mangklabruks, A., Paz-Pacheco,A., Rauff, A., & Ladenson, P.W. "Well-Differentiated Epithelial Thyroid Cancer Management in the Asia Pacific Region: A Report and Clinical Practice Guideline." Thyroid, 2006: 461-469. PMID: 16756468.
17. Lo, T. E., Uy, A. T., & Maningat, P. D. (2016). Well-Differentiated Thyroid Cancer: The Philippine General Hospital Experience. Endocrinology and metabolism (Seoul, Korea), 31(1), 72–79. https://doi.org/10.3803/EnM.2016.31.1.72

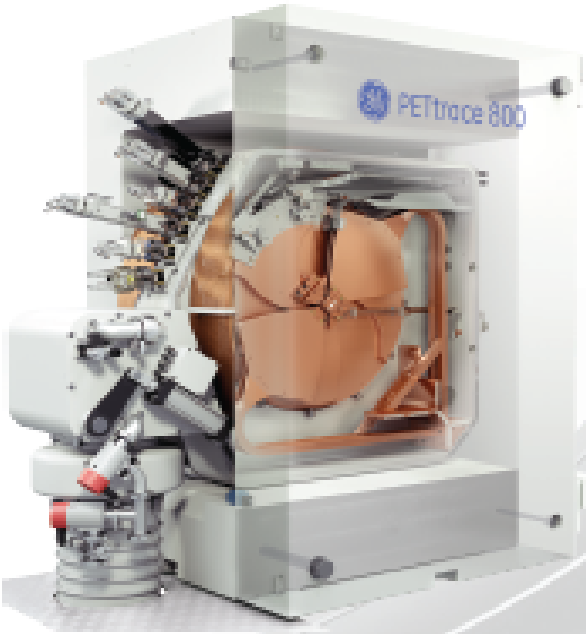
Connecting Every Step in Molecular Medicine

DISCOVERY	DIAGNOSIS	TREATMENT
for development of **targeted** tracers and therapy	with **breakthrough** PET and SPECT imaging solutions	using **predictive** and advanced monitoring tools

We are focused on core challenges

At GE Healthcare we believe we can uniquely help connect the teams, data, and decisions in every step from discovery to diagnosis to treatment with intelligent efficient innovations that will ultimately help you deliver precise, personalized care. This is our vision.

There is so much data available across molecular imaging today. There are so many connection opportunities and loops to close. That's where we see significant opportunity. And we believe we're in a unique position to fulfill this vision because we are the only vendor with pharmaceutical diagnostics, cyclotrons, chemistry synthesis, PET/CT, SPECT/CT, advanced digital solutions and pharma partnerships to cover the breadth of steps from discovery to treatment.



Please contact your local GE Healthcare representative to learn more or visit: https://www.gehealthcare.com.sg/

Role of Early Dynamic PET/CT Scan Imaging with ^{18}F-PSMA-1007 in Staging and Restaging Prostate Cancer in a Tertiary Private Hospital

Arrene Joy B. Baldonado, MD

Department of Radiological Sciences, Section of Nuclear Medicine and PET-CT center, Cardinal Santos Medical Center, E-mail address: baldonadoarrene@gmail.com

ABSTRACT

Introduction:

^{18}F-PSMA-1007 is a novel prostate-specific membrane antigen (PSMA)-based radiopharmaceutical for imaging prostate cancer. The recommended imaging time is 60 minutes post-injection of the radiotracer. However, during this time there is a physiologic accumulation of the radiotracer in the urinary bladder which sometimes may obscure lesions adjacent to it.

Objective:

This study aims to determine if early dynamic imaging in addition to the recommended 60-minute post-injection static imaging can improve the detection of PSMA-avid lesions in the staging and restaging of prostate cancer.

Methods:

This is a retrospective cross-sectional study of the detection rate of early dynamic and static imaging using ^{18}F-PSMA-1007 PET/CT scan in patients with prostate cancer (PCa) who were referred for initial staging or restaging. The McNemar test was used to compare the detection rate between the two imaging. Spearman correlation was used to determine the correlation of Gleason score (GS), PSA, and SUVmax values.

Results:

^{18}F-PSMA-1007 PET/CT scans of 53 patients with prostate cancer, were referred for either staging (22/53) or restaging (31/53), all of whom had undergone both early dynamic and static imaging. Among the 53 patients, 5 had 2 lesions each, for a total of 58 lesions were included in the analysis. There were 48/58 lesions detected on both early dynamic and static imaging, 2/58 lesions were only detected in the early imaging, 1/58 lesions was only detected in the static imaging, and 7/58 were not detected on both imaging. McNemar the test was not statistically significant (p = 1.000) in the detection rate of the two methods. There is a positive correlation between serum PSA levels and SUVmax measurements for all the patients. Only the correlation between the GS and SUVmax in the static imaging of the staging group was statistically significant.

Conclusion:

Early dynamic imaging may be an adjunctive procedure in detecting PSMA-avid lesions, particularly in the basal segment of the prostate gland near the urinary bladder. However, it is not recommended as a standard component of the comprehensive protocol for imaging using ^{18}F-PSMA-1007 PET/CT in patients with PCa.

Keywords: *^{18}F-PSMA-1007 PET/CT scan, dynamic imaging, static imaging, prostate cancer*

INTRODUCTION

Prostatic malignancy is the fourth most common cancer worldwide, and the third leading malignancy site in males in the Philippines, with an age-standardized incidence rate (ASR) of 21.9 per 100,000 population [1, 2]. The gold standard for the diagnosis of prostate carcinoma is a histological assessment usually obtained by transrectal ultrasound-guided systematic core needle biopsy [3]. An important histopathologic parameter is the Gleason Score (GS) which reflects the grade of differentiation of prostate cancer (PCa) and thus correlates with tumor aggressiveness [4]. The standard treatment of PCa involves prostatectomy and radiotherapy. Although this treatment is curative for some, 20 to 30% experience a recurrence typically detected when there is a rise in serum prostate-specific antigen (PSA) levels after initial treatment [5, 6, 7, 8].

The role of conventional imaging modalities (CT, MRI, 99mTc-MDP bone scan) is limited in PCa recurrence assessment, as well as in the detection of nodal and distant disease [9, 10, 11, 12]. Due to this, there is a need to develop new diagnostic methods to allow an accurate means of staging and restaging of PCa. In recent years, Positron Emission Tomography (PET) with fluorodeoxyglucose and choline-based radiotracers have been introduced for the diagnosis and staging of PCa with the eventual development of targeted imaging using prostate-specific membrane antigen (PSMA) [13, 14, 15]. PSMA is a transmembrane glycoprotein over-expressed in prostate cancer cells and shows low expression in benign prostatic tissue [16]. It has been used in various clinical management of PCa, such as staging primary tumors, localization of biochemical relapse, planning for radiotherapy, prediction, and assessment of treatment response. Several radiolabeled PSMA probes were developed including the most widely used Gallium-68 Prostate Specific Membrane Antigen-110 (^{68}Ga-PSMA-110 [15, 17, 18, 19]. It is superior to conventional imaging and choline-based PET/CT for evaluating PCa patients with biochemical recurrence but also for staging purposes [20]. However, the disadvantage of ^{68}Ga-PSMA as a radiotracer is its high accumulation in the urinary bladder which may influence the uptake evaluation of the prostate bed [21]. The introduction of a novel PSMA - based radiopharmaceutical, Fluorine-18 Prostate Specific Membrane Antigen-1007 (^{18}F-PSMA-1007), offers several advantages over ^{68}Ga-PSMA [14, 15]. It is primarily excreted in the hepatobiliary tract and shows a relatively lower urinary bladder activity, hence it can be used in evaluating cases of local tumor recurrence and unclear lesions adjacent to the urinary bladder [15, 19]. Moreover, ^{18}F-PSMA has a longer half-life (T1/2 = 109 min), higher physical spatial resolution, and larger dose since it is produced by a cyclotron as compared to Gallium-68 which is derived from elution of 68Ge/68Ga generators [15]. The larger dose produced via cyclotron leads to a greater number of patients that can be accommodated. Studies show that ^{18}F-PSMA is useful in staging PCa and positively correlates the SUV values of PSMA-avid lesions with PSA level and GS [14, 15, 22].

In the Philippines, several institutions are utilizing ^{18}F-PSMA-1007 PET/CT scans in the diagnosis of PCa namely: Cardinal Santos Medical Center (CSMC), Chinese General Hospital and Medical Center (CGHMC), Centuria Medical Makati (CMM), iScan Diagnostic Center, National Kidney and Transplant Institute (NKTI), and The Medical City (TMC); all of which are located in Metro Manila. Amongst the said institutions, only the CSMC performed early dynamic imaging in ^{18}F-PSMA-1007 PET/CT scans from August 2020 to March 2022. With the current clinical data, there is no standard protocol in the imaging time of ^{18}F-PSMA-1007 published, however, studies recommend static imaging ranging from 40 to 90 minutes post-injection, most commonly used is 60 minutes, to allow for radiotracer uptake [15, 22, 23, 24, 25, 26].

Dynamic PET/CT scan is a modality that allows registration of pharmacokinetic information over time, while classical static whole-body PET/CT protocols enable the acquisition of patient images only at a one-time point after tracer injection [27]. In the dynamic PET/CT scan, studies have shown that in the first few minutes post-injection, radiotracer uptake of local prostate cancer lesions is visible before its accumulation in the urinary bladder [16]. Hence, the use of dynamic imaging can aid in the detection of PSMA-avid lesions within the proximity of the urinary bladder. When viewed in static imaging, these lesions can be obscured by the physiologic activity in the bladder. A study by Barakat et al. demonstrated that early dynamic imaging using ^{68}Ga-PSMA PET/CT scan increases the detection of PSMA-avid lesions in the anterior portion of the prostate and is suggestive of prostate cancer or its recurrence [28]. There is limited published data on the diagnostic efficacy of ^{18}F-PSMA-1007 PET/CT scan and no study on its use with early dynamic imaging. This study aims to show the detection rate of ^{18}F-PSMA PET/CT scan in early dynamic imaging as an adjunct to the standard 60-minute post-injection static imaging, in the staging and restaging of prostate cancer. Furthermore, to analyze

the correlation between the GS and PSA value with the SUV level of the lesions.

MATERIALS AND METHODS

Ethical approval:

This study was approved by the Ethics Review Board (ERB) of CSMC with RERC CODE 2021-023. The need for written informed consent was waived by the ERB due to the study's use of the retrospective method of data collection. This study was conducted in strict compliance with the provisions of the Philippine Data Privacy Act of 2012 (Republic Act of No. 10173). All imaging procedures performed were following the tenets of the Declaration of Helsinki and its amendments.

Patient:

A pilot sample of 53 patients was retrospectively analyzed in this study. Patients with biopsy-proven prostate carcinoma, who underwent both early dynamic and standard static ^{18}F-PSMA-1007 PET/CT scans between August 2020 to March 2021 in a single-center hospital (Cardinal Santos Medical Center) were included. All the patients were referred for baseline staging or restaging of prostate cancer. Patients with multiple primary cancer, not-biopsy-proven prostate cancer, and those with technically inadequate studies (e.g. motion artifacts, incomplete study) were excluded. Eligible data such as the patient's age, GS, serum PSA levels (ng/mL), previous and ongoing treatments (radical prostatectomy, radiotherapy, hormonal therapy), and ^{18}F-PSMA-1007 PET/CT findings were assessed. The scans were assigned to two groups according to clinical purpose: (1) staging of disease and (2) restaging. Patients who underwent ^{18}F-PSMA PET/CT scans before initiation of treatment were classified under the "staging group" while patients who underwent PET/CT scans after treatment were classified as the "restaging group". Of 53 patients, 22 were included in the staging group and 31 in the re-staging group.

Radiotracer:

^{18}F-PSMA-1007 radiotracers provided by the Khealth Corporation were utilized in this study.

Image acquisition:

Each patient received 148-444 MBq of ^{18}F-PSMA intravenously depending on the computed dose base on the body weight. Data acquisition consists of two parts: early dynamic imaging which is followed by static imaging (whole-body PET/CT). Early dynamic studies were performed over the lower abdomen to the pelvic area and acquired at the time of injection up to 6 minutes post-injection; while the static (whole-body) imaging was acquired at 60 minutes post-injection. Patients were asked to urinate before both imaging modalities.

^{18}F-PSMA-1007 PET/CT acquisition was performed on a GE Discovery 71.0 scanner. This system has a time-of-flight (TOF) capable technology with a full three-dimensional PET and a 64-slice CT. PET acquisition time was at 2 minutes per bed position for the whole body and 0.5 to 1 minute per bed position for the lower extremities using the TOF-PET technique. The exact CT parameters used for unenhanced acquisition include 0.8 pitch, 0.75-second rotation time, and effective tube current-time product range of 50-300 milliamperes (mA); these were dependent on the body thickness and tube voltage of 120 kilovoltage peak (kVp). Finally, image reconstruction was performed at a slice thickness of 3.75 mm for PET/CT.

Image Evaluation:

Images were interpreted using the dedicated commercially available Autonomous Database Warehouse (ADW) Linux software which provides PET, CT, and fused PET/CT imaging data in the axial, coronal, and sagittal planes. All PET imaging was attenuation corrected (AC) and had undergone Q-clear technology using the GE Discovery 7.10 scanner. The images were reconstructed and the maximum standardized uptake value (SUVmax) of the detected focal lesions was measured. The SUVmax measured in the early dynamic and static images was used for the quantification of tracer data. The visible prostate cancer lesions in the early dynamic and static images were independently reviewed, qualitatively, and quantitatively scored by two nuclear medicine physicians and two radiologists. All disagreements in the interpretation of the results in the provided images were resolved through consensus. Any focal uptake in the prostate or prostate bed with SUVmax greater than 2.5 g/ml or greater than the background was considered pathologic and suggestive of malignancy. In this study, such lesions were categorized as 'positive' while lesions that did not show an increase in tracer uptake in comparison to the surrounding tissue, or those with undetectable lesions were categorized as 'negative'.

Statistical Analysis:

Age was reported as mean ± standard deviation (SD). Serum PSA and SUVmax were reported as median with

interquartile range (IQR). Stata 16.1 software was used for data processing and analysis. Continuous variables based on data distribution were presented as mean ± SD or median IQR. Categorical variables were reported as frequencies and percentages. The McNemar test was used to compare the detection rate between early dynamic and static imaging. Shapiro-Wilk test was utilized to test whether the data provided in the GS, serum PSA and SUVmax levels were normally distributed. The null hypothesis of the Shapiro-Wilk test was used as a reference in assessing if the data collected were normally distributed. A significance level of 0.05 was used as the threshold to determine whether to accept or reject the null hypothesis. The result of the Shapiro-Wilk test showed that the data provided on the above-mentioned parameters were not normally distributed hence, Spearman correlation was performed. The following were used to determine the correlation of GS and serum PSA with the SUVmax of the lesions: 0-0.10 (negligible correlation), 0.10-0.39 (weak correlation), 0.40-0.69 (moderate correlation), 0.70-0.89 (strong correlation), and 0.90-1.00 (very strong correlation). P values less than 0.05 were considered statistically significant.

RESULTS

Out of the 53 patients, 22 (41.51%) were referred for initial staging and 31 (58.49%) for restaging. The same 22 patients had not received any treatment at the time of the study. Among those who had undergone treatment, 5 had hormonal therapy only, 14 had radical prostatectomy only, 7 had radiotherapy only, and 5 had received two of the previous treatments. The mean age of all the patients was 68.72 years with an interquartile range (IQR) from 63 to 76 years old. The median serum PSA level of all patients is 11.90 ng/mL (IQR: 2.20 to 68.50 ng/mL), while the median GS is 7.00 (IQR: 6 to 8) (Table 1).

On one hand, the staging group had a median GS of 7 (IQR: 6.00 – 9.00), median serum PSA of 16.68 ng/ml (IQR: 10.10 – 51.83), median early dynamic imaging SUVmax of 5.3 (IQR: 4.75 – 7.55) and median static imaging SUVmax of 25.6 (IQR: 22.43 – 43.20). On the other hand, the restaging group had a median GS of 7 (IQR: 6.00 – 8.00), median serum PSA of 6.21 ng/ml (IQR: 0.80 – 68.50), median early dynamic imaging SUVmax of 4.4 (IQR: 3.45 – 5.15) and median static imaging SUVmax of 11.0 (IQR: 5.80 – 26.05), respectively (Table 2).

It must be noted that although there were 53 patients, 5 patients had two lesions each, making a total of 58 lesions included in the analysis (Table 3).

Out of the 58 lesions in the sample, 48 (84.48%) were detected by both early dynamic and static imaging, 2 (3.45%) were detected by early dynamic imaging and not seen on static imaging, 1 lesion (1.72%) was not appreciated on early dynamic imaging but was observed on static imaging, while 7 lesions (12.07%) were not observed on both early dynamic and static imaging (Table 3). McNemar's test showed that there is no statistical difference in the detection rate of lesions on both early dynamic and static imaging.

From the staging group, 3 out of the 22 patients had two lesions each, with a total of 25 lesions. Out of the 25 lesions, 23 (92%) were detected by both early dynamic and static imaging, and 2 (8%) were detected by early dynamic imaging but were undetected by static imaging. In total, all of the lesions in the patients from the staging group were detected by early dynamic imaging (Table 4). McNemar's test shows no significant difference in the detection rate of both imaging.

Thirty-three lesions were included in the restaging group; of which, 25 (75.76%) were detected by both early dynamic and static imaging, 1 (3.03%) was detected by static imaging but undetected by early dynamic imaging, and 7 (21.21%) were undetected by both methods. In total, 25 out of the 33 (75.76%) lesions were detected by early dynamic imaging, which is slightly lower (78.79%) than the detection rate of static imaging (Table 5). McNemar's test shows that there is no significant difference in the detection rate of the two methods.

In the Shapiro-Wilk test, the GS, PSA, SUVmax of early dynamic imaging, and SUVmax of static imaging are not normally distributed ($p = 0.0000$).

Based on the results in Table 6, only the correlation between GS and SUVmax levels using static imaging for patients in the staging group is statistically significant. The correlation implies that there is a positive and moderate correlation between GS and levels of SUVmax for static imaging of patients in the staging group. The rest of the parameters for GS correlation are statistically insignificant (weak).

For the serum PSA (Table 6), only the correlation between the values of serum PSA and SUVmax in the restaging group of patients on early dynamic and static imaging, as well as in the staging group for the early

TABLE 1. Clinical profile and indication

		Mean (±Standard Deviation)	Median	Interquartile Range
Age (years)		68.72 (± 9.73)	69.00	63.00 – 76.00
Serum PSA (ng/ml) [1]		45.72 (± 91.23)	11.90	2.20 – 68.50
Gleason Score[2]		7.25 (± 1.46)	7.00	6.00 – 8.00
Profiles			**Frequency**	**Percentage**
Indication				
	Staging		22	41.51%
	Restaging		31	58.49%
Treatment				
	Hormonal Therapy only		5	9.43%
	Radical Prostatectomy only		14	26.42%
	Radiotherapy only		7	13.21%
	Hormonal Therapy and Radical Prostatectomy		3	5.66%
	Hormonal Therapy and Radiotherapy		2	3.77%
	Radical Prostatectomy and Radiotherapy		–	–
	Hormonal Therapy, Radical Prostatectomy, and Radiotherapy		–	–

[1] No PSA level data were obtained from four (4) patients.
[2] No Gleason Score data were obtained from nine (9) patients.

TABLE 2. Gleason Scores, PSA Levels, and SUV Levels of patients per indication

Profiles	Staging Group	Restaging Group
PSA level (ng/ml)		
Mean (± Standard Deviation)	35.75 (± 36.73)	52.60 (± 115.03)
Median	16.68	6.21
Interquartile Range	10.10 – 51.83	0.80 – 68.50
Gleason score		
Mean (± Standard Deviation)	7.38 (± 1.24)	7.13 (± 1.66)
Median	7.00	7.00
Interquartile Range	6.00 – 9.00	6.00 – 8.00
SUV level – Early Dynamic Imaging		
Mean (± Standard Deviation)	6.99 (± 4.08)	4.60 (± 1.52)
Median	5.30	4.40
Interquartile Range	4.75 – 7.55	3.45 – 5.15
SUV level – Static Imaging		
Mean (± Standard Deviation)	36.51 (± 28.91)	31.05 (± 87.33)
Median	25.60	11.00
Interquartile Range	22.43 – 43.20	5.80 – 26.05

TABLE 3. Number of lesions detected on early dynamic and static imaging using ^{18}F-PSMA-1007 PET/CT scan in all patients

Lesion Detection (n = 58)		Static Imaging		
		+	-	Total
Early Dynamic Imaging	+	48 (82.76%)	2 (3.45%)	**50 (86.21%)**
	-	1 (1.72%)	7 (12.07%)	**8 (13.79%)**
	Total	**49 (84.48%)**	**9 (15.52%)**	**58 (100.00%)**
Test Statistics		**0.330**		
P-value		**0.564**		

TABLE 4. Number of lesions in patients from the staging group detected on early dynamic and static imaging using ^{18}F-PSMA-1007 PET/CT scan

Lesion Detection (n = 25)		Static Imaging		
		+	-	Total
Early Dynamic Imaging	+	23 (92.00%)	2 (8.00%)	**25 (100.00%)**
	-	–	–	**–**
	Total	**23 (92.00%)**	**2 (8.00%)**	**25 (100.00%)**
Test Statistics		**2.000**		
P-value		**0.157**		

TABLE 5. Number of lesions in patients from the restaging group detected on early dynamic and static imaging using 18F-PSMA-1007 PET/CT scan

Lesion Detection (n = 33)		Static Imaging		
		+	-	Total
Early Dynamic Imaging	+	25 (75.76%)	–	**25 (75.76%)**
	-	1 (3.03%)	7 (21.21%)	**8 (24.24%)**
	Total	**26 (78.79%)**	**7 (21.21%)**	**33 (100.00%)**
Test Statistics		**1.000**		
P-value		**0.564**		

dynamic imaging were statistically significant. The three correlation coefficients imply that there are positive and moderate correlations between the said variables.

DISCUSSION

^{18}F-PSMA-1007 PET/CT is a new radiotracer used in diagnosing patients with prostate cancer [14, 15]. The recommended imaging time post-injection of the radiotracer is 40-90 minutes [15, 22, 23, 24, 25, 26]. Currently, there is no study on the use of early dynamic imaging using 18F-PSMA-1007 PET/CT in detecting PCa lesions. However, studies using 18F-Choline PET/CT in diagnosing PCa, showed that early dynamic acquisition in the pelvic region within 3-8 minutes post injection helped to distinguish avid lesion from the urinary bladder activity. The basis behind the use of early dynamic imaging is that cancer lesions show radiotracer accumulation before radiotracer buildup in the urinary

TABLE 6. Correlation between the Gleason Scores, PSA, and SUVmax values of all patients

Comparison	**Correlation Coefficient**	**Correlation Interpretation**	**Sig.**	**p-value Interpretation**
Gleason Scores vs. SUV Levels – **all ptients**	0.114	Weak	0.315	Insignificant
Gleason Scores vs. SUV Levels – **Early dynamic imaging in the Staging group**	0.157	Weak	0.341	Insignificant
Gleason Scores vs. SUV Levels - **Static imaging in the staging group**	0.550	Moderate	0.010	Significant
Gleason Scores vs. SUV Levels – **Early dynamic imaging in the restaging group**	–0.182	Weak	0.470	Insignificant
Gleason Scores vs. SUV Levels - **Static imaging in the restaging group**	–0.103	Weak	0.666	Insignificant
PSA level vs. SUV Levels – **All patients**	0.334	Weak	0.001	Significant
PSA level vs. SUV Levels – **Early dynamic imaging in the staging group**	0.236	Weak	0.317	Insignificant
PSA level vs. SUV Levels - **Static imaging in the staging group**	0.403	Moderate	0.079	Insignificant
PSA level vs. SUV Levels –**Early dynamic imaging in the restaging group**	0.606	Moderate	0.003	Significant
PSA level vs. SUV Levels - **Static imaging in the restaging group**	0.457	Moderate	0.017	Significant

[1] Correlation coefficients between | 0.00 | and | 0.10 | are interpreted as having negligible correlation, | 0.10 | and | 0.39 | as weak, | 0.40 | and | 0.69 | as moderate, | 0.70 | and | 0.89 | as strong, and | 0.90 | and | 1.00 | as very strong.
[2] Criteria: sig-value > 0.05 Not Significant (Accept Null Hypothesis)
sig-value < 0.05 Significant (Reject Null Hypothesis)

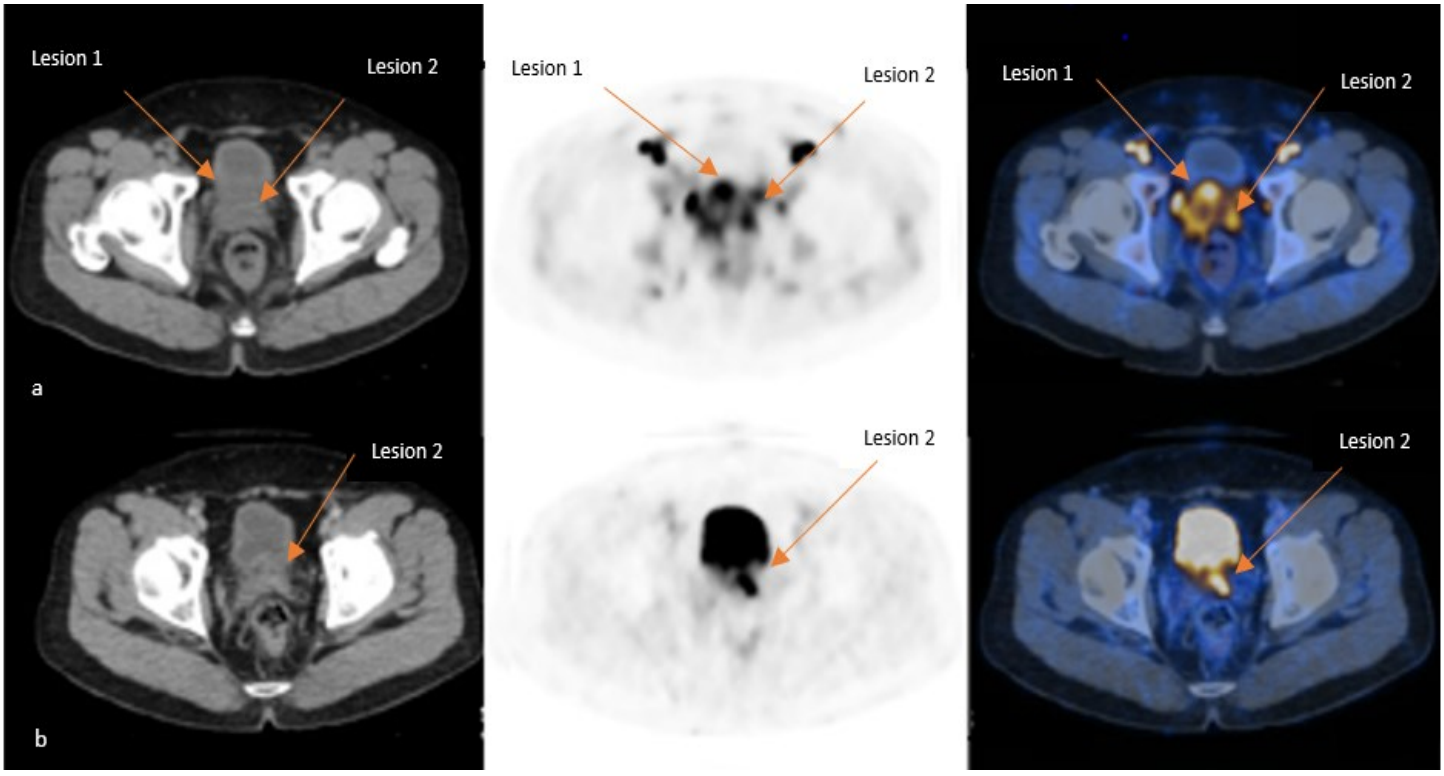


FIGURE 1. Axial cut of the ^{18}F-PSMA-1007 PET/CT scan of a 55-year-old, male, for prostate cancer staging. (a) Two foci of increased tracer activity were noted in the prostate gland in the early dynamic imaging (orange arrows). (b) In comparison to static imaging, one of the lesions (lesion 1) was obscured by the radiotracer accumulation in the uri-

bladder, which may obscure these lesions in the standard imaging at 60 minutes. In a similar study by Barakat et al., early imaging of Ga-68 PSMA PET/CT in 115 patients with PCa was examined. They acquired images at 3- and 60-minutes post-injection. In their study, 106 out of 115 lesions were detected on both early and static imaging; while 8 out of 115 lesions were only seen on early dynamic imaging. It showed a statistically significant increase in the detected rate from 64% using standard imaging to 68% when performing early imaging [28]. The author concluded that early images may help increase the detection rate of PSMA-avid lesions, in the anterior transition zone of the prostate or anterior aspect of the prostate bed [28]. In contrast to this study, there is no statistically significant difference between the detection rate of PSMA-avid lesions in both the early dynamic and static imaging in all the patients. Although there is no significant difference, the detection of 2/58 lesions in the early dynamic imaging, which was not seen in the static imaging may imply that early dynamic imaging may still aid in assessing lesions in PCa. Upon review of the images, these lesions were located in the basal segment of the prostate gland which is near the urinary bladder neck (Figure 1). This also demonstrates that prostate to bladder ratio is higher in the early dynamic images than in the static-60-minute imaging, due to the lesser radiotracer activity in the bladder during the initial phase.

At present, the recommended methods to reduce radioactivity in the bladder include voiding before the scan or the administration of diuretics like furosemide to wash out accumulated radioactivity within the bladder. However, these approaches are time-consuming, do not sufficiently reduce bladder activity, and the use of diuretics is often avoided in patients with renal impairment [17]. The study of Perveen et al. demonstrated that even after administration of diuretics before imaging at 60 minutes post-injection, the bladder activity remained increased compared to the early imaging [29]. PSMA is a transmembrane glycoprotein overexpressed in prostate cancer cells and shows low expression in benign prostatic tissue, setting forth the rationale for using ^{18}F-PSMA-1007 PET/CT in detecting prostate cancer lesions [13, 14, 15, 16]. GS reflects the grade of differentiation of PCa and thus correlates with tumor aggressiveness [4]. While PSA is a serine protease enzyme produced by the columnar epithelium of the prostatic tissue which is used in screening and monitoring patients with PCa [5, 6, 7]. The most common parameter used to measure tracer accumulation in PET is the standardized uptake value (SUV). It is a semi-quantitative measure of normalized radioactivity concentration in PET images. A SUVmax of 2.5 or higher than the background is generally considered to be indicative of malignant tissue [29, 30, 31, 32,33, 34]. Several studies using Ga68-PSMA and ^{18}F-PSMA-1007 PET/CT scans showed a statistically significant correlation between serum PSA and SUVmax in primary tumors [35, 36, 37, 38]. Therefore, as serum PSA levels increase, the SUVmax value also increases and thus results in a higher probability of detecting lesions in cases of prostate cancer. Reflected in this study, is the positive correlation between serum PSA levels and SUVmax values in all of the patients and restaging group patients. Patients who were referred for restaging following treatment had lower median serum PSA and SUVmax values compared to the group for initial staging. No significant correlation was observed between GS and SUVmax values in both early dynamic or static imaging when all patients were analyzed. A significant correlation was only seen among the initial staging group when stratified. In contrast to the current study, one conducted by Hong et al. showed a strong correlation between the SUVmax and GS values in PCa patients using ^{18}F-PSMA-1007 PET/CT scan, wherein the value of SUVmax was higher in GS > 7 [35]. In this study, there were only 17 patients with a GS > 7, which may relate to the lack of correlation due to the small number of patients with GS of 8 and 9.

Of the 58 lesions, only 1 lesion was detected in the static imaging alone which was not seen in the early dynamic imaging. This patient was under the restaging group and had already received hormonal treatment, with a serum PSA of 0.07 ng/ml. This lesion was not well delineated in the early dynamic imaging but was detectable in the static imaging with an SUVmax of 4.6. In this case, the low serum PSA and SUVmax measurements may reflect that the patient's therapy is effective. The detected lesion may not be seen in the early dynamic imaging probably because the lesion is not as metabolically active when compared to an untreated lesion, hence standard static imaging of 60 minutes may improve and enable lesion detection when a sufficient amount of radiotracer has accumulated and the SUV value may be determined.

Out of the 53 patients, 7 had no detectable lesion on both early dynamic and static imaging. These patients were referred for restaging, after receiving treatment. The majority of these patients had undergone prostatectomy and hormonal therapy. These negative findings may relate to favorable or effective treatment responses.

Limitations of this study:

The major limitation of this study is the small sample size and its retrospective nature. A prospective study in multiple institutions with a larger sample size is recommended to strengthen the findings of this research paper. Studies in assessing the sensitivity and specificity of ^{18}F-PSMA-1007 in the diagnosis of PCa, as well as in evaluating treatment response are likewise recommended. In addition, the correlation of GS, serum PSA, and SUVmax needs further evaluation.

CONCLUSION

Early dynamic imaging may serve as an adjunctive procedure for detecting PSMA-avid lesions; however, it is not recommended as a standard component of the overall imaging protocol with prostate cancer using ^{18}F-PSMA-1007 PET/CT. It may be particularly useful for identifying lesions near the base of the prostate gland, in proximity to the urinary bladder.

Disclosure. The author declares no conflict of interest relevant to the conduct and authorship of this study.

Acknowledgment. The author expresses her utmost gratitude to Ms. Ma. Kristine Joy S. Calvario and Dr. Candice Genuino-Montaño for their guidance in the construction of the protocol up to the finalization of the manuscript.

REFERENCES

1. Fitzmaurice, C. et al. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: a systematic analysis for the global burden of disease study. JAMA oncology 3, 524–548 (2017).
2. Redaniel MTM, Laudico AV, Lumague MRM, et al. Cancer in the Philippines vol. IV. Part I - Cancer incidence 1998-2002. Philippine Cancer Society, Manila 2008. 2
3. De Visschere P, Oosterlinck W, De Meerleer G, Villeirs G. Clinical and imaging tools in the early diagnosis of prostate cancer, a review. JBR-BTR. 2010 Mar-Apr;93 (2):62-70. doi: 10.5334/jbr-btr.121. PMID: 20524513.
4. Candefjord S., Ramser K., Lindahl O.A.: Technologies for localization and diagnosis of prostate cancer. J Med Eng Technol, 2009, 33: 585-603.
5. Han M, Partin AW, Zahurak M, et al. Biochemical (prostate -specific antigen) recurrence probability following radical prostatectomy for clinically localized prostate cancer. J Urol. 2003;169(2): 517-523.
6. Punnen S, Cooperberg MR, D'Amico AV, et al. Management of biochemical recurrence after primary treatment of prostate cancer: a systematic review of the literature. Eur Urol. 2013;64(6): 905-915.
7. Rosenbaum E, Partin A, Eisenberger MA. Biochemical relapse after primary treatment for prostate cancer: studies on natural history and therapeutic considerations. J Natl ComprCancNetw. 2004;2(3):249-256.
8. Simmons MN, Stephenson AJ, Klein EA. Natural history of biochemical recurrence after radical prostatectomy: risk assessment for secondary therapy. Eur Urol. 2007;51 (5):1175-1184.
9. Bott SR. Management of recurrent disease after radical prostatectomy. Prostate Cancer Prostatic Dis. 2004;7:211–6.
10. Beer AJ, Eiber M, Souvatzoglou M, Schwaiger M, Krause BJ. Radionuclide and hybrid imaging of recurrent prostate cancer. Lancet Oncol. 2011;12:181–91.
11. Johnstone PA, Tarman GJ, Riffenburgh R, Rohde DC, Puckett ML, Kane CJ. Yield of imaging and scintigraphy assessing biochemical failure in prostate cancer patients. Urol Oncol. 1997;3:108–12.
12. Cher ML, Bianco Jr FJ, Lam JS, et al. Limited role of radionuclide bone scintigraphy in patients with prostate specific antigen elevations after radical prostatectomy. J Urol. 1998;160:1387–91.
13. Maurer, T., Eiber, M., Schwaiger, M. &Gschwend, J. E. Current use of PSMA-PET in prostate cancer management. Nature reviews. Urology 13, 226 (2016).
14. Kesch C, Kratochwil C, Mier W, Kopka K, Giesel FL. 68Ga or 18F for prostate cancer imaging? J Nucl Med. 2017;58:687 -688.
15. Giesel FL, Hadaschik B, Cardinale J, et al. F-18 labelled PSMA-1007: biodistribution, radiation dosimetry and histopathological validation of tumor lesions in prostate cancer patients. Eur J Nucl Med Mol Imaging. 2017;44:678 -688.
16. Kasperzyk JL, Finn SP, Flavin R, Fiorentino M, Lis R, Hendrickson WK, et al. Prostate-specific membrane antigen protein expression in tumor tissue and risk of lethal prostate cancer. Cancer Epidemiol Biomark Prev. 2013;22:2354–63.
17. Afshar-Oromieh A, Avtzi E, Giesel FL, Holland-Letz T, Linhart HG, Eder M, et al. The diagnostic value of PET/CT imaging with the (68)Ga-labelled PSMA ligand HBED-CC in the diagnosis of recurrent prostate cancer. Eur J Nucl Med Mol Imaging. 2015;42:197–209.
18. Rahman LA, Rutagengwa D, Lin P, Lin M, Yap J, Lai K, et al. High negative predictive value of 68Ga PSMA PET-CT for local lymph node metastases in high risk primary prostate cancer with histopathological correlation. Cancer Imaging. 2019;19:86.
19. Rahbar K, Weckesser M, Huss S, Semjonow A, Breyholz HJ, Schrader AJ, et al. Correlation of intraprostatic tumor extent with 68Ga-PSMA distribution in patients with prostate cancer. J Nucl Med. 2016;57:563–7.

20. Afshar-Oromieh A, Zechmann CM, Malcher A, et al. Comparison of PET imaging with a (68)Ga-labelled PSMA ligand and (18)F-choline-based PET/CT for the diagnosis of recurrent prostate cancer. Eur J Nucl Med Mol Imaging. 2014;41:11-20
21. Afshar-Oromieh A, Malcher A, Eder M, Eisenhut M, Linhart HG, Hadaschik BA, et al. PET imaging with a [68Ga]gallium -labelled PSMA ligand for the diagnosis of prostate cancer: biodistribution in humans and first evaluation of tumour lesions. Eur J Nucl Med Mol Imaging. 2013;40:486–95.
22. Giesel FL, Knorr K, Spohn F, Will L, Maurer T, Flechsig P, et al. Detection efficacy of (18)F-PSMA-1007 PET/CT in 251 patients with biochemical recurrence of prostate cancer after radical prostatectomy. J Nucl Med. 2019;60:362–8.
23. Kuten, J., Fahoum, I., Savin, Z., Shamni, O., Gitstein, G., Hershkovitz, D., … Even-Sapir, E. (2019). Head- to head Comparison of 68Ga-PSMA-11 with 18F-PSMA-1007 PET/ CT in Staging Prostate Cancer Using Histopathology and Immunohistochemical Analysis as Reference-Standard. Journal of Nuclear Medicine jnumed. 119.234187. doi:10.2967/jnumed.119.234187
24. Giesel, F. L., Will, L., Lawal, I., Lengana, T., Kratochwil, C., Vorster, M., … Sathekge, M. (2017). Intraindividual Comparison of18F-PSMA-1007 and18F-DCFPyL PET/CT in the Prospective Evaluation of Patients with Newly Diagnosed Prostate Carcinoma: A Pilot Study. Journal of Nuclear Medicine, 59(7), 1076–1080.
25. Cardinale J, Schafer M, Benesova M, et al. Preclinical evaluation of 18F-PSMA-1007, a new prostate-specific membrane antigen ligand for prostate cancer imaging. J Nucl Med. 2017;58:425-431.
26. Cardinale J, Martin R, Remde Y, et.al. Procedures for the GMP-Compliant Production and Quality Control of [18F] PSMA-1007: A Next Generation Radiofluorinated Tracer for the Detection of Prostate Cancer. Pharmaceuticals (Basel). 2017 Sep 27;10(4):77. doi: 10.3390/ph10040077. PMID: 28953234; PMCID: PMC5748634.
27. Dimitrakopoulou-Strauss A, Pan L, Strauss LG. Quantitative approaches of dynamic FDG-PET and PET/CT studies (dPET/CT) for the evaluation of oncological patients. Cancer Imaging. 2012;12(1):283-289. Published 2012 Sep 28. doi:10.1102/1470-7330.2012.0033
28. Barakat, A., Yacoub, B., Homsi, M. E., Saad Aldine, A., El Hajj, A., & Haidar, M. B. (2020). Role of Early PET/CT Imaging with 68Ga-PSMA in Staging and Restaging of Prostate Cancer. Scientific Reports, 10(1). doi:10.1038/ s41598-020-59296-6
29. Perveen, G., Arora, G., Damle, N. A., Prabhu, M., Arora, S., Tripathi, M., Bal, C., Kumar, P., Kumar, R., & Singh, P. (2018). Role of Early Dynamic Positron Emission Tomography/Computed Tomography with 68Ga-prostate-specific Membrane Antigen-HBED-CC in Patients with Adenocarcinoma Prostate: Initial Results. Indian journal of nuclear medicine : IJNM : the official journal of the Society of Nuclear Medicine, India, 33(2),112–117.
30. Ziessman, H. A., O'Malley, J. P., & Thrall, J. H. (2006). Nuclear medicine: The requisites in radiology. Philadelphia: Mosby Elsevier
31. Jadvar Hossein et. Al,(2017): Appropriate Use Criteria for 18-F FDG PET/CT in restaging and Treatment Response Assessment of Malignant Disease. J Nuc Med: 201758 (12) 2026-2037
32. Whitehead A, Julious SA, Cooper CL, Campbell MJ. Estimating the sample size for a pilot randomised trial to minimise the overall trial sample size for the external pilot and main trial for a continuous outcome variable. Statistical Methods in Medical Research. 2016; 25(3):1057 -1073
33. Giesel FL, Sterzing F, Schlemmer HP, Holland-Letz T, Mier W, Rius M, Afshar-Oromieh A, Kopka K, Debus J, Haberkorn U, Kratochwil C. Intra-individual comparison of (68)Ga-PSMA-11-PET/CT and multi-parametric MR for imaging of primary prostate cancer. Eur J Nucl Med Mol Imaging. 2016 Jul;43(8):1400-6. doi: 10.1007/s00259-016-3346-0. Epub 2016 Mar 14. PMID: 26971788; PMCID: PMC4906063.
34. Hoffmann MA, Müller-Hübenthal J, Rosar F, Fischer N, von Eyben FE, Buchholz HG, Wieler HJ, Schreckenberger M. Primary Staging of Prostate Cancer Patients with [18F] PSMA-1007 PET/CT Compared with [68Ga]Ga-PSMA-11 PET/CT. J Clin Med. 2022 Aug 29;11(17):5064. doi: 10.3390/jcm11175064. PMID: 36078994; PMCID: PMC9457380.
35. Hong, J. J., Liu, B. L., Wang, Z. Q., Tang, K., Ji, X. W., Yin, W. W., Lin, J., & Zheng, X. W. (2020). The value of 18F-PSMA-1007 PET/CT in identifying non-metastatic high-risk prostate cancer. EJNMMI research, 10(1), 138.
36. PET/CT Imaging in Primary Prostate Cancer. Clinical Nuclear Medicine 41, e473–e479 (2016).
37. Fendler, W. P. et al. 68Ga-PSMA PET/CT detects the location and extent of primary prostate cancer. Journal of Nuclear Medicine 57,1720–1725 (2016).
38. Uprimny, C. et al. 68 Ga-PSMA-11 PET/CT in primary staging of prostate cancer: PSA and Gleason score predict the intensity of tracer accumulation in the primary tumour. European journal of nuclear medicine and molecular imaging 44, 941–949 (2017).

Thyroid Disorder Classification using Machine Learning Models

Vincent Peter C. Magboo, MD, MS, Ma. Sheila A. Magboo, MS

Department of Physical Sciences and Mathematics, University of the Philippines Manila
E-mail address: vcmagboo@up.edu.ph, mamagboo@u.edu.ph

ABSTRACT

Introduction:

Thyroid hormones are produced by the thyroid gland and are essential for regulating the basal metabolic rate. Abnormalities in the levels of these hormones lead to two classes of thyroid diseases – hyperthyroidism and hypothyroidism. Detection and monitoring of these two general classes of thyroid diseases require accurate measurement and interpretation of thyroid function tests. The clinical utility of machine learning models to predict a class of thyroid disorders has not been fully elucidated.

Objective:

The objective of this study is to develop machine learning models that classify the type of thyroid disorder on a publicly available thyroid disease dataset extracted from a machine learning data repository.

Methods:

Several machine learning algorithms for classifying thyroid disorders were utilized after a series of pre-processing steps applied on the dataset.

Results:

The best performing model was obtained by with XGBoost with a 99% accuracy and showing very good recall, precision, and F1-scores for each of the three thyroid classes. Generally, all models with the exception of Naïve Bayes did well in predicting the negative class generating over 90% in all metrics. For predicting hypothyroidism, XGBoost, decision tree and random forest obtained the most superior performance with metric values ranging from 96-100%. On the other end in predicting hyperthyroidism, all models have lower classification performance as compared to the negative and hypothyroid classes Needless to say, XGBoost and random forest did obtain good metric values ranging from 71-89% in predicting hyperthyroid class.

Conclusion:

The findings of this study were encouraging and had generated useful insights in the application and development of faster automated models with high reliability which can be of use to clinicians in the assessment of thyroid diseases. The early and prompt clinical assessment coupled with the integration of these machine learning models in practice can be used to determine prompt and precise diagnosis and to formulate personalized treatment options to ensure the best quality of care to our patients.

***Keywords**: thyroid disorders, machine learning, feature importance, SMOTE, XGBoost*

INTRODUCTION

The thyroid is a butterfly-shaped organ producing the thyroid hormones the levels of which play an important function of regulating the basal metabolic rate. Sufficient levels of these metabolic thyroid hormones are crucial for protein synthesis, for fetal and childhood tissue development and growth, for normal development of the nervous system in utero, in early childhood and continuing further to support neurological function in adults [1]. Derangement in the levels of these hormones lead to two classes of thyroid diseases namely: hyperthyroidism and hypothyroidism characterized by hyperfunction and hypofunction of the thyroid gland, respectively. In hypothyroidism, many patients complain of fatigue, weight gain and intolerance to cold temperature while anxiety, weight loss and sensitivity to heat are common symptoms of hyperthyroidism [2]. Needless to say, detection and monitoring of these two general classes of thyroid diseases require accurate measurement and interpretation of thyroid function tests [3].

The technological advancements in data mining techniques including processing and computation, machine learning (ML) approaches can also be applied to classify thyroid diseases [4]. Mollica et al., applied machine learning approach coupled with oversampling techniques and Bayesian networks framework on classification of thyroid tumors on histopathological images [5]. Authors concluded that integrating ML models in clinical practice could help reduce a pathologist's workload on top of improving disease diagnosis. In the study by Alyas et al., researchers applied several ML algorithms like decision tree, random forest algorithm, k-Nearest Neighbors (KNN), and artificial neural networks (ANN) on a thyroid disease dataset [6]. Results showed random forest with the highest classification accuracy at 94.8%. In [7], authors applied ML machine learning-based techniques to predict hypothyroidism namely: decision tree, random forest, naive Bayes, and ANN. Results showed decision tree and random forest generated the highest classification performance with an accuracy of 99.6% and 99.3%, respectively.

The objective of this study is to develop machine learning models that classify the type of thyroid disorder on a publicly available thyroid disease dataset extracted from a machine learning data repository. The clinical utility of ML models to predict a class of thyroid disorders has not been fully elucidated. The main contribution of this research is to find robust and reliable prediction models that will assist healthcare professionals in assessing the type of thyroid disease.

METHODOLOGY

The study was performed in several steps. The first step was loading of the dataset. Pre-processing techniques applied to the dataset include data cleaning, imputation method, and utilizing Synthetic Minority Oversampling TEchnique (SMOTE) for handling data imbalance. The next step was the application of several ML algorithms followed by assessment of classification performance. The machine learning pipeline for this study is shown in Figure 1.

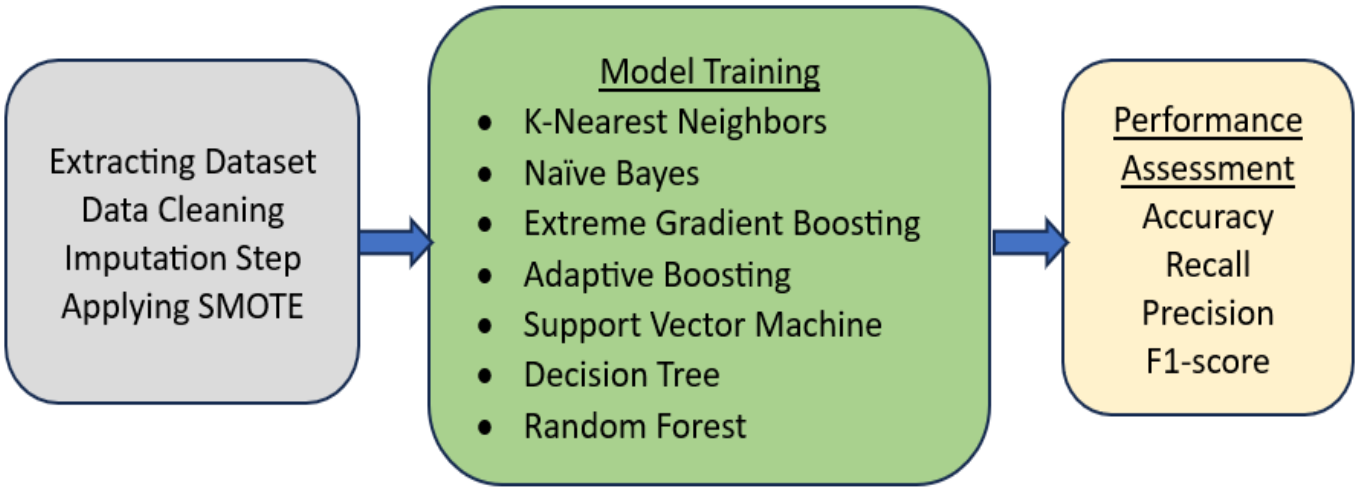


FIGURE 1. Machine Learning Pipeline for Thyroid Disorder Classification

TABLE 1. Independent attributes for the Thyroid Disorder Classification

Attribute	Data Type	Attribute	Data Type	Attribute	Data Type
patient_id	object	sick	object	FTI	float
age	integer	lithium	object	TBG	float
on_thyroxine	object	goitre	object	TSHmeasured	object
query_on_thyroxine	object	tumor	object	T3measured	object
on_antithyroid_meds	object	hypopituitary	object	TT4measured	object
query_hyperthyroid	object	psych	object	T4Umeasured	object
pregnant	object	TSH	float	FTImeasured	object
thyroid_surgery	object	T3	float	TBGmeasured	object
I131_treatment	object	TT4	float	referralsource	object
query_hypothyroid	object	T4U	float	binaryclass	object

Dataset Description

In this study, a publicly available thyroid disease extracted from a publicly available machine learning repository (University of California Irvine Machine Learning Repository) was used [8]. This dataset contains 9,172 anonymized thyroid disease cases from Garavan Institute, Sydney, Australia. The dataset consisted of 31 columns including the target variable, diagnosis. The listing of the independent attributes is seen in Table 1.

Pre-processing Steps

To prepare the dataset for machine learning, data cleaning and pre-processing methods were applied. Redundant and irrelevant variables such as 'TSHmeasured', 'T3measured', 'TT4measured', 'T4Umeasured', 'FTImeasured', 'TBGmeasured', 'patient_id', 'referralsource' were dropped from the dataset as they were mainly boolean variables with no predictive capability. Rows with inconsistent values, particularly those age over 100 years old, were likewise gnoses (negative, hypothyroid, and hyperthyroid) were retained for ML application as other diagnoses were deemed not relevant to the main focus of this research study. Hence, as a result of these data cleaning, the dataset was reduced to 7,142 records. The dataset has a severe imbalance with negative class comprising 89.4% (6,384 records) while the hypothyroid and hyperthyroid classes comprised 8.1% (582 records) and 2.5% (175 records) respectively. To address this severe imbalance, SMOTE was utilized.

Machine Learning Models

The dataset was split into 25% testing and 75% training with 10-fold cross validation. Python 3.8 and its ML libraries (scikit-learn, pandas, Matplotlib, seaborn, and NumPy) were used. Several ML models were utilized to predict thyroid diseases namely: k-Nearest Neighbors (kNN), Naïve Bayes (NB), Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), Adaptive Boosting (AdaBoost), and Extreme Gradient Boosting (XGBoost).

Performance Metrics

Metrics such as accuracy, recall, precision, and F1 score were computed to assess classification performance. Accuracy refers to the ability of the ML model to predict the classes of the dataset correctly and assess how close or near the predicted value is to the actual or theoretical value [9]. Recall is the ratio of the correctly classified number of positive instances to the number of all instances whose actual class is positive [10]. Recall is also called the true positive rate or sensitivity rate. The precision, sometimes called the positive predictive value, denotes the proportion of the retrieved samples which are relevant and is calculated as the ratio between correctly classified samples and all samples assigned to that class [11]. F1—score is defined as the harmonic

TABLE 2. Performance Metrics of the ML Models for Thyroid Disorder Classification

ML Model	Accuracy	Target Class	Recall	Precision	F1-score
XGBoost	99	Negative	99	100	99
		Hypothyroid	100	97	99
		Hyperthyroid	89	74	80
Decision Tree	98	Negative	99	99	99
		Hypothyroid	99	98	98
		Hyperthyroid	80	73	76
Random Forest	98	Negative	99	100	99
		Hypothyroid	100	96	98
		Hyperthyroid	89	71	79
AdaBoost	98	Negative	99	99	99
		Hypothyroid	100	73	96
		Hyperthyroid	70	99	75
Support Vector Machine	94	Negative	94	99	96
		Hypothyroid	97	84	90
		Hyperthyroid	93	35	51
k-Nearest Neighbors	92	Negative	93	98	96
		Hypothyroid	86	77	81
		Hyperthyroid	80	34	47
Naïve Bayes	36	Negative	31	93	47
		Hypothyroid	67	46	54
		Hyperthyroid	84	4	7

mean of precision and recall and as such, to generate a high F1-score, necessarily require to have high values of recall and precision [12]. Additionally, the feature importance scores of the best performing model was also generated.

RESULTS AND DISCUSSION

The performance metrics of the various ML models are shown in Table 2. The best performing model is XGBoost with a 99% accuracy. XGBoost also generated the highest recall, precision, and F1-score for each of the three thyroid classes. Following XGBoost is random forest with an accuracy of 98% and showing very good recall, precision, and F1-scores for each of the three thyroid classes with XGBoost. AdaBoost and decision tree also obtained excellent accuracy rates of 98%. On the other hand, Naïve Bayes performed the worst with a measly accuracy rate of 36%. Additionally, its predictive capability for all the three thyroid classes were below par indicating its inability to predict thyroid disorders.

Generally, all ML models with the exception of Naïve Bayes did well in predicting the negative class generating over 90% in all metrics. This is expected as the negative class had the greatest number of instances in the dataset. For predicting hypothyroidism, XGBoost, decision tree and random forest obtained the most superior performance with metric values ranging from 96-100%. On the other end in predicting hyperthyroidism, all models have lower classification performance as compared to the negative and hypothyroid classes. Note that the hyperthyroid class only constituted 2.5% of the entire dataset. Nonetheless, XGBoost and random forest did obtain good metric values ranging from 71-89% in predicting hyperthyroid class. Likewise, AdaBoost and decision tree yielded fairly acceptable metric values in predicting hyperthyroid class. However, support vector machine, k-Nearest Neighbors and Naïve Bayes generated poor predictive capability in classifying hyperthyroidism more prominently with its very low precision and F1-scores. Nonetheless, our results highlight the importance of addressing the severe data imbalance to obtain a more reliable diagnostic performance.

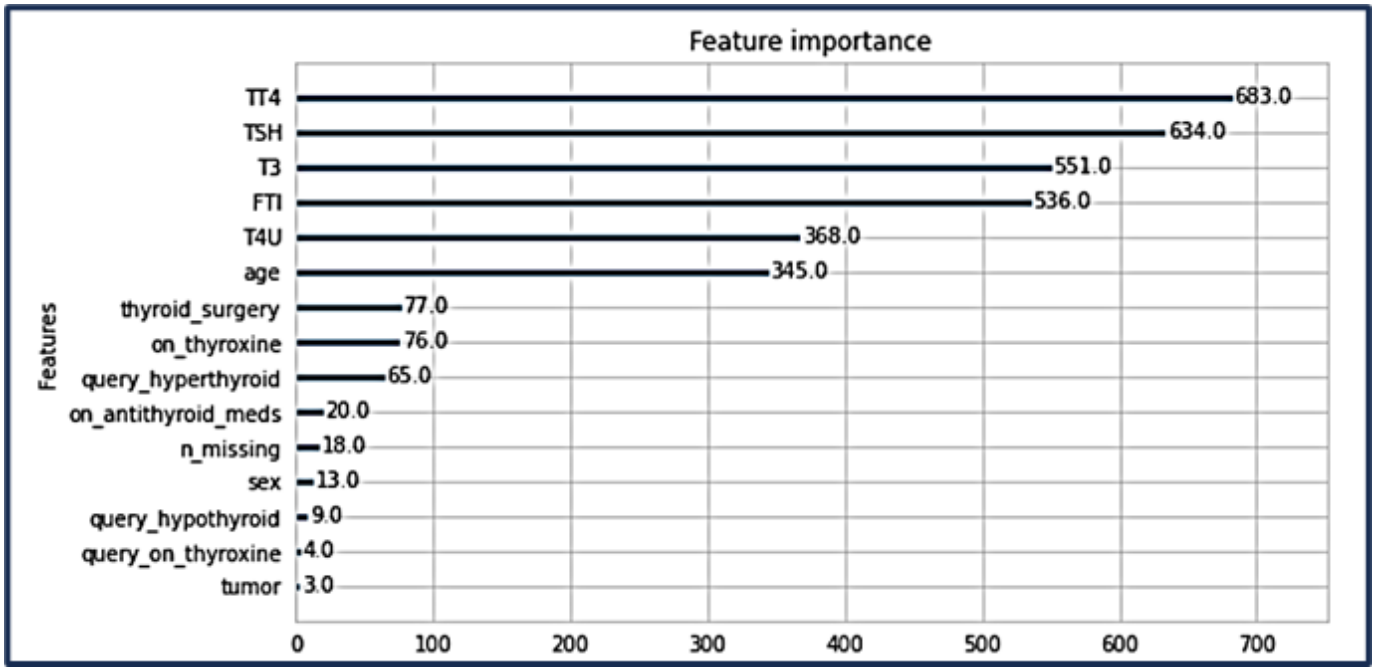


FIGURE 2. Feature Importance of Attributes of the Best Performing Model

As a measure to address the severe imbalance in this dataset, SMOTE was utilized. SMOTE can sufficiently increase the instances of minority samples so that the classification algorithm can increase the learning of minority samples during the training of the data [13]. Machine learning algorithms can be biased to favor the majority class in the presence of an imbalance [14]. SMOTE as an oversampling technique is a common measure utilized in ML to handle imbalanced datasets by creating copies of the minority class instances to balance the dataset which effectively led to a reduction in the bias and in the improvement of the classification accuracy of the model [13, 14, 15, 16].

The feature importance scores of attributes for the best performing model (XGBoost) are seen in Figure 2. The top important features were the hormone test level measurements (TT4, TSH, T3, FTI, T4U) while surprisingly though that the other attributes did not perform well in predicting our target variable. This confirmed our clinical suspicion that the hormone tests are the most helpful in our aim to predict target diagnosis as seen in the correlation heatmap in Figure 3. Only the attributes TT4, TSH, T3 and FTI had a strong positive correlation with the target variable. Feature importance highlights which attributes utilized by the ML model have higher predictive capability as compared to the other attributes. The identification of these features can aid in the model explainability [17]. Feature importance scores also provide insight into the data and the model by identifying most and least relevant in predicting the target variable. It also serves as a basis for dimensionality reduction by removing those attributes with lowest feature importance scores. This act simplifies the model which consequently lead to faster machine execution and also improved diagnostic performance of the model.

As to the metrics and best performing models, our results are comparable with other studies [4, 5, 6, 7, 18] which utilized traditional ML models applied to thyroid disease dataset. These findings suggest the feasibility of applying the machine learning approaches to predict thyroid disorders with acceptable results. The clinical utility of this study is even more highlighted with robust models that can provide faster and with high reliability to assist healthcare professional in predicting thyroid disorders as well as enable clinicians to propose personalized treatment options for our patients [19].

CONCLUSION

Alterations in the levels of thyroid hormones generally lead to two classes of thyroid diseases namely: hyperthyroidism and hypothyroidism characterized by hyperfunction and hypofunction of the thyroid gland, respectively. In this study, several machine learning

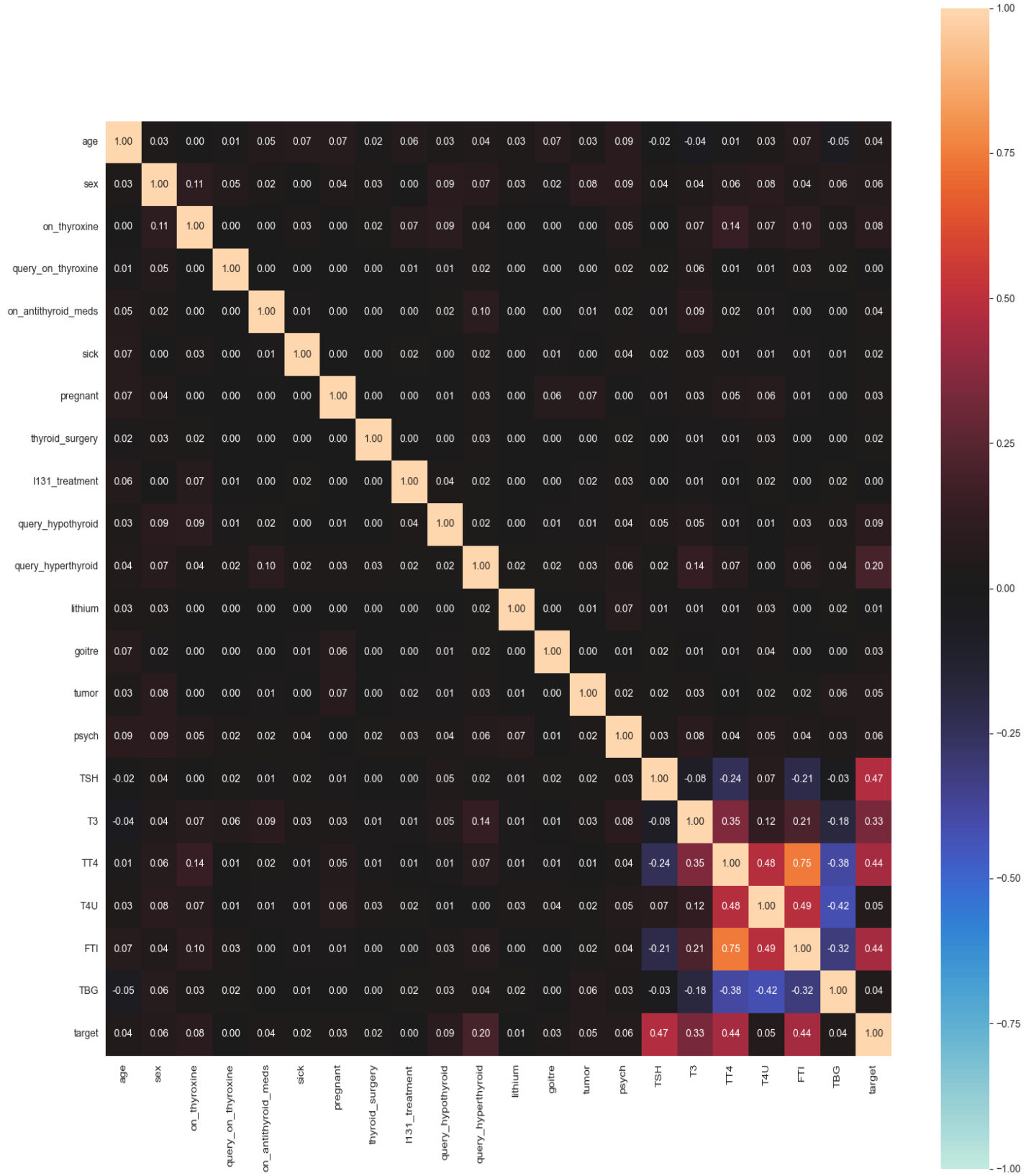


FIGURE 3. Correlation Heatmap of Predictor Variables for Thyroid Disorder Classification

models for classifying thyroid disorders were applied on a publicly available thyroid disease dataset from a machine learning data repository. The best performing model was obtained by with XGBoost with a 99% accuracy and showing very good recall, precision, and F1-scores for each of the three thyroid classes. Generally, all ML models with the exception of Naïve Bayes did well in predicting the negative class generating over 90% in all metrics. For predicting hypothyroidism, XGBoost, decision tree and random forest obtained the most superior performance with metric values ranging from 96-100%. On the other end in predicting hyperthyroidism, all models have lower classification performance as compared to the negative and hypothyroid classes Needless to say XGBoost and random forest did obtain good metric values ranging from 71-89% in predicting hyperthyroid class. Likewise, AdaBoost and decision tree yielded fairly acceptable metric values in predicting hyperthyroid class.

Future enhancement should include explainable artificial intelligence tools for better understanding of the models by the clinicians. Additionally, ML models could also be applied to larger datasets which combines patient symptoms, comorbidities, and radiographic features coming in the quest for excellent diagnostic accuracy. The findings of this study were encouraging and had generated useful insights in the application and development of faster automated models with high reliability which can be of use to clinicians in the assessment of thyroid diseases. The early and prompt clinical assessment coupled with the integration of these ML models in practice can be used to determine prompt and precise diagnosis and to formulate personalized treatment options to ensure the best quality of care to our patients.

REFERENCES

1. Gordon Betts J, Young K.A., Wise J, et al. (2022). The Thyroid Gland. In Anatomy and Physiology, 2nd Ed. OpenStax. https://openstax.org/books/anatomy-and-physiology-2e/pages/17-4-the-thyroid-gland.
2. Kwang-Sig Lee, Hyuntae Park. (2022). Machine learning on thyroid disease: a review. Front. Biosci. (Landmark Ed) 27 (3), 101. https://doi.org/10.31083/j.fbl2703101.
3. Andersen, S., Karmisholt, J., Bruun, N.H. et al. (2022). Interpretation of TSH and T4 for diagnosing minor alterations in thyroid function: a comparative analysis of two separate longitudinal cohorts. Thyroid Res 15, 19. https://doi.org/10.1186/s13044-022-00137-1.
4. Chaganti R, Rustam F, De La Torre Díez I, et al. (2022). Thyroid Disease Prediction Using Selective Features and Machine Learning Techniques. Cancers (Basel). 2022 Aug 13;14(16):3914. doi: 10.3390/cancers14163914. PMID: 36010907; PMCID: PMC9405591.
5. Mollica G, Francesconi D, Costante G, et al. (2022). Classification of Thyroid Diseases Using Machine Learning and Bayesian Graph Algorithms. IFAC-PapersOnLine, 55 (40) : 67 – 72. https://doi.org/10.1016/j.ifacol.2023.01.050.
6. Alyas T, Hamid M, Alissa K, et al. (2022). Empirical Method for Thyroid Disease Classification Using a Machine Learning Approach. BioMed Research International, vol. 2022, Article ID 9809932, 10 pages. https://doi.org/10.1155/2022/9809932.
7. Guleria K, Sharma S, Kumar K, Tiwari S. (2022). Early prediction of hypothyroidism and multiclass classification using predictive machine learning and deep learning. Measurement: Sensors, Volume 24, 2022, 100482. https://doi.org/10.1016/j.measen.2022.100482.
8. Quinlan, Ross. (1987). Thyroid Disease. UCI Machine Learning Repository. https://doi.org/10.24432/C5D010.
9. Debal DA, Sitote TM. (2022). Chronic kidney disease prediction using machine learning techniques. J Big Data 9, 109. https://doi.org/10.1186/s40537-022-00657-5.
10. Yağcı, M. (2022). Educational data mining: prediction of students' academic performance using machine learning algorithms. Smart Learn. Environ. 9, 11. https://doi.org/10.1186/s40561-022-00192-z.
11. Hicks SA, Strümke I, Thambawita V, et al. (2022). On evaluation metrics for medical applications of artificial intelligence. Sci Rep 12, 5979 (2022). https://doi.org/10.1038/s41598-022-09954-8.
12. De Diego IM, Redondo AR, Fernández RR, et al. (2022). General Performance Score for classification problems. Appl Intell 52, 12049–12063. https://doi.org/10.1007/s10489-021-03041-7.
13. Wei Du. (2022). Application of Improved SMOTE and XGBoost Algorithm in the Analysis of Psychological Stress Test for College Students. Journal of Electrical and Computer Engineering, vol. 2022, Article ID 2760986, 8 pages. https://doi.org/10.1155/2022/2760986.
14. Nishat MM, Faisal F, Ratul IJ, et al. (2022). A Comprehensive Investigation of the Performances of Different Machine Learning Classifiers with SMOTE-ENN Oversampling Technique and Hyperparameter Optimization for Imbalanced Heart Failure Dataset. Scientific Programming, vol. 2022, Article ID 3649406, 17 pages. https://doi.org/10.1155/2022/3649406.
15. Yakshit, Kaur G, Kaur V, Sharma Y, Bansal B. (2022). Analyzing various Machine Learning Algorithms with SMOTE and ADASYN for Image Classification having Imbalanced Data. 2022 IEEE International Conference on Current Development in Engineering and Technology (CCET), Bhopal, India, 2022, pp. 1-7, doi: 10.1109/CCET56606.2022.10080783.

16. Priyadarshinee S, Panda P. (2022). Improving Prediction of Chronic Heart Failure using SMOTE and Machine Learning. 2022 Second International Conference on Computer Science, Engineering and Applications (ICCSEA), Gunupur, India, 2022, pp. 1-6, doi: 10.1109/ ICCSEA54677.2022.9936470.
17. Collaris D, Weerts H, Miedema D, van Wijk J, Pechenizkiy M. (2022). Characterizing Data Scientists' Mental Models of Local Feature Importance. In Nordic Human-Computer Interaction Conference (NordiCHI '22). Association for Computing Machinery, New York, NY, USA, Article 9, 1–12. https://doi.org/10.1145/3546155.3546670.
18. Pal M, Parija S, Panda G. (2022). Enhanced Prediction of Thyroid Disease Using Machine Learning Method. 2022 IEEE VLSI Device Circuit and System (VLSI DCS), Kolkata, India, 2022, pp. 199-204, doi: 10.1109/ VLSIDCS53788.2022.9811472.
19. Magboo VC, Magboo MS. (2021). Machine Learning Classifiers on Breast Cancer Recurrences. In: Watrobski, J., Salabun, W., Toro, C., Zanni-Merk, C., Howlett, R., Jain, L. (eds.) 25th International Conference on Knowledge-Based and Intelligent Information & Engineering System 2021, Procedia Computer Science, vol 192, pp. 2742–2752. Elsevier, Warsaw, Poland. https://doi.org/10.1016/ j.procs.2021.09.044.

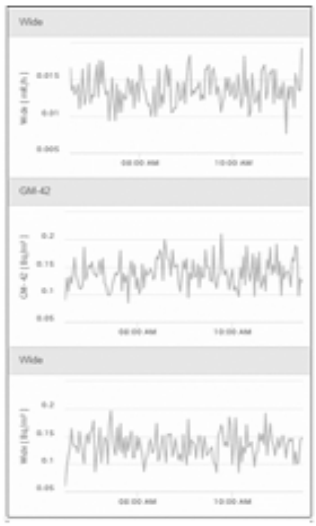
ROTEM
Rotem Industries Ltd

PERFECTION MADE IN DETECTION ISRAEL

If radiation detection and real-time monitoring is of concern, ROTEM offers quality global solutions. Our systems are already commissioned within nuclear medicine departments, cyclotrons, accelerators, and radioisotope laboratories worldwide. Please contact our local partner for more information.

Portable Survey Meters.

WebiSmarts - On Line Area & Stack Monitoring Systems.



WebiSmarts - Coincidence Stack System is capable of differentiating between various isotopes.

I-131 · Ga-68 · F-18 · Lu-177 · C-11 · N-13

ROTEM
Rotem Industries Ltd

ALVTECHNOLOGIES
PHILIPPINES INCORPORATED

ROTEM INDUSTRIES LTD.
Radiation Detection Division
Mishor Yamin, D.N. Arava 8680600, Israel
Tel. +972-8-6564780/1, Fax. +972-8-6573252
E-mail. sales@rotemi.co.il Web: www.rotem-radiation.co.il

YOUR LOCAL CONTACT:
ALVTechnologies Philippines Incorporated
Unit 2308 Medical Plaza Ortigas San Miguel Ave Pasig
Tel. +632-8-6382317 to 19, Fax. +632-8-6382318
E-mail. sales@alvtechnologies.com.ph

aspen
PHILIPPINES

Levothyroxine sodium

Eltroxin

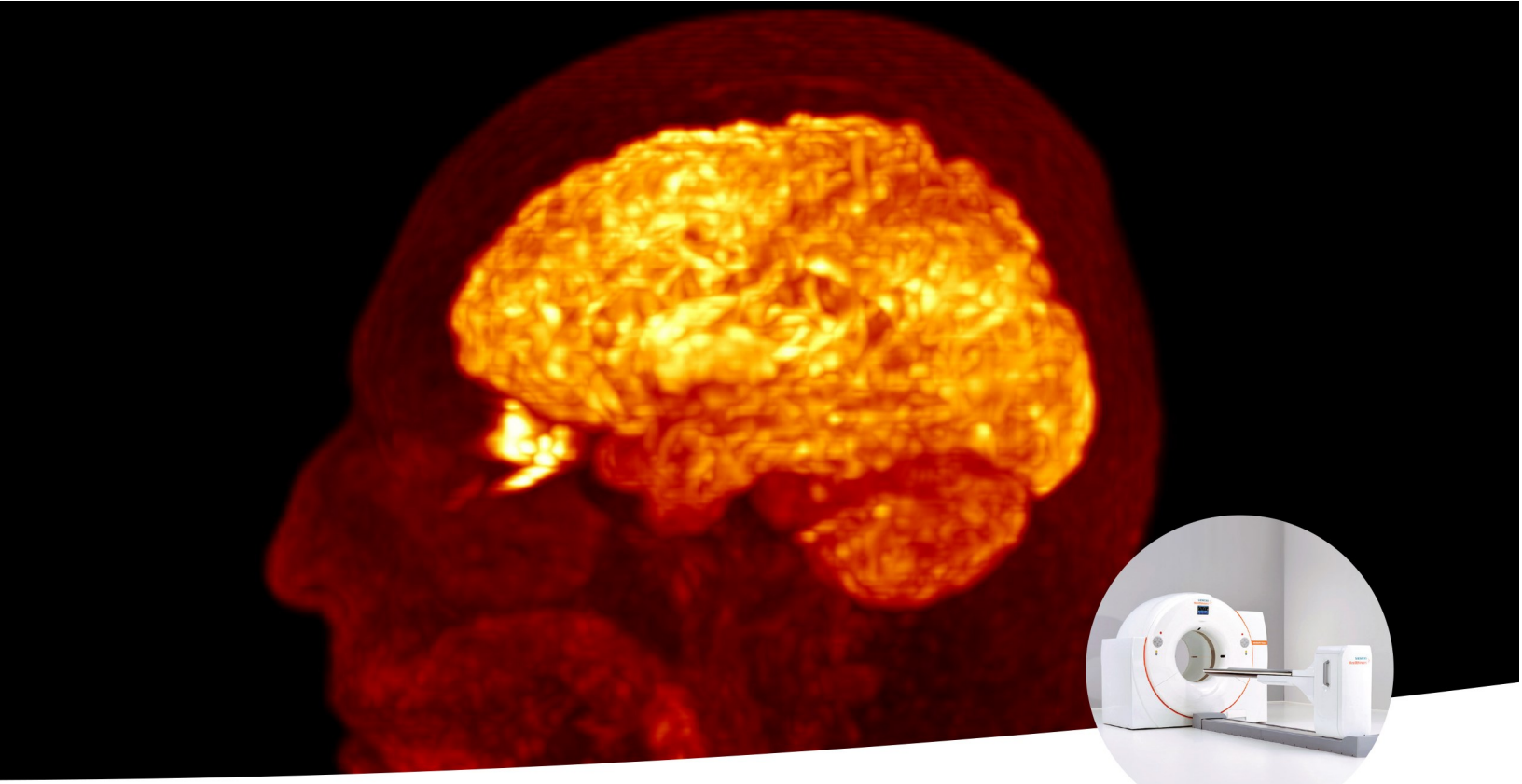
Thiamazole

Strumazol

aspen
PHILIPPINES
Healthcare. We Care.

Unit 1001-1002, 10th Floor
Trade and Financial Tower
7th Avenue cor. 32nd Street
Bonifacio Global City, Taguig City

APH-THM-12-21-00001

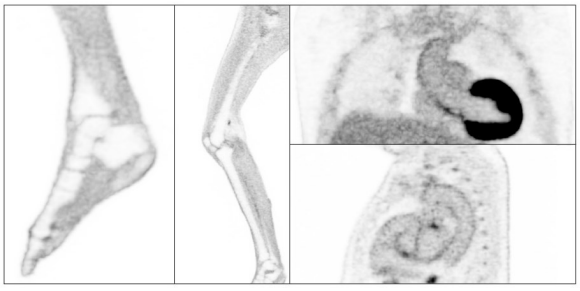


Biograph Vision
see a whole new world
of precision

Expanding precision medicine.

The healthcare market operates in a constant state of flux, which can be challenging as you strive for better clinical outcomes, faster workflow, and consistent quality of results.

What if you were able to visualize smaller lesions, get information for more accurate staging and patient risk stratification, and manage operational inefficiencies to facilitate the most appropriate treatment strategy[1]?



Impressive clarity and remarkable delineation are hallmarks of Biograph Vision scans. Data courtesy of University Medical Center Groningen, Groningen, The Netherlands

3.2 mm LSO crystals[1]
Fast time of flight at 214 ps[1]
High effective sensitivity at 100 cps/kBq[1]
100% sensor coverage[1]

Accuracy to reveal the bigger picture.
Performance to maximize efficiency.
Reproducibility to understand disease progression.

Learn more about Biograph Vision at siemens.com/vision

[1]Based on internal measurements (resolution and time of flight) compared to current systems. Data on file. Biograph Vision is not commercially available in all countries. Its future availability cannot be guaranteed. Please contact your local Siemens Healthineers organization for further details.

siemens.com/vision

SIEMENS Healthineers