Jurnal Global Ilmiah
Vol. 4, No. 2, November 2026, pp 161-180
e-ISSN: 3026-5207
DOI: 10.55324/jgi.v4i2.445

TIRADS 1-4: Antara Kecurigaan, Kepastian, dan Validasi Histopatologis Komprehensif dalam Penatalaksanaan Nodul Tiroid

Marcel Hariman*, Denny Saleh
Universitas Sam Ratulangi Manado, Indonesia
Email: harimanmarcel@gmail.com*, dennysaleh@yahoo.co.uk

Keywords:
diagnostic accuracy; histopathology; thyroid malignancy; thyroid nodules; TIRADS

This is an open access article under the CC BY-SA license.

Copyright © 2026 by Author. Published by Internasional Journal Labs

Abstract

Thyroid nodules are highly prevalent, yet only a small proportion are malignant, making accurate risk stratification essential to avoid unnecessary invasive procedures. The Thyroid Imaging Reporting and Data System (TIRADS) is widely used, but its agreement with histopathological findings requires validation in local populations. This study aimed to assess the association between TIRADS categories 1–4 and malignancy, evaluate diagnostic performance, and identify the optimal category threshold in a local sample. A retrospective cross-sectional diagnostic study was conducted using paired ultrasonography and pathology reports from 21 eligible patients at a tertiary referral hospital in Manado, Indonesia. Each patient was assigned the highest reported TIRADS category, with histopathology and one Bethesda II cytology result as reference standards. Analyses included exact testing, diagnostic performance estimates, and receiver operating characteristic (ROC) curves. Malignancy was found in 10 patients (47.6%), predominantly papillary thyroid carcinoma. TIRADS categories 3 and 4 were most frequent (28.6% each). The association between TIRADS categories and pathological findings was not statistically significant ($p = 0.282$). A threshold of TIRADS ≥4 yielded 60.0% sensitivity, 72.7% specificity, and 66.7% accuracy, with an area under the curve of 0.641 (95% CI: 0.395–0.887; $p = 0.275$) and an unnecessary biopsy rate of 33.3%. Four malignancies occurred below this threshold, indicating limited sensitivity. In conclusion, TIRADS discrimination was not statistically established in this sample. These preliminary findings should be interpreted cautiously due to the small sample size, incomplete data, and heterogeneous reporting; larger studies with standardized ultrasonography and nodule-to-specimen matching are needed.

Kata Kunci:
akurasi diagnostik; histopatologi; keganasan tiroid; nodul tiroid; TIRADS

Abstrak

Nodul tiroid sangat prevalen, namun hanya sebagian kecil yang bersifat ganas, sehingga stratifikasi risiko yang akurat diperlukan untuk menghindari tindakan invasif yang tidak perlu. Thyroid Imaging Reporting and Data System (TIRADS) banyak digunakan, tetapi kesesuaiannya dengan hasil histopatologi masih memerlukan validasi pada populasi lokal. Penelitian ini bertujuan menilai hubungan kategori TIRADS 1–4 dengan keganasan, mengukur kinerja diagnostik, serta mengidentifikasi ambang kategori optimal pada sampel lokal. Studi diagnostik retrospektif dengan pendekatan potong lintang ini menganalisis pasangan laporan ultrasonografi dan patologi anatomi dari 21 pasien yang memenuhi kriteria di rumah sakit rujukan tersier di Manado, Indonesia. Kategori TIRADS tertinggi digunakan sebagai kategori setiap pasien, dengan histopatologi dan satu hasil sitologi Bethesda II sebagai baku

emas. Analisis meliputi uji eksak, penghitungan parameter diagnostik, dan kurva receiver operating characteristic (ROC). Keganasan ditemukan pada 10 pasien (47,6%), terutama karsinoma papiler tiroid. Kategori TIRADS 3 dan 4 paling banyak ditemukan (masing-masing 28,6%). Hubungan kategori TIRADS dengan hasil patologi tidak bermakna secara statistik (p = 0,282). Ambang TIRADS ≥4 memberikan sensitivitas 60,0%, spesifisitas 72,7%, dan akurasi 66,7%, dengan luas area di bawah kurva 0,641 (IK 95%: 0,395–0,887; p = 0,275) dan proporsi biopsi tidak perlu 33,3%. Empat keganasan berada di bawah ambang tersebut, menunjukkan sensitivitas yang terbatas. Kesimpulannya, kemampuan diskriminasi TIRADS belum terbukti bermakna secara statistik pada sampel ini. Temuan pendahuluan ini perlu ditafsirkan dengan hati-hati karena sampel kecil, data tidak lengkap, dan heterogenitas pelaporan; penelitian lanjutan dengan standardisasi ultrasonografi dan pencocokan nodul dengan spesimen diperlukan.

PENDAHULUAN

Nodul tiroid merupakan salah satu kelainan endokrin yang paling sering dijumpai dalam praktik klinis sehari-hari, baik pada pelayanan primer, poliklinik bedah, maupun poliklinik penyakit dalam. Nodul tiroid didefinisikan sebagai lesi diskret di dalam kelenjar tiroid yang secara radiologis dapat dibedakan dari parenkim tiroid di sekitarnya. Data epidemiologis global menunjukkan bahwa prevalensi nodul tiroid sangat bervariasi tergantung pada metode deteksi yang digunakan. Sebuah tinjauan sistematis dan meta-analisis terhadap lebih dari sembilan juta subjek dari seluruh dunia melaporkan bahwa prevalensi nodul tiroid pada populasi umum mencapai 24,83%, dengan kecenderungan meningkat dari 21,53% pada periode 2000–2011 menjadi 29,29% pada periode 2012–2022, serta lebih banyak dijumpai pada perempuan (36,51%) dibandingkan laki-laki (23,47%) (Mu et al., 2022). Temuan serupa dilaporkan pada studi kohort pemeriksaan kesehatan berskala besar di Tiongkok yang melibatkan hampir tujuh juta peserta, dengan prevalensi nodul tiroid sebesar 36,9% (Y. Li et al., 2021). Peningkatan prevalensi ini sebagian besar didorong oleh semakin luasnya penggunaan ultrasonografi resolusi tinggi dalam pemeriksaan kesehatan rutin, sehingga banyak nodul yang sebelumnya tidak terdeteksi kini ditemukan secara insidental (Netter, 2027; Ovalle & Nahirney, 2021).

Meskipun sebagian besar nodul tiroid bersifat jinak, kekhawatiran utama dalam pengelolaan klinisnya adalah kemungkinan keganasan yang menyertainya. Konsensus SFE-AFCE-SFMN tahun 2022 tentang tata laksana nodul tiroid menegaskan bahwa lebih dari 90% nodul tiroid bersifat jinak dan tidak memerlukan tindakan spesifik, namun tujuan utama evaluasi klinis adalah mendeteksi nodul ganas yang berisiko mengalami rekurensi maupun kematian, nodul toksik yang menyebabkan hipertiroidisme, serta nodul kompresif yang memerlukan terapi (Borson-Chazot et al., 2022). Oleh karena jumlah nodul tiroid yang harus dievaluasi sangat besar sementara hanya sebagian kecil yang ganas, dibutuhkan suatu sistem stratifikasi risiko yang efisien, terstandarisasi, dan dapat direproduksi agar tindakan biopsi aspirasi jarum halus (*fine needle aspiration biops*y/FNAB) hanya dilakukan pada nodul yang benar-benar mencurigakan, sehingga mengurangi tindakan invasif yang tidak perlu.

Untuk menjawab kebutuhan tersebut, dikembangkanlah *Thyroid Imaging Reporting and Data System* (TIRADS), suatu sistem klasifikasi berbasis ultrasonografi yang

mengkategorikan nodul tiroid berdasarkan karakteristik sonografisnya menjadi beberapa kategori risiko keganasan. Beberapa versi TIRADS telah dikembangkan oleh berbagai organisasi profesi di dunia, di antaranya *American College of Radiology* TIRADS (ACR-TIRADS), *European Thyroid Association* TIRADS (EU-TIRADS), *Korean Thyroid Imaging Reporting and Data System* (K-TIRADS), serta *Chinese Thyroid Imaging Reporting and Data System* (C-TIRADS) yang diterbitkan pada tahun 2020 (Zhou et al., 2020). Setiap sistem memiliki kriteria penilaian komponen sonografi yang sedikit berbeda, namun pada dasarnya menilai komposisi, ekogenisitas, bentuk, batas tepi, dan fokus ekogenik nodul untuk kemudian dijumlahkan menjadi suatu skor yang menentukan kategori risiko keganasan, mulai dari kategori 1 (jinak) hingga kategori 5 (sangat mencurigakan keganasan).

Kategori TIRADS 1 hingga 4 merupakan area yang paling menantang dalam praktik klinis. Kategori 1 dan 2 pada umumnya diasosiasikan dengan risiko keganasan yang sangat rendah sehingga tidak memerlukan FNAB, sedangkan kategori 5 memiliki risiko keganasan tinggi yang hampir selalu mengindikasikan biopsi. Namun demikian, kategori 3 dan khususnya kategori 4 berada pada zona abu-abu diagnostik, dengan risiko keganasan yang bertingkat dan tumpang tindih dengan kategori lain, sehingga keputusan klinis mengenai perlu tidaknya biopsi maupun tindak lanjut menjadi tidak selalu jelas. Studi oleh (Hekimsoy et al., 2021) yang memvalidasi ACR-TIRADS dan EU-TIRADS terhadap 251 nodul dengan diagnosis histopatologis definitif melaporkan sensitivitas dan spesifisitas ACR-TIRADS masing-masing sebesar 71% dan 75%, sedangkan EU-TIRADS sebesar 73% dan 80%, dengan luas area di bawah kurva (AUC) sebesar 0,78 dan 0,80. Studi tersebut menyimpulkan bahwa kinerja diagnostik kedua sistem tergolong moderat, sehingga klinisi perlu memahami kekuatan dan keterbatasan masing-masing sistem dalam pengelolaan optimal nodul tiroid.

Tantangan lain yang memperkuat perlunya validasi histopatologis yang komprehensif adalah adanya variasi kinerja diagnostik antar sistem TIRADS. Zhu et al. (2021) yang membandingkan C-TIRADS, ACR-TIRADS, pedoman *American Thyroid Association* (ATA), dan K-TIRADS pada lebih dari dua ribu nodul tiroid dengan diagnosis histopatologis maupun histopatologis menyimpulkan bahwa C-TIRADS menunjukkan kinerja diagnostik yang tinggi dengan tingkat biopsi yang tidak perlu relatif rendah dibandingkan tiga pedoman lainnya. Sejalan dengan hal tersebut, suatu *network meta-analysis* oleh (Słowińska-Klencka et al., 2022; Yang et al., 2023) yang membandingkan EU-TIRADS, C-TIRADS, ACR-TIRADS, dan perangkat lunak bantu diagnosis berbasis kecerdasan buatan (S-Detect) turut menegaskan adanya perbedaan bermakna dalam sensitivitas, spesifisitas, dan akurasi di antara berbagai sistem stratifikasi risiko ultrasonografi tersebut. Perbedaan kinerja ini menimbulkan konsekuensi klinis yang penting, karena keputusan biopsi maupun pembedahan pada akhirnya akan mempengaruhi luaran pasien serta beban biaya kesehatan.

Standar emas dalam menentukan sifat jinak atau ganas suatu nodul tiroid tetaplah pemeriksaan histopatologis, baik yang diperoleh melalui FNAB maupun spesimen operasi definitif. Klasifikasi histopatologis nodul tiroid sendiri telah mengalami pembaruan signifikan melalui edisi kelima klasifikasi *World Health Organization* (WHO) tahun 2022, yang memperkenalkan kategori baru berupa tumor jinak, neoplasma berisiko rendah, dan neoplasma ganas, disertai sistem gradasi histologis berdasarkan hitung mitosis, nekrosis, dan indeks Ki-67 (Juhlin et al., 2022; Jung et al., 2022). Selain itu, sistem pelaporan sitopatologi

Bethesda juga telah direvisi menjadi edisi ketiga pada tahun 2023, dengan penyederhanaan subkategori *atypia of undetermined significance* serta penyesuaian risiko keganasan (*risk of malignancy*) pada setiap kategori diagnostik berdasarkan data mutakhir (Ali & Cibas, 2023). Pembaruan-pembaruan histopatologis ini menjadikan validasi ulang sistem TIRADS terhadap standar baku emas histopatologis yang termutakhir menjadi semakin relevan dan mendesak untuk dilakukan.

Permasalahan lain yang tidak kalah penting adalah variabilitas interobserver dalam penerapan sistem TIRADS. Tinjauan sistematis dan meta-analisis oleh W. Li et al., (2022) terhadap tiga belas studi dengan lebih dari lima ribu nodul melaporkan bahwa kesepakatan antar pengamat (*inter-reader agreement*) untuk klasifikasi ACR TI-RADS secara keseluruhan hanya berada pada kategori sedang (kappa = 0,51; interval kepercayaan 95% 0,42–0,59), dengan tingkat keganasan nodul sebagai faktor yang berpengaruh signifikan terhadap heterogenitas hasil. Variabilitas ini menegaskan bahwa akurasi diagnostik TIRADS sangat bergantung pada pengalaman operator serta konteks populasi yang diperiksa, sehingga hasil validasi dari satu senter tidak selalu dapat digeneralisasikan ke senter lain dengan karakteristik populasi dan sumber daya yang berbeda.

Di Indonesia, khususnya di wilayah Sulawesi Utara, data mengenai kesesuaian kategori TIRADS dengan hasil histopatologis definitif pada nodul tiroid, terutama pada kategori 1 hingga 4, masih sangat terbatas. Sebagian besar bukti ilmiah yang tersedia berasal dari populasi Asia Timur, Eropa, dan Timur Tengah dengan karakteristik demografis, prevalensi defisiensi iodium, dan pola rujukan klinis yang berbeda dari populasi lokal. Padahal, keputusan klinis mengenai perlu tidaknya FNAB pada nodul kategori 3 dan 4 memiliki implikasi langsung terhadap beban tindakan invasif, biaya pelayanan kesehatan, serta kecemasan pasien. Ketidaksesuaian antara kategori TIRADS yang ditetapkan secara sonografis dengan hasil histopatologis definitif berpotensi menyebabkan dua konsekuensi klinis yang merugikan, yaitu keterlambatan diagnosis keganasan pada nodul yang dianggap berisiko rendah (negatif palsu), maupun tindakan biopsi dan pembedahan yang tidak perlu pada nodul yang sesungguhnya jinak (positif palsu).

Berdasarkan uraian di atas, penelitian ini bermaksud melakukan validasi histopatologis yang komprehensif terhadap sistem klasifikasi TIRADS kategori 1 hingga 4 pada pasien dengan nodul tiroid yang menjalani FNAB dan/atau tindakan operatif di Rumah Sakit Pendidikan Fakultas Kedokteran Universitas Sam Ratulangi Manado yaitu RSUP Prof Dr. R. D. Kandou. Penelitian ini diharapkan dapat memberikan bukti ilmiah lokal mengenai akurasi diagnostik TIRADS, mengidentifikasi ambang batas kategori yang paling optimal untuk indikasi biopsi, serta memberikan rekomendasi tatalaksana berbasis bukti yang lebih tepat dalam pengelolaan nodul tiroid di lingkungan pelayanan kesehatan setempat, sehingga dapat meminimalkan tindakan diagnostik dan operatif yang tidak perlu tanpa mengorbankan sensitivitas deteksi keganasan.

Berdasarkan latar belakang yang telah diuraikan, penelitian ini berupaya menjawab beberapa pertanyaan utama. Pertama, penelitian ini akan mengkaji bagaimana distribusi kategori TIRADS 1–4 pada pasien dengan nodul tiroid yang menjalani pemeriksaan ultrasonografi dan FNAB dan/atau tindakan operatif. Kedua, penelitian ini akan menganalisis korelasi antara kategori TIRADS 1–4 dengan hasil pemeriksaan histopatologis definitif pada nodul tiroid. Ketiga, penelitian ini bertujuan untuk mengukur nilai sensitivitas,

spesifisitas, nilai duga positif, nilai duga negatif, akurasi, serta luas area di bawah kurva (AUC) dari sistem klasifikasi TIRADS dalam memprediksi keganasan nodul tiroid jika dibandingkan dengan baku emas histopatologis. Keempat, penelitian ini akan mengevaluasi proporsi keganasan pada masing-masing kategori TIRADS 1, 2, 3, dan 4, dengan perhatian khusus pada kategori 4 yang memiliki heterogenitas risiko paling tinggi. Terakhir, penelitian ini akan mengidentifikasi apakah terdapat ambang batas (cut-off) kategori TIRADS yang optimal untuk menentukan indikasi FNAB pada populasi penelitian ini. Secara umum, penelitian ini bertujuan untuk mengetahui validitas dan akurasi diagnostik sistem klasifikasi TIRADS kategori 1–4 dibandingkan dengan hasil pemeriksaan histopatologis sebagai baku emas dalam penatalaksanaan nodul tiroid. Secara khusus, penelitian ini memiliki beberapa tujuan yang lebih terperinci.

Penelitian ini bertujuan untuk mengetahui karakteristik demografis dan klinis pasien dengan nodul tiroid yang menjadi subjek penelitian, serta mengetahui distribusi frekuensi kategori TIRADS 1, 2, 3, dan 4 pada subjek tersebut. Selain itu, penelitian ini juga bertujuan untuk mengetahui distribusi hasil histopatologis (jinak dan ganas) berdasarkan klasifikasi WHO 2022 pada setiap kategori TIRADS. Lebih lanjut, penelitian ini akan menghitung nilai sensitivitas, spesifisitas, nilai duga positif, nilai duga negatif, dan akurasi diagnostik TIRADS terhadap hasil histopatologis, serta menganalisis luas area di bawah kurva Receiver Operating Characteristic (AUC-ROC) sistem TIRADS dalam memprediksi keganasan nodul tiroid. Penelitian ini juga akan menentukan ambang batas kategori TIRADS yang optimal sebagai indikasi FNAB berdasarkan analisis titik potong (cut-off point) pada populasi penelitian, dan menganalisis proporsi biopsi yang tidak perlu (unnecessary FNAB rate) apabila indikasi biopsi didasarkan pada masing-masing kategori TIRADS.

METODE

Penelitian ini menggunakan desain studi diagnostik dengan pendekatan potong lintang analitik (*cross-sectional analytic diagnostic study*), yang membandingkan hasil klasifikasi TIRADS berdasarkan pemeriksaan ultrasonografi tiroid dengan hasil pemeriksaan histopatologis definitif sebagai baku emas. Pengambilan data dilakukan secara retrospektif terhadap rekam medis dan arsip hasil pemeriksaan pasien yang telah menjalani ultrasonografi tiroid disertai FNAB dan/atau tindakan operatif dalam periode waktu tertentu, serta dapat dilengkapi dengan pengambilan data prospektif pada periode penelitian berjalan untuk memenuhi target jumlah sampel minimal.

Tempat dan Waktu Penelitian

Penelitian ini akan dilaksanakan di RSUP Prof. Dr. R. D. Kandou Manado dan/atau rumah sakit jejaring pendidikan Fakultas Kedokteran Universitas Sam Ratulangi Manado yang memiliki fasilitas ultrasonografi tiroid, tindakan FNAB, layanan bedah onkologi, serta laboratorium patologi anatomi. Penelitian direncanakan berlangsung selama kurang lebih 1 bulan, terhitung sejak penyusunan proposal, pengumpulan data, hingga analisis dan penyusunan laporan akhir, pada tahun 2026.

Populasi dan Sampel Penelitian

Populasi target penelitian ini adalah seluruh pasien dengan nodul tiroid yang menjalani pemeriksaan ultrasonografi tiroid. Populasi terjangkau adalah pasien dengan nodul tiroid yang menjalani pemeriksaan ultrasonografi tiroid disertai FNAB dan/atau tindakan

operatif di RSUP Prof. Dr. R. D. Kandou Manado dan rumah sakit jejaring pada periode penelitian. Sampel penelitian dipilih dari populasi terjangkau menggunakan metode consecutive sampling, yaitu seluruh subjek yang memenuhi kriteria inklusi dan tidak memenuhi kriteria eksklusi diikutsertakan secara berurutan hingga jumlah sampel minimal terpenuhi.

Kriteria Sampel

Kriteria Inklusi

1. Pasien dengan nodul tiroid yang telah menjalani pemeriksaan ultrasonografi tiroid dengan pelaporan kategori TIRADS yang lengkap.
2. Pasien yang telah menjalani FNAB dan/atau tindakan operatif (tiroidektomi parsial/total) dengan hasil pemeriksaan histopatologis definitif yang tersedia dan lengkap dalam rekam medis.
3. Interval waktu antara pemeriksaan ultrasonografi dan tindakan FNAB/operasi tidak lebih dari 3 bulan.
4. Usia pasien ≥ 18 tahun.
5. Bersedia menjadi subjek penelitian dan menandatangani lembar persetujuan (informed consent) untuk komponen data yang diambil secara prospektif.

Kriteria Eksklusi

6. Data rekam medis tidak lengkap, terutama tidak tersedianya gambar/laporan ultrasonografi asli atau laporan histopatologis definitif.
7. Pasien dengan riwayat tindakan pada kelenjar tiroid sebelumnya (pasca tiroidektomi parsial, ablasi radiofrekuensi, atau terapi iodium radioaktif) yang dapat memengaruhi gambaran sonografis maupun histopatologis.
8. Pasien dengan kualitas gambar ultrasonografi yang tidak adekuat untuk penilaian kategori TIRADS secara retrospektif.
9. Pasien dengan diagnosis karsinoma tiroid meduler atau anaplastik yang telah diketahui sebelum pemeriksaan ultrasonografi indeks (untuk menghindari bias verifikasi).
10. Pasien hamil pada saat pemeriksaan (dipertimbangkan terpisah mengingat perubahan fisiologis kelenjar tiroid selama kehamilan).

Besar Sampel

Perhitungan Jumlah Sampel Minimal

Perhitungan besar sampel pada penelitian diagnostik ini menggunakan rumus (Buderer, 1996; Chen et al., 2022) yang mengintegrasikan perkiraan sensitivitas, spesifisitas, dan prevalensi keganasan pada populasi penelitian, sebagaimana telah diadopsi secara luas pada studi validasi diagnostik kontemporer. Rumus yang digunakan adalah sebagai berikut:

$$n\ (\text{sensitivitas}) = Z^2 \times Sn \times (1 - Sn) / [W^2 \times P]$$

$$n\ (\text{spesifisitas}) = Z^2 \times Sp \times (1 - Sp) / [W^2 \times (1 - P)]$$

Keterangan:

Z = nilai baku distribusi normal pada tingkat kepercayaan 95% (Z = 1,96); Sn = perkiraan sensitivitas;

Sp = perkiraan spesifisitas;

W = presisi absolut yang diinginkan (lebar interval kepercayaan); P = perkiraan prevalensi keganasan pada populasi penelitian.

Berdasarkan tinjauan pustaka, perkiraan sensitivitas TIRADS kategori ≥4 terhadap

keganasan histopatologis ditetapkan sebesar 80% dan spesifisitas sebesar 75%, mengacu pada rentang nilai yang dilaporkan oleh Hekimsoy et al. (2021) dan studi validasi serupa. Perkiraan prevalensi keganasan pada populasi nodul tiroid yang menjalani FNAB dan/atau operasi ditetapkan sebesar 25%.

Mengingat keterbatasan waktu dan biaya pada tahap awal penelitian (studi pendahuluan/pilot), presisi absolut (*W*) tidak lagi ditetapkan sebesar 10%, melainkan diperlonggar menjadi 30% (0,30)agar jumlah sampel tetap memenuhi kaidah perhitungan dan lebih feasible secara operasional. Tingkat kepercayaan tetap dipertahankan pada 95% (Z = 1,96).

Perhitungan sampel untuk sensitivitas (*W* = 30%): $n = (1{,}96)^2 \times 0{,}80 \times 0{,}20 / (0{,}30^2 \times 0{,}25)$

$n = 0{,}6147 / 0{,}0225 \approx 27{,}3 \rightarrow$ dibulatkan menjadi 28 subjek.

Perhitungan sampel untuk spesifisitas (*W* = 30%): $n = (1{,}96)^2 \times 0{,}75 \times 0{,}25 / (0{,}30^2 \times 0{,}75)$

$n = 0{,}7203 / 0{,}0675 \approx 10{,}7 \rightarrow$ dibulatkan menjadi 11 subjek.

Berdasarkan kedua perhitungan tersebut, digunakan nilai yang lebih besar sebagai jumlah sampel minimal yang dibutuhkan, yaitu 28 subjek. Dengan mempertimbangkan kemungkinan data tidak lengkap atau *drop-out* sebesar 10%, maka jumlah sampel minimal yang ditetapkan adalah:

$28 + (10\% \times 28) = 28 + 2{,}8 = 30{,}8 \approx 30$ subjek (nodul tiroid).

Cara Pengambilan Data

Data penelitian dikumpulkan melalui penelusuran rekam medis elektronik maupun manual pasien yang memenuhi kriteria inklusi, meliputi data demografis (usia, jenis kelamin), hasil laporan ultrasonografi tiroid beserta kategori TIRADS yang telah ditetapkan oleh dokter spesialis radiologi, hasil sitologi FNAB (bila tersedia), serta hasil pemeriksaan histopatologis definitif dari dokter spesialis patologi anatomi. Untuk data yang diambil secara retrospektif, gambar ultrasonografi yang tersimpan dalam sistem *Picture Archiving and Communication System* (PACS) atau arsip digital akan ditinjau ulang secara independen oleh dua orang penilai (dokter spesialis radiologi/bedah yang terlatih) yang tidak mengetahui hasil histopatologis (blinded), untuk menentukan atau mengonfirmasi kategori TIRADS sesuai kriteria ACR-TIRADS, guna meminimalkan bias pelaporan awal dan menilai konsistensi antar penilai (interobserver agreement) menggunakan statistik kappa. Seluruh data kemudian dicatat dalam formulir pengumpulan data (*case report form*) yang telah distandarisasi sebelum dilakukan entri dan analisis statistik.

Analisis Data

Analisis data dilakukan menggunakan perangkat lunak statistik standar (SPSS/STATA/R). Karakteristik demografis dan klinis subjek penelitian dianalisis secara deskriptif, dengan data kategorik disajikan dalam bentuk frekuensi dan persentase, sedangkan data numerik disajikan dalam bentuk rerata dan simpangan baku (apabila berdistribusi normal) atau median dan rentang interkuartil (apabila terdistribusi tidak normal), yang ditentukan melalui uji normalitas *Kolmogorov-Smirnov* atau *Shapiro-Wilk*. Analisis korelasi antara kategori TIRADS dengan hasil histopatologis dilakukan menggunakan uji *Chi-Square* atau uji eksak Fisher apabila terdapat sel dengan nilai harapan

kurang dari 5, disertai analisis tren (*Chi-Square for trend*) untuk menilai pola peningkatan proporsi keganasan seiring peningkatan kategori TIRADS.

Kinerja diagnostik sistem TIRADS dihitung menggunakan tabel kontingensi 2x2 untuk memperoleh nilai sensitivitas, spesifisitas, nilai duga positif, nilai duga negatif, rasio kemungkinan positif dan negatif (*likelihood ratio*), serta akurasi diagnostik keseluruhan, disertai interval kepercayaan 95%. Analisis *kurva Receiver Operating Characteristic* (ROC) dilakukan untuk menghitung luas area di bawah kurva (AUC) serta menentukan titik potong (*cut-off*) kategori TIRADS yang optimal berdasarkan indeks Youden. Kesepakatan antar penilai (*interobserver agreement*) dalam penentuan kategori TIRADS dianalisis menggunakan statistik kappa Cohen, dengan interpretasi kekuatan kesepakatan mengikuti kriteria Landis dan Koch. Nilai $p < 0,05$ dianggap bermakna secara statistik pada seluruh analisis.

Alur Penelitian



Gambar 1 Alur Penelitian

HASIL DAN PEMBAHASAN

Unit analisis adalah pasien. Karena hasil patologi anatomi pada data tersedia per pasien atau per spesimen dan tidak selalu per nodul, setiap pasien diklasifikasikan menurut kategori TIRADS tertinggi yang dilaporkan pada ultrasonografi (pendekatan *worst-category*). Bila laporan menuliskan rentang kategori (misalnya TIRADS 2–3), digunakan batas atas. Kategori TIRADS diambil sebagaimana tertulis pada laporan ultrasonografi asal tanpa penilaian ulang, sehingga versi sistem TIRADS yang digunakan pemeriksa tidak dapat diverifikasi.

Karakteristik Sampel Penelitian

Seleksi subjek ditunjukkan pada Tabel 5.1 dan Gambar 5.1. Dari 30 entri, satu entri

merupakan duplikat sehingga terdapat 29 pasien unik. Delapan pasien dikeluarkan karena tidak memenuhi kriteria inklusi dan eksklusi pada Subbab 4.4, sehingga 21 pasien dianalisis (10 ganas dan 11 jinak).

Tabel 1 Alur Seleksi Subjek Penelitian

Tahap	n	Keterangan
Entri data awal	30	Seluruh entri pada daftar subjek
Duplikat	1	Entri identik (hasil USG dan PA sama persis)
Pasien unik	29	
Dikeluarkan: kategori TIRADS tidak dilaporkan	4	Kriteria inklusi 1 tidak terpenuhi
Dikeluarkan: PA bukan dari nodul tiroid atau non-diagnostik	2	Satu PA kelenjar getah bening; satu biopsi tanpa sarang tumor
Dikeluarkan: PA tidak berpasangan dengan nodul indeks	1	USG dilakukan setelah lobektomi; nodul yang dinilai USG bukan spesimen PA
Dikeluarkan: selang USG–PA > 3 bulan	1	Selang sekitar 12 bulan menurut tanggal USG dan tanggal operasi
Subjek dianalisis	**21**	**Ganas 10; jinak 11**

Keterangan: tanggal PA hanya tercatat pada 5 dari 21 subjek yang dianalisis (selang USG–PA 11–64 hari; pada satu subjek PA mendahului USG sekitar 3 minggu). Pada subjek lain selang waktu tidak dapat diverifikasi.

Karakteristik subjek disajikan pada Tabel 5.2. Data usia dan jenis kelamin tidak tercantum pada daftar subjek sehingga tidak dapat dianalisis; kedua variabel ini perlu dilengkapi dari rekam medis sebelum laporan akhir. Diameter terbesar nodul hanya terdokumentasi pada 10 dari 21 subjek. Uji Shapiro–Wilk menunjukkan distribusi normal ($W = 0{,}946$; $p = 0{,}626$), sehingga ukuran disajikan sebagai rerata dan simpangan baku. Kategori TIRADS 5 ditemukan pada 3 pasien; karena termasuk definisi operasional (Tabel 4.1) dan untuk menghindari seleksi data, kategori ini turut dianalisis. Analisis yang dibatasi pada kategori 1–4 sesuai judul penelitian disajikan sebagai analisis sensitivitas pada Subbab 2.

Tabel 2 Karakteristik Subjek Penelitian (n = 21)

Variabel	Hasil
Usia	Tidak tercatat pada data
Jenis kelamin	Tidak tercatat pada data
Kategori TIRADS tertinggi, n (%)	
TIRADS 1	2 (9,5)
TIRADS 2	4 (19,0)
TIRADS 3	6 (28,6)
TIRADS 4	6 (28,6)
TIRADS 5	3 (14,3)
Hasil histopatologi, n (%)	
Ganas	10 (47,6)
Jinak	11 (52,4)
Diameter terbesar nodul (n = 10), cm	

Rerata ± SD	3,59 ± 1,60
Median (IQR)	3,60 (2,97–4,48)
Rentang	0,93–5,90
Jinak (n = 5), median (IQR)	3,60 (3,60–3,80)
Ganas (n = 5), median (IQR)	3,50 (2,80–5,60)
Perbandingan jinak vs ganas	Mann–Whitney U = 12,0; p = 1,000

Keterangan: SD = simpangan baku; IQR = rentang interkuartil.

Analisis kesepakatan antar penilai (kappa Cohen) tidak dapat dilaksanakan karena daftar subjek hanya memuat satu kategori TIRADS per pasien dari laporan asal dan tidak memuat penilaian ulang oleh penilai kedua yang tersamar. Analisis ini perlu dikerjakan pada tahap penelitian lanjutan.

Distribusi Kategori TIRADS 1–4 dan Hasil Histopatologis

Distribusi kategori TIRADS pada 21 subjek adalah TIRADS 1 sebanyak 2 pasien (9,5%), TIRADS 2 sebanyak 4 pasien (19,0%), TIRADS 3 sebanyak 6 pasien (28,6%), TIRADS 4 sebanyak 6 pasien (28,6%), dan TIRADS 5 sebanyak 3 pasien (14,3%). Sebanyak 12 pasien (57,1%) berada pada kategori 3–4 yang merupakan zona abu-abu diagnostik. Secara histopatologis, 10 pasien (47,6%; IK 95% 25,7–70,2) menderita keganasan dan 11 pasien (52,4%) jinak.

Rincian diagnosis histopatologis berdasarkan terminologi WHO 2022 disajikan pada Tabel 5.3. Karsinoma papiler tiroid (seluruh varian) merupakan keganasan terbanyak (8 dari 10; 80,0%), dengan varian folikuler pada 4 pasien. Pada kelompok jinak, struma adenomatosa merupakan diagnosis terbanyak (5 dari 11; 45,5%). Data individual seluruh subjek disajikan pada Tabel 3.

Tabel 3 Distribusi Diagnosis Histopatologis Subjek Penelitian

Diagnosis	n	% dalam kelompok
Jinak (n = 11)		
Struma adenomatosa	5	45,5
Struma koloides (termasuk kistik)	2	18,2
Adenoma folikuler (satu disertai inflamasi kronik)	2	18,2
Follicular nodular disease	1	9,1
Lesi folikuler jinak (sitologi Bethesda II)	1	9,1
Ganas (n = 10)		
Karsinoma papiler varian folikuler (termasuk invasive encapsulated)	4	40,0
Karsinoma papiler klasik/solid/tidak dispesifikasi	3	30,0
Mikrokarsinoma papiler	1	10,0
Karsinoma berdiferensiasi buruk	1	10,0
Karsinoma tiroid, tidak dispesifikasi	1	10,0

Keterangan: satu subjek jinak (P01) memiliki rujukan sitologi FNAB (Bethesda II); subjek lain memiliki rujukan histopatologi.

Tabel 4 Data Individual Subjek Penelitian (n = 21)

Kode	TIRADS tertinggi	Diagnosis patologi anatomi	Hasil
P01	1	Benign follicular lesion (Bethesda II)	Jinak
P02	5	Struma adenomatosa	Jinak
P03	4	Papillary microcarcinoma (lobus & isthmus)	Ganas
P04	3	Thyroid follicular nodular disease	Jinak
P05	3	Adenoma folikuler tiroid	Jinak
P06	4	Poorly differentiated carcinoma	Ganas
P07	4	Follicular variant PTC (bilateral)	Ganas
P08	2	Follicular variant PTC (bilateral)	Ganas
P09	2	Struma adenomatosa	Jinak
P10	4	Adenoma folikuler + inflamasi kronik	Jinak
P11	3	Struma adenomatosa	Jinak
P12	5	Struma adenomatosa	Jinak
P13	4	Invasive encapsulated follicular variant PTC	Ganas
P14	3	Struma koloides cystica	Jinak
P15	2	Struma koloides	Jinak
P16	1	Struma adenomatosa	Jinak
P17	3	Papillary carcinoma	Ganas
P18	5	Papillary carcinoma	Ganas
P19	2	Follicular variant PTC	Ganas
P20	3	Papillary carcinoma (solid & classic variant)	Ganas
P21	4	Carcinoma thyroid	Ganas

Keterangan: kode subjek menggantikan identitas pasien untuk menjaga kerahasiaan.

Korelasi Kategori TIRADS 1–4 dengan Hasil Histopatologis Definitif

Tabulasi silang kategori TIRADS dengan hasil histopatologi disajikan pada Tabel 5.5. Proporsi keganasan berturut-turut 0% (TIRADS 1), 50,0% (TIRADS 2), 33,3% (TIRADS 3), 83,3% (TIRADS 4), dan 33,3% (TIRADS 5).

Tabel 5 Tabulasi Silang Kategori TIRADS dan Hasil Histopatologi

Kategori TIRADS	Jinak, n (%)	Ganas, n (%)	Total
TIRADS 1	2 (100,0)	0 (0,0)	2
TIRADS 2	2 (50,0)	2 (50,0)	4
TIRADS 3	4 (66,7)	2 (33,3)	6
TIRADS 4	1 (16,7)	5 (83,3)	6
TIRADS 5	2 (66,7)	1 (33,3)	3
Total	**11 (52,4)**	**10 (47,6)**	**21**

Keterangan: persentase dihitung per baris.

Sebanyak 10 sel (100%) memiliki frekuensi harapan kurang dari 5 (minimum 0,95), sehingga uji Chi-Square Pearson tidak valid dan sesuai rencana analisis digunakan uji eksak

Fisher–Freeman–Halton (Tabel 5.6). Uji eksak menunjukkan tidak terdapat hubungan bermakna antara kategori TIRADS dan hasil histopatologi (p = 0,282). Uji *Chi-Square for trend* (*linear-by-linear association*) juga tidak bermakna (χ^2 = 1,251; p = 0,263; uji eksak p = 0,289). Korelasi Spearman antara kategori TIRADS dan keganasan lemah dan tidak bermakna (ρ = 0,251; p = 0,272; Tabel 5.7). Cramér's V sebesar 0,518 menunjukkan besar efek sedang, tetapi tanpa kemaknaan statistik akibat jumlah sampel kecil.

Tabel 6 Hasil Uji Chi-Square

Uji	Nilai	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)
Pearson Chi-Square	5,632	4	0,228	—
Likelihood Ratio	6,655	4	0,155	—
Fisher–Freeman–Halton Exact Test	—	—	—	0,282
Linear-by-Linear Association	1,251	1	0,263	0,289
N of Valid Cases	21			

Keterangan: 10 sel (100,0%) memiliki frekuensi harapan < 5; frekuensi harapan minimum 0,95.

Tabel 7 Ukuran Asosiasi

Ukuran	Nilai	Signifikansi	N
Spearman correlation (ρ)	0,251	0,272	21
Cramér's V	0,518	0,228 (aproksimasi)	21

Berdasarkan hasil tersebut, hipotesis H_1 (korelasi bermakna antara kategori TIRADS dan histopatologi) dan H_2 (peningkatan proporsi keganasan yang bermakna seiring peningkatan kategori) tidak didukung secara statistik pada sampel ini. Pola proporsi juga tidak meningkat secara monoton karena kategori 2 memiliki proporsi keganasan 50,0% dan kategori 5 hanya 33,3%.

Kinerja Diagnostik TIRADS terhadap Baku Emas Histopatologis

Kinerja diagnostik dihitung dengan cut-off TIRADS ≥4 sebagai kategori positif (mencurigakan) sesuai definisi operasional pada Tabel 4.1. Tabel kontingensi 2×2 disajikan pada Tabel 5.8 dan ukuran kinerja pada Tabel 5.9.

Tabel 8 Tabel Kontingensi 2×2 TIRADS ≥4 terhadap Histopatologi

TIRADS	Ganas	Jinak	Total
≥ 4 (positif)	6 (TP)	3 (FP)	9
< 4 (negatif)	4 (FN)	8 (TN)	12
Total	**10**	**11**	**21**

Keterangan: uji Fisher eksak 2×2 p = 0,198; odds ratio = 4,00.

Tabel 9 Kinerja Diagnostik TIRADS ≥4 terhadap Baku Emas Histopatologi (n = 21)

Parameter	Nilai	IK 95%
Sensitivitas	60,0%	26,2–87,8%

Spesifisitas	72,7%	39,0–94,0%
Nilai duga positif	66,7%	29,9–92,5%
Nilai duga negatif	66,7%	34,9–90,1%
Akurasi	66,7%	43,0–85,4%
Rasio kemungkinan positif (LR+)	2,20	0,74–6,54
Rasio kemungkinan negatif (LR−)	0,55	0,24–1,28

Keterangan: IK 95% sensitivitas, spesifisitas, nilai duga, dan akurasi menggunakan metode Clopper–Pearson; IK 95% LR menggunakan metode logaritmik.

Pada cut-off ≥4, sensitivitas 60,0% dan spesifisitas 72,7% dengan akurasi 66,7%. Empat dari 10 keganasan (40,0%) bernilai negatif palsu, yaitu dua karsinoma papiler varian folikuler pada TIRADS 2 dan dua karsinoma papiler pada TIRADS 3; tiga pasien jinak bernilai positif palsu (dua struma adenomatosa pada TIRADS 5 dan satu adenoma folikuler dengan inflamasi kronik pada TIRADS 4). Seluruh interval kepercayaan lebar sehingga estimasi bersifat tidak presisi.

Tabel 10 Kinerja Diagnostik pada Berbagai Cut-off Kategori TIRADS

Cut-off	Sensitivitas (%)	Spesifisitas (%)	NDP (%)	NDN (%)	Akurasi (%)
TIRADS ≥ 2	100,0	18,2	52,6	100,0	57,1
TIRADS ≥ 3	80,0	36,4	53,3	66,7	57,1
TIRADS ≥ 4	**60,0**	**72,7**	**66,7**	**66,7**	**66,7**
TIRADS ≥ 5	10,0	81,8	33,3	50,0	47,6

Keterangan: NDP = nilai duga positif; NDN = nilai duga negatif.

Analisis sensitivitas yang dibatasi pada kategori TIRADS 1–4 (n = 18; ganas 9) disajikan pada Tabel 5.11. Uji eksak tetap tidak bermakna (Pearson χ^2 = 5,333; df = 3; p = 3,000), sedangkan uji tren mendekati batas kemaknaan (χ^2 = 3,400; p = 0,065). Pada cut-off ≥4 spesifisitas dan nilai duga positif meningkat, sedangkan sensitivitas tetap rendah.

Tabel 11 Analisis Sensitivitas: Kinerja Diagnostik pada TIRADS 1–4 (n = 18)

Parameter	Cut-off ≥ 3	Cut-off ≥ 4
Sensitivitas, % (IK 95%)	77,8 (40,0–97,2)	55,6 (21,2–86,3)
Spesifisitas, % (IK 95%)	44,4 (13,7–78,8)	88,9 (51,8–99,7)
Nilai duga positif, %	58,3	83,3
Nilai duga negatif, %	66,7	66,7
Akurasi, %	61,1	72,2
Indeks Youden	0,222	0,444
AUC (kategori 1–4)		0,747

Analisis ROC, AUC, dan Penentuan Cut-off Optimal

Analisis kurva ROC pada seluruh subjek menghasilkan luas area di bawah kurva

(AUC) sebesar 0,641 (galat baku 0,126; IK 95% 0,395–0,887; Tabel 5.12). Nilai ini tidak berbeda bermakna dari 0,500 (p = 0,275; uji Mann–Whitney U = 70,5; p = 0,277) dan batas bawah IK 95% berada di bawah 0,70. Dengan demikian hipotesis H_3 (AUC > 0,70) tidak didukung pada sampel ini (uji satu arah p = 0,684). Pada analisis sensitivitas TIRADS 1–4, AUC sebesar 0,747.

Tabel 12 Luas Area di Bawah Kurva ROC

Variabel hasil uji	Area	Galat baku	Asymp. Sig.	IK 95% asimtotik
Kategori TIRADS	0,641	0,126	0,275	0,395–0,887

Keterangan: galat baku menurut DeLong; hipotesis nol AUC = 0,500; variabel status positif = ganas.

Koordinat kurva ROC dan indeks Youden disajikan pada Tabel 5.13 dan Gambar 5.2. Indeks Youden tertinggi dicapai pada cut-off TIRADS ≥4 (J = 0,327; sensitivitas 60,0%; spesifisitas 72,7%), sehingga ambang kategori ≥4 merupakan titik potong optimal pada populasi penelitian ini. Cut-off ≥5 memberikan indeks Youden negatif karena sensitivitas hanya 10,0%.

Tabel 13 Koordinat Kurva ROC dan Indeks Youden

Cut-off	Sensitivitas	1 − Spesifisitas	Spesifisitas	Indeks Youden
≥ 1	1,000	1,000	0,000	0,000
≥ 2	1,000	0,818	0,182	0,182
≥ 3	0,800	0,636	0,364	0,164
≥ 4	0,600	0,273	0,727	0,327
≥ 5	**0,100**	**0,182**	**0,818**	**-0,082**
≥ 6	0,000	0,000	1,000	0,000

Proporsi Keganasan dan Unnecessary FNAB Rate

Proporsi keganasan pada setiap kategori TIRADS beserta interval kepercayaan eksak disajikan pada Tabel 5.14 dan Gambar 5.3. Kategori TIRADS 4 memiliki proporsi keganasan tertinggi (83,3%; 5 dari 6 pasien), tetapi interval kepercayaannya lebar (35,9–99,6%) dan tumpang tindih dengan kategori lain.

Tabel 14 Proporsi Keganasan per Kategori TIRADS

Kategori TIRADS	n	Ganas, n	Proporsi (%)	IK 95% (Clopper–Pearson)
TIRADS 1	2	0	0,0	0,0–84,2
TIRADS 2	4	2	50,0	6,8–93,2
TIRADS 3	6	2	33,3	4,3–77,7
TIRADS 4	6	5	83,3	35,9–99,6
TIRADS 5	3	1	33,3	0,8–90,6
Total	**21**	**10**	**47,6**	**25,7–70,2**

Unnecessary FNAB rate didefinisikan sebagai proporsi nodul jinak di antara seluruh pasien yang memenuhi indikasi biopsi menurut cut-off tertentu (positif palsu dibagi seluruh

positif). Ukuran pelengkap yang dilaporkan adalah proporsi pasien jinak yang akan terbiopsi (1 − spesifisitas) dan proporsi keganasan yang terlewat (Tabel 5.15).

Tabel 15 Beban Biopsi dan Unnecessary FNAB Rate menurut Cut-off TIRADS

Cut-off	Terbiopsi, n (%)	TP	FP	Unnecessary FNAB rate (%)	k terbiopsi (%)	Ganas terlewat, n (%)
≥ 2	19 (90,5)	10	9	47,4	81,8	0 (0,0)
≥ 3	15 (71,4)	8	7	46,7	63,6	2 (20,0)
≥ 4	**9 (42,9)**	**6**	**3**	**33,3**	**27,3**	**4 (40,0)**
≥ 5	3 (14,3)	1	2	66,7	18,2	9 (90,0)

Keterangan: TP = positif benar; FP = positif palsu. Pada cut-off ≥4 IK 95% unnecessary FNAB rate adalah 7,5–70,1%.

Pada cut-off ≥4, sebanyak 9 dari 21 pasien (42,9%) memenuhi indikasi biopsi dan 33,3% di antaranya jinak (3 dari 9). Cut-off yang lebih rendah (≥3 dan ≥2) menaikkan beban biopsi menjadi 71,4% dan 90,5% dengan unnecessary FNAB rate 46,7% dan 47,4%, sedangkan cut-off ≥5 menurunkan beban biopsi menjadi 14,3% tetapi melewatkan 9 dari 10 keganasan. Pada analisis TIRADS 1–4, unnecessary FNAB rate pada cut-off ≥4 sebesar 16,7% (1 dari 6).

Makna Temuan Utama dalam Konteks Klinis dan Teoretis

Penelitian ini mengevaluasi 21 pasien nodul tiroid yang memiliki pasangan data kategori TIRADS dan hasil patologi anatomi. Temuan utamanya adalah sebagai berikut. Pertama, proporsi keganasan mencapai 47,6%. Kedua, tidak terdapat hubungan bermakna antara kategori TIRADS dan hasil histopatologi (uji eksak p = 0,282) maupun tren linear (p = 0,263). Ketiga, pada cut-off ≥4 diperoleh sensitivitas 60,0%, spesifisitas 72,7%, dan akurasi 66,7%. Keempat, AUC sebesar 0,641 (IK 95% 0,395–0,887) tidak mencapai ambang hipotesis H_3. Kelima, cut-off ≥4 memberikan indeks Youden tertinggi dengan unnecessary FNAB rate 33,3%.

Secara teoretis, temuan bahwa kategori 4 memiliki proporsi keganasan tertinggi (83,3%) sejalan dengan konsep pada Subbab 2.1.3 bahwa kategori 4 merupakan zona abu-abu dengan risiko bermakna. Namun pola risiko tidak meningkat bertahap karena kategori 2 memiliki proporsi keganasan 50,0% dan kategori 5 hanya 33,3%, berbeda dari asumsi stratifikasi risiko bertingkat yang mendasari hipotesis H_2. Mengingat jumlah sampel (21) lebih kecil daripada sampel minimal yang dihitung pada Subbab 4.5 (30 subjek) dan interval kepercayaan lebar, hasil ini harus dibaca sebagai data awal (*pilot*) dan bukan sebagai bukti bahwa TIRADS tidak bermanfaat.

Karakteristik Sampel dan Distribusi TIRADS 1–4 pada Populasi Lokal

Sebanyak 57,1% subjek berada pada kategori TIRADS 3–4, yang menegaskan bahwa zona abu-abu diagnostik merupakan bagian terbesar dari kasus yang menjalani pemeriksaan patologi anatomi. Proporsi keganasan 47,6% jauh lebih tinggi daripada perkiraan 25% yang digunakan pada perhitungan besar sampel (Subbab 4.5) dan daripada gambaran umum bahwa lebih dari 90% nodul tiroid bersifat jinak (Borson-Chazot et al., 2022). Prevalensi nodul pada populasi umum memang tinggi dan sebagian besar jinak (Mu et al., 2022; Li et al., 2021),

sehingga angka 47,6% lebih mencerminkan sampel rujukan tersier dengan konfirmasi patologi (bias seleksi dan bias verifikasi sebagaimana dibahas pada Subbab 2.3.5) daripada populasi nodul tiroid secara keseluruhan. Kanker tiroid sendiri termasuk keganasan yang paling sering didiagnosis di dunia (Sung et al., 2021), sehingga data lokal tetap relevan.

Secara histopatologis, karsinoma papiler tiroid mendominasi kelompok ganas (80,0%), dan varian folikuler ditemukan pada 4 dari 10 keganasan. Terminologi pada hasil patologi anatomi, misalnya *follicular nodular disease* dan *invasive encapsulated follicular variant papillary thyroid carcinoma*, konsisten dengan klasifikasi WHO 2022 (Baloch et al., 2022). Data usia dan jenis kelamin tidak tersedia sehingga perbandingan demografis dengan literatur tidak dapat dilakukan, dan laporan ultrasonografi berasal dari beberapa fasilitas pencitraan yang berbeda sehingga heterogenitas pemeriksa dan alat tidak dapat dikesampingkan.

Korelasi TIRADS dengan Histopatologi dan Proporsi Keganasan per Kategori

Tidak adanya hubungan bermakna antara kategori TIRADS dan hasil histopatologi perlu ditafsirkan hati-hati. Besar efek yang teramati sedang (Cramér's V = 0,518), tetapi uji tidak bermakna akibat jumlah sampel kecil, 10 dari 10 sel dengan frekuensi harapan kurang dari 5, dan interval kepercayaan proporsi yang lebar. Analisis yang dibatasi pada kategori 1–4 memperlihatkan tren yang mendekati kemaknaan (p = 0,065), yang mengisyaratkan bahwa kategori 5 ikut mengganggu pola trend pada sampel ini.

Beberapa data individual membantu menjelaskan pola yang tidak monoton. Pertama, dua keganasan pada TIRADS 2 adalah karsinoma papiler varian folikuler. Sesuai Subbab 2.2.3, neoplasma folikuler cenderung tampil dengan gambaran sonografi yang ambigu sehingga mudah mendapat kategori rendah. Kedua, dua dari tiga pasien TIRADS 5 ternyata jinak (struma adenomatosa); kalsifikasi pada struma lama dapat memberi skor tinggi bila terbaca sebagai fokus ekogenik, padahal pada ACR TI-RADS makrokalsifikasi hanya bernilai 1 poin dibandingkan 3 poin untuk mikrokalsifikasi (Tabel 2.1). Ketiga, karena kategori yang digunakan adalah kategori tertinggi per pasien sementara patologi anatomi berasal dari spesimen lobus atau seluruh kelenjar, kategori tinggi pada satu nodul tidak selalu berpasangan dengan nodul yang menjadi sumber diagnosis.

Kinerja Diagnostik TIRADS: Sensitivitas, Spesifisitas, Akurasi, dan AUC

Pada cut-off ≥4, sensitivitas 60,0% dan spesifisitas 72,7% lebih rendah daripada asumsi perhitungan sampel (80% dan 75%) dan daripada hasil validasi ACR-TIRADS oleh Hekimsoy et al. (2021) pada 251 nodul (sensitivitas 71%, spesifisitas 75%, AUC 0,78). AUC pada penelitian ini (0,641) juga lebih rendah, meskipun interval kepercayaan sensitivitas (26,2–87,8%) mencakup nilai 71% sehingga perbedaan tersebut belum dapat dinyatakan bermakna. Sebaliknya, unnecessary FNAB rate pada cut-off ≥4 (33,3%) lebih rendah daripada 61–64% yang dilaporkan Hekimsoy et al. (2021); perbandingan ini terbatas karena perbedaan definisi, populasi, dan jumlah sampel.

Nilai duga negatif 66,7% menunjukkan bahwa kategori di bawah 4 belum dapat digunakan untuk menyingkirkan keganasan pada sampel ini: 4 dari 10 keganasan tidak terjaring, dan menurunkan cut-off menjadi ≥3 hanya mengurangi kasus terlewat menjadi 2 dari 10 dengan peningkatan beban biopsi menjadi 71,4%. Cut-off ≥5 tidak layak karena sensitivitas hanya 10,0%. Temuan bahwa ≥4 merupakan titik potong optimal menurut indeks Youden sejalan dengan batas kategori positif pada definisi operasional penelitian, dan

analisis TIRADS 1–4 memperkuat pola ini (spesifisitas 88,9%; nilai duga positif 83,3%) walaupun sensitivitas tetap rendah (55,6%).

Konsistensi dengan Literatur

Hasil penelitian ini sejalan dengan literatur pada beberapa hal. Kinerja diagnostik TIRADS yang moderat dilaporkan oleh Hekimsoy et al. (2021), dan variasi kinerja antar sistem serta antarpopulasi ditegaskan oleh (Zhu et al., 2021) dan Yang et al. (2023). Kinerja dan tingkat biopsi tidak perlu juga bergantung pada ambang ukuran nodul yang diterapkan (Huh et al., 2020) pada penelitian ini ambang ukuran tidak diterapkan karena data ukuran tidak lengkap, sehingga kinerja yang teramati merupakan kinerja berbasis kategori saja. Konsensus Korean Society of Thyroid Radiology 2021 merevisi kriteria biopsi terutama untuk mengurangi biopsi tidak perlu pada nodul kecil (1–2 cm) dengan tetap mempertahankan sensitivitas (Ha et al., 2021), yang menunjukkan bahwa pertimbangan ukuran dan karakteristik nodul perlu dipadukan dengan kategori.

Variabilitas antar penilai juga relevan. Li et al. (2022) melaporkan kesepakatan antarpengamat ACR TI-RADS hanya sedang (kappa = 0,51). Pada penelitian ini kategori berasal dari laporan berbagai pemeriksa dan tidak dinilai ulang, sehingga variabilitas tersebut dapat ikut mempengaruhi hasil tetapi tidak dapat diukur. Hasil yang berbeda dari literatur adalah AUC yang lebih rendah dan tidak adanya gradasi risiko yang monoton; keduanya paling mungkin dijelaskan oleh jumlah sampel kecil, bias seleksi, dan pendekatan satu kategori per pasien, bukan oleh perbedaan biologis yang telah terbukti. Pembaruan klasifikasi WHO 2022 dan Bethesda 2023 (Ali & Cibas, 2023) tidak dapat dievaluasi lebih jauh karena hanya satu subjek memiliki rujukan sitologi.

Kebaruan dan Implikasi Penelitian

Sepengetahuan penulis berdasarkan penelusuran literatur yang terbatas, penelitian ini merupakan salah satu data awal mengenai kesesuaian kategori TIRADS dengan hasil patologi anatomi pada pasien nodul tiroid di lingkungan pelayanan Sulawesi Utara. Kebaruan utamanya adalah dokumentasi kinerja TIRADS berbasis data lokal beserta identifikasi celah data yang perlu dibenahi, yaitu usia, jenis kelamin, ukuran nodul, tanggal pemeriksaan patologi, dan skor komponen TIRADS yang tidak tercatat lengkap.

Implikasi praktisnya adalah sebagai berikut. Pertama, kategori TIRADS rendah (2–3) tidak boleh dipakai sendirian untuk menyingkirkan keganasan, terutama pada pasien dengan kecurigaan klinis; pada data ini empat keganasan berada pada kategori 2–3. Kedua, cut-off ≥4 dapat dipakai sebagai ambang kerja sementara karena memberi keseimbangan terbaik, namun belum dapat dijadikan dasar perubahan pedoman. Ketiga, penyeragaman format pelaporan USG tiroid (versi sistem, skor komponen, dan ukuran nodul) perlu dilakukan agar validasi lanjutan dapat dikerjakan pada tingkat nodul.

Kekuatan dan Keterbatasan Penelitian

Kekuatan penelitian ini adalah penggunaan data riil dengan rujukan patologi anatomi, penerapan seleksi subjek yang eksplisit dan dapat ditelusuri, penggunaan uji eksak yang sesuai untuk frekuensi harapan kecil, pelaporan interval kepercayaan untuk seluruh ukuran kinerja, serta analisis sensitivitas yang dibatasi pada kategori 1–4 sesuai judul.

Keterbatasan penelitian ini adalah sebagai berikut.

1. Jumlah sampel 21 subjek lebih kecil daripada sampel minimal 30 pada Subbab 4.5, sehingga daya statistik rendah dan interval kepercayaan lebar.

2. Desain retrospektif dengan bias seleksi dan bias verifikasi karena hanya pasien yang menjalani pemeriksaan patologi yang dianalisis; proporsi keganasan (47,6%) tidak mewakili populasi nodul tiroid umum.
3. Kategori TIRADS diambil dari laporan asal tanpa penilaian ulang tersamar; versi sistem TIRADS tidak seragam atau tidak tercantum, dan kesepakatan antarpenilai (kappa) tidak dapat dihitung.
4. Unit analisis adalah pasien dengan kategori tertinggi, bukan nodul; satu kategori yang dilaporkan sebagai rentang (TIRADS 2–3) ditetapkan pada batas atas, dan satu kategori diambil dari diagnosis klinis pada laporan.
5. Usia, jenis kelamin, dan ukuran nodul (hanya 10 dari 21 subjek) tidak lengkap sehingga faktor perancu tidak dapat dikendalikan.
6. Jenis rujukan tidak seragam (satu subjek dengan sitologi Bethesda II) dan selang USG–PA hanya dapat diverifikasi pada 5 subjek.
7. Kategori TIRADS 5 ikut dianalisis walaupun judul membatasi pada kategori 1–4, dan analisis multivariat tidak dapat dilakukan.

KESIMPULAN

Berdasarkan analisis terhadap 21 pasien, kategori terbanyak adalah TIRADS 3 dan 4 (masing-masing 28,6%), sedangkan keganasan ditemukan pada 10 pasien (47,6%), terutama karsinoma papiler tiroid. Data usia dan jenis kelamin tidak tersedia, sementara rerata diameter nodul pada 10 pasien yang terdokumentasi sebesar 3,59 ± 1,60 cm. Tidak ditemukan hubungan bermakna antara kategori TIRADS dan histopatologi (uji eksak $p = 0,282$; uji tren $p = 0,263$). Cut-off optimal TIRADS ≥4 menghasilkan sensitivitas 60,0%, spesifisitas 72,7%, nilai duga positif dan negatif serta akurasi masing-masing 66,7%, indeks Youden 0,327, dan *unnecessary FNAB rate* 33,3%. Nilai AUC sebesar 0,641 (IK 95% 0,395–0,887; $p = 0,275$), sehingga ketiga hipotesis penelitian belum didukung secara statistik. Temuan ini bersifat pendahuluan dan tidak dapat digeneralisasi karena jumlah sampel belum mencapai target.

PENDANAAN

Penelitian ini tidak menerima pendanaan khusus dari lembaga pemerintah, swasta, maupun organisasi lainnya.

KONFLIK KEPENTINGAN

Para penulis menyatakan bahwa tidak terdapat konflik kepentingan dalam pelaksanaan penelitian, penulisan, maupun publikasi artikel ini.

KONTRIBUSI PENULIS

Seluruh penulis berkontribusi dalam penyusunan artikel ini. Penulis 1 Marcel Hariman berperan dalam konseptualisasi penelitian, berperan dalam pengumpulan dan analisis data, berperan dalam interpretasi hasil dan penyusunan naskah, sedangkan penulis lainnya berkontribusi dalam peninjauan, penyuntingan, dan persetujuan akhir naskah.

PERNYATAAN PENGGUNAAN KECERDASAN ARTIFISIAL GENERATIF

Penulis menyatakan bahwa tidak menggunakan kecerdasan artifisial generatif dalam penyusunan artikel ini.

REFERENSI

Ali, S. Z., & Cibas, E. S. (2023). The 2023 Bethesda System for Reporting Thyroid Cytopathology. *Thyroid*, *33*(9), 1039–1044.

Baloch, Z. W., Asa, S. L., Barletta, J. A., Ghossein, R. A., Juhlin, C. C., & Jung, C. K. (2022). Overview of the 2022 WHO classification of thyroid neoplasms. *Endocrine Pathology*, *33*(1), 27–63.

Borson-Chazot, F., Borget, I., Mathonnet, M., & Leenhardt, L. (2022). SFE-AFCE-SFMN 2022 consensus on the management of thyroid nodules: Epidemiology and challenges in the management of thyroid nodules. *Annales d'Endocrinologie*, *83*(6), 367–372.

Buderer, N. M. F. (1996). Statistical methodology: I. Incorporating the prevalence of disease into the sample size calculation for sensitivity and specificity. *Academic Emergency Medicine*, *3*(9), 895–900.

Chen, Y., Gao, Z., He, Y., Mai, W., Li, J., & Zhou, M. (2022). An artificial intelligence model based on ACR TI-RADS characteristics for US diagnosis of thyroid nodules. *Radiology*, *303*(3), 613–619.

Ha, E. J., Chung, S. R., & Na, D. G. (2021). 2021 Korean Thyroid Imaging Reporting and Data System and imaging-based management of thyroid nodules: Korean Society of Thyroid Radiology consensus statement and recommendations. *Korean Journal of Radiology*, *22*(12), 2094–2123.

Hekimsoy, İ., Öztürk, E., Ertan, Y., Orman, M. N., Kavukçu, G., & Özgen, A. G. (2021). Diagnostic performance rates of the ACR-TIRADS and EU-TIRADS based on histopathological evidence. *Diagnostic and Interventional Radiology*, *27*(4), 511–518.

Huh, S., Lee, H. S., Yoon, J., Kim, E.-K., Moon, H. J., & Yoon, J. H. (2020). Diagnostic performances and unnecessary US-FNA rates of various TIRADS after application of equal size thresholds. *Scientific Reports*, *10*(1), 10632.

Juhlin, C. C., Mete, O., & Baloch, Z. W. (2022). The 2022 WHO classification of thyroid tumors: Novel concepts in nomenclature and grading. *Endocrine-Related Cancer*, *30*(2), e220293.

Jung, C. K., Bychkov, A., & Kakudo, K. (2022). Update from the 2022 World Health Organization classification of thyroid tumors: A standardized diagnostic approach. *Endocrinology and Metabolism*, *37*(5), 703–718.

Li, W., Sun, Y., Xu, H., Shang, W., & Dong, A. (2022). Systematic review and meta-analysis of American College of Radiology TI-RADS inter-reader reliability for risk stratification of thyroid nodules. *Frontiers in Oncology*, *12*, 840516.

Li, Y., Jin, C., Li, J., Tong, M., Wang, M., & Huang, J. (2021). Prevalence of thyroid nodules in China: A health examination cohort-based study. *Frontiers in Endocrinology*, *12*, 676144.

Mu, C., Ming, X., Tian, Y., Liu, Y., Yao, M., & Ni, Y. (2022). Mapping global epidemiology of thyroid nodules among general population: A systematic review and meta-analysis. *Frontiers in Oncology*, *12*, 1029926.

Netter, F. H. (2027). *Netter atlas of human anatomy* (8th ed.). Elsevier.

Ovalle, W. K., & Nahirney, P. C. (2021). *Netter's essential histology* (3rd ed.). Elsevier.

Słowińska-Klencka, D., Klencki, M., Duda-Szymańska, J., & Popowicz, B. (2022). Optimization of the management of category III thyroid nodules using repeat FNA and TIRADS. *Cancers*, *14*(18), 4489.

Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, *71*(3), 209–249. https://doi.org/10.3322/caac.21660

Yang, L., Li, C., Chen, Z., He, S., Wang, Z., & Liu, J. (2023). Diagnostic efficiency among Eu-/C-/ACR-TIRADS and S-Detect for thyroid nodules: A systematic review and network meta-analysis. *Frontiers in Endocrinology*, *14*, 1227339.

Zhou, J., Yin, L., Wei, X., Zhang, S., Song, Y., & Luo, B. (2020). 2020 Chinese guidelines for ultrasound malignancy risk stratification of thyroid nodules: The C-TIRADS. *Endocrine*, *70*(2), 256–279.

Zhu, H., Yang, Y., Wu, S., Chen, K., Luo, H., & Huang, J. (2021). Diagnostic performance of US-based FNAB criteria of the 2020 Chinese guideline for malignant thyroid nodules: Comparison with the 2017 American College of Radiology guideline, the 2015 American Thyroid Association guideline, and the 2016 Korean Thyroid Association guideline. *Quantitative Imaging in Medicine and Surgery*, *11*(8), 3604–3618.