

Fuzzy-DATA: A Strength- and Certainty-Calibrated Fuzzy TOPSIS Framework for Decision Support in Differentiated Thyroid Cancer

Fuzzy-DATA: Diferansiye Tiroid Kanserinde Karar Desteği için Öneri Gücü ve Kanıt Kesinliği ile Kalibre Edilmiş Bulanık TOPSIS Çerçevesi

Doğan Şengül^{1*}

1) İstanbul Sabahattin Zaim Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Yazılım Mühendisliği Bölümü, İstanbul, Türkiye

ORCID: 0000-0002-2285-3907, dogan.sengul@izu.edu.tr

Geliş Tarihi: 11 Ağustos 2026 - **Kabul Tarihi:** 1 Eylül 2026

DOI: 10.55205/joctensa.5220262015551

ATIF: Şengül, D., (2026). Fuzzy-DATA: A strength- and certainty-calibrated fuzzy TOPSIS framework for decision support in differentiated thyroid cancer. *Cihannüma Teknoloji Fen ve Mühendislik Bilimleri Dergisi*, 5(2), 112-130.

ÖZET

Klinik uygulama kılavuzları, önerilerin yönünü ve gücünü bildirmekle birlikte, bu kategorik ifadeler bireysel hastaya özgü ve tercihe duyarlı kararları doğrudan çözememektedir. Bu çalışmada, diferansiye tiroid kanseri yönetiminde 2025 Amerikan Tiroid Birliği kılavuzundaki öneri gücünü, kanıt kesinliğini, hastaya uygulanabilirliği ve bağlamsal karar kriterlerini birlikte kullanan açıklanabilir bir bulanık karar çerçevesi (Fuzzy-DATA) geliştirilmektedir. Çerçeve, telafi edilemeyen uygunluk katmanının ardından, öneri gücü ve kanıt kesinliği ile kalibre edilmiş bulanık TOPSIS sıralamasını kullanmaktadır. Öneri gücü, kanıt kesinliği ve uygulanabilirlik üçgensel bulanık sayılarla temsil edilmiş ve yedi fayda yönlü kriter kullanılmaktadır. Çalışmada iki sentetik örnek, küçük ve düşük riskli papiller tiroid kanseri yönetimini ve yüksek riskli hastalıkta ameliyat sonrası radyoaktif iyot seçimini ele almaktadır. Değerlendirme; tek yönlü ağırlık duyarlılığı, model bileşeni çıkarma analizi, Öneri 15 kapsam eşlemesi stres testleri, çalıştırılabilir uygunluk kapısı testleri ve her örnek için 30.000 iterasyonlu koşullu sayısal sağlamlık benzetimini içermektedir. Birinci örnekte aktif izlem 0,709 yakınlık katsayısıyla ilk sırada yer almış ve koşullu birinci sıra sıklığı %99,3 olmuştur. Birincil kapsam-sınırlı eşlemede lobektomi, yaklaşık 0,680 kılavuz uyumu ağırlığında aktif izlemi geçmiştir; ancak bu eşik alternatif Öneri 15 eşlemelerinde değişmiş, güvenilir izlem mümkün olmadığından aktif izlem uygunluk kapısıyla hariç tutulmuştur. İkinci örnekte 100-150 miliküri adjuvan iyot-131 tedavisi 0,686 yakınlık katsayısıyla ilk sırada yer almış ve koşullu birinci sıra sıklığı %92,7 olmuştur. Dirichlet yoğunluğu 50 ile 300 arasında değiştirildiğinde ilk sıradaki seçenek değişmemiştir. Çerçeve, örneklendirici girdiler altında içsel olarak tutarlı ve varsayımları görünür bir davranış göstermektedir. Bulgular, klinik geçerlilikten ziyade hesaplamalı uygulanabilirliği ortaya koymaktadır; bu nedenle, çok disiplinli kalibrasyon ve ileriye dönük doğrulama gereği önem arz etmektedir.

Anahtar Kelimeler: Bulanık TOPSIS, Çok kriterli karar verme, Diferansiye tiroid kanseri, Klinik uygulama kılavuzu, Benzetim

* Sorumlu Yazar: dogan.sengul@izu.edu.tr

ABSTRACT

Clinical practice guidelines communicate both the direction and strength of recommendations, but these categorical labels do not directly resolve preference-sensitive choices for an individual patient. This study develops an interpretable fuzzy decision framework, Fuzzy-DATA, that incorporates recommendation strength, evidence certainty, patient-level applicability and contextual decision criteria within the 2025 American Thyroid Association management pathway for differentiated thyroid cancer. The framework uses a noncompensatory eligibility layer followed by a strength- and certainty-calibrated fuzzy Technique for Order Preference by Similarity to Ideal Solution (TOPSIS). Recommendation strength, certainty and applicability were represented by triangular fuzzy numbers and seven benefit-form criteria were used. Two synthetic showcases addressed a small low-risk papillary thyroid carcinoma and postoperative radioactive iodine selection for high-risk disease. Evaluation comprised one-way weight sensitivity, model ablation, Recommendation 15 scope-mapping stress tests, executable eligibility-gate tests and 30,000-iteration conditional numerical robustness simulation per showcase. In Showcase 1, active surveillance ranked first (closeness coefficient 0.709; conditional top-rank frequency 99.3%). Under the primary scope-qualified mapping, lobectomy overtook active surveillance near a guideline-concordance weight of 0.680, but this threshold changed under alternative scope mappings; unreliable follow-up excluded active surveillance. In Showcase 2, adjuvant iodine-131 at 100-150 millicuries ranked first (closeness coefficient 0.686; conditional top-rank frequency 92.7%). The leading alternative remained unchanged when the Dirichlet concentration varied from 50 to 300. The framework showed internally coherent and assumption-transparent behaviour under the illustrative inputs. The findings establish computational feasibility rather than clinical validity; multidisciplinary calibration and prospective validation are required.

Keywords: Fuzzy TOPSIS, Multi-criteria decision making, Differentiated thyroid cancer, Clinical practice guideline, Simulation

INTRODUCTION

Management of differentiated thyroid cancer (*DTC*) contains several decision nodes at which more than one defensible strategy may exist. Low-risk papillary thyroid carcinoma can involve active surveillance, percutaneous ablation or surgery, whereas postoperative decisions may involve whether radioactive iodine (*RAI*) should be used and, when used, which activity range is appropriate. The 2025 American Thyroid Association (*ATA*) guideline addresses these decisions through the *DATA* pathway (Diagnosis, risk/benefit Assessment, Treatment decisions and response Assessment) and explicitly emphasises individualised, shared patient-clinician decision-making (Ringel et al., 2025).

The guideline reports recommendation direction, recommendation strength and certainty of evidence. Strong recommendations are intended for all or nearly all relevant persons or situations when benefits clearly outweigh harms, generally with at least moderate certainty. Conditional recommendations signal a closer balance of benefits and harms, lower certainty, variation in values and preferences or concerns about feasibility, acceptability, resource use or equity (Andrews et al., 2013; Guyatt et al., 2025; Ringel et al., 2025). These labels are essential for interpretation, but they are not numerical probabilities and do not by

themselves specify how to combine guideline force with patient-specific clinical characteristics, treatment burdens, follow-up reliability, access and preferences.

Fuzzy set theory was developed to represent concepts whose boundaries are gradual rather than crisp (Zadeh, 1965). The Technique for Order Preference by Similarity to Ideal Solution (*TOPSIS*) ranks alternatives according to their proximity to a positive ideal and distance from a negative ideal (Hwang and Yoon, 1981). Its fuzzy extension permits linguistic ratings and criterion importance to be represented using fuzzy numbers, retaining imprecision that would otherwise be lost if ratings were converted to single values at an early stage (Chen, 2000). Fuzzy *TOPSIS* has subsequently been used across diverse decision domains, including healthcare and remains attractive when multiple qualitative criteria must be integrated transparently (Salih et al., 2019; Palczewski and Sařabun, 2019; Alharbi et al., 2024).

This methodological study proposes Fuzzy-DATA, an independent computational extension of the ATA DATA pathway. The framework does not create new clinical recommendations. Instead, it translates reported recommendation metadata into an explicit fuzzy guideline-concordance component, applies noncompensatory eligibility gates, combines this component with patient-context criteria and reports both a ranking and its uncertainty. The objectives were to define the framework mathematically, demonstrate it in two contrasting simulated decision nodes and evaluate whether its behaviour is consistent with the intended distinction between conditional and strong recommendations.

MATERIALS AND METHODS

Study Design and Guideline Basis

This is a methodological development and simulation study. The 2025 ATA adult *DTC* guideline is the sole source for the clinical decision nodes, recommendation direction, recommendation strength, certainty labels, relevant eligibility conditions and typical RAI activity ranges (Ringel et al., 2025). The first showcase uses Recommendations 11, 14 and 15: active surveillance and ultrasound-guided percutaneous ablation may be offered or considered for selected cT1aN0M0 papillary thyroid cancers under conditional recommendations with low-certainty evidence, whereas lobectomy is strongly recommended with moderate-certainty evidence when resection is chosen. Because Recommendation 15 addresses surgical extent after resection has been selected, its use for lobectomy in the flat parent-node showcase was treated as a scope-qualified partial proxy and was stress-tested against alternative scope mappings. The second showcase used Recommendation 32 and the accompanying RAI summary table: RAI is strongly recommended with moderate-certainty evidence after total thyroidectomy in ATA high-risk disease, with a typical iodine-131 (I-131) activity of 100-150 millicuries (mCi). No real patient records, identifiable data or outcomes were used.

Design Principles

Five design requirements were prespecified. First, the method had to preserve the distinction between recommendation direction, normative strength, epistemic certainty and patient-level applicability. Second, explicit contraindications or ineligibility conditions had to operate as noncompensatory gates rather than as low-scoring criteria that could be offset by

unrelated advantages. Third, preference-sensitive factors had to influence ranking when the guideline itself signalled legitimate variation. Fourth, the output had to be explainable through criterion-level ratings, recommendation metadata, ranking margins and uncertainty analyses. Fifth, the model had to permit iterative updating after response assessment, consistent with the *DATA* pathway.

Fuzzy-DATA Architecture

The framework has seven sequential stages (Figure 1). A clinical decision node and its feasible alternatives are first defined. Hard safety and eligibility rules are then applied. For each remaining alternative, recommendation strength, certainty, direction and applicability are fuzzified. Patient-context performance is rated on six additional criteria. Strength-certainty-calibrated fuzzy *TOPSIS* produces a rank and closeness coefficient. Conditional numerical robustness and deterministic one-way sensitivity analyses quantify assumption-dependent uncertainty. Finally, the result is presented for shared decision-making and the model is rerun when new response data become available, re-entering the pathway at the safety and eligibility gates.

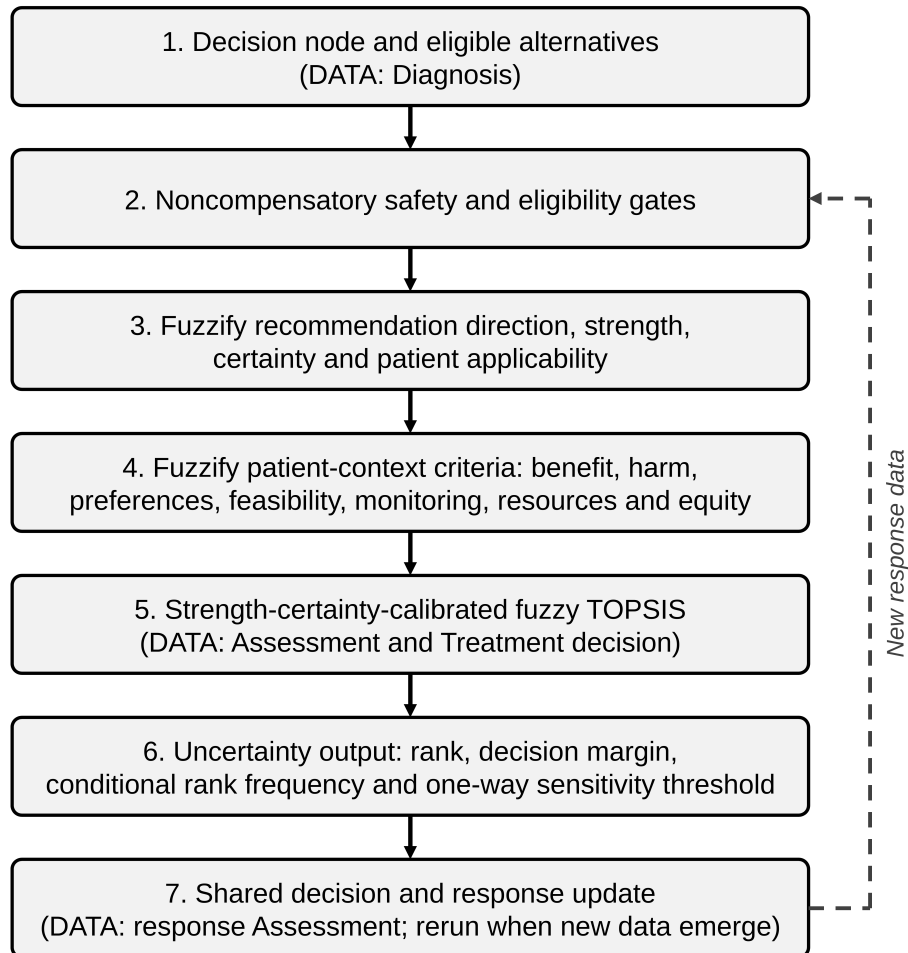


Figure 1. Fuzzy-DATA decision-support architecture. The dashed arrow indicates a rerun with new response data, re-entering at the safety and eligibility gates.

Fuzzy Representation of Recommendation Metadata

A triangular fuzzy number (*TFN*) was written as $\tilde{x} = (l, m, u)$, where l , m and u are the lower, most plausible and upper values on the unit interval $[0,1]$. Each such *TFN* has the membership function

$$\mu_{\tilde{x}}(t) = \begin{cases} \frac{t-l}{m-l} & l \leq t \leq m, \\ \frac{u-t}{u-m} & m < t \leq u, \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

with the convention that a degenerate side ($l = m$ or $m = u$) collapses to the point value. Overlapping *TFNs* were used because the categories are ordered but not equally spaced and because a categorical label does not imply a precise numerical magnitude. The prespecified encoding is shown in Table 1. These values are methodological design parameters, not probabilities of benefit or probabilities that a recommendation is correct.

Table 1. Prespecified triangular fuzzy encodings used in the simulation

Dimension	Category	<i>TFN</i>	Interpretation
Recommendation strength	Strong	(0.80, 0.95, 1.00)	High normative force
Recommendation strength	Conditional	(0.45, 0.65, 0.85)	Substantial but preference-sensitive force
Recommendation strength	Good Practice Statement	(0.80, 0.95, 1.00)	Direct force; not multiplied by certainty
Recommendation strength	No recommendation	(0.00, 0.20, 0.40)	Insufficient guideline support
Evidence certainty	High	(0.85, 0.95, 1.00)	Very high confidence
Evidence certainty	Moderate	(0.65, 0.80, 0.90)	Moderate confidence
Evidence certainty	Low	(0.40, 0.60, 0.75)	Limited confidence
Evidence certainty	Very low	(0.20, 0.40, 0.60)	Very limited confidence
Evidence certainty	Insufficient	(0.00, 0.20, 0.40)	No reliable conclusion
Applicability	Exact	(0.85, 0.95, 1.00)	Population/action closely matches
Applicability	Partial	(0.45, 0.65, 0.85)	Recommendation applies conditionally
Applicability	Poor action/dose match	(0.15, 0.35, 0.55)	Relevant topic but discordant action details
Performance rating	Very low	(0.00, 0.00, 0.25)	<i>VL</i>
Performance rating	Low	(0.00, 0.25, 0.50)	<i>L</i>
Performance rating	Moderate	(0.25, 0.50, 0.75)	<i>M</i>
Performance rating	High	(0.50, 0.75, 1.00)	<i>H</i>
Performance rating	Very high	(0.75, 1.00, 1.00)	<i>VH</i>

Note. Good Practice Statements were represented directly because the *ATA* guideline states that they are not graded with the Grading of Recommendations Assessment, Development and Evaluation (*GRADE*) approach. All encodings are transparent simulation parameters requiring empirical calibration before clinical use. *ATA*, American Thyroid Association; *TFN*, triangular fuzzy number.

For a graded recommendation in favour of alternative i , fuzzy recommendation support was calculated as the componentwise product of recommendation strength $\tilde{S}_i$ and evidence

certainty $\tilde{E}_i$. A recommendation against an alternative used the complement of this product. Patient-specific applicability $\tilde{A}_i$ was then applied:

$$\tilde{B}_i = \tilde{S}_i \otimes \tilde{E}_i \quad (2)$$

where $\otimes$ is an author-defined vertex-wise calibration operator that multiplies corresponding vertices, so that $\tilde{B}_i = (l_{S,i}l_{E,i}, m_{S,i}m_{E,i}, u_{S,i}u_{E,i})$. This vertex-wise operation is a transparent calibration rule and is not the extension-principle product of fuzzy numbers. The applicability term then gives

$$\tilde{G}_i = \tilde{B}_i \otimes \tilde{A}_i \quad (3)$$

For a recommendation against the alternative, $\tilde{B}_i$ was replaced by its complement

$$\tilde{B}_i^- = (1 - u_{B,i}, 1 - m_{B,i}, 1 - l_{B,i}) \quad (4)$$

A Good Practice Statement bypassed the evidence-certainty multiplication and was multiplied only by applicability. The multiplicative form was intentionally conservative: a recommendation received high concordance only when normative strength, evidence certainty and applicability were jointly high. Because strength is partly informed by certainty, the calculation should be interpreted as a transparent calibration rule rather than as statistically independent evidence multiplication. *Proposition 1* confirms that these operations always return valid *TFNs* on the unit interval.

Proposition 1 (closure of the calibration operations). Let $\tilde{S} = (l_S, m_S, u_S)$, $\tilde{E}$ and $\tilde{A}$ be *TFNs* whose components lie in $[0,1]$. Then $\tilde{B} = \tilde{S} \otimes \tilde{E}$, the complement $(1 - u_B, 1 - m_B, 1 - l_B)$ and $\tilde{G} = \tilde{B} \otimes \tilde{A}$ are *TFNs* whose components lie in $[0,1]$.

Proof. Write $\tilde{S} = (l_S, m_S, u_S)$ and $\tilde{E} = (l_E, m_E, u_E)$ with $0 \leq l_S \leq m_S \leq u_S \leq 1$ and likewise for $\tilde{E}$. Because the product of two numbers in $[0,1]$ remains in $[0,1]$, every component of $\tilde{B}$ lies in $[0,1]$. For the ordering, multiplication by a nonnegative number preserves inequalities, so $l_S l_E \leq m_S l_E \leq m_S m_E$ and $m_S m_E \leq u_S m_E \leq u_S u_E$. Hence the components of $\tilde{B}$ are ordered and $\tilde{B}$ is a valid *TFN*. For the complement, the map $t \mapsto 1 - t$ is decreasing and maps $[0,1]$ onto itself; applying it to $l_B \leq m_B \leq u_B$ and reversing the order gives $1 - u_B \leq 1 - m_B \leq 1 - l_B$, again a valid *TFN* in $[0,1]$. The argument for $\tilde{G} = \tilde{B} \otimes \tilde{A}$ repeats the product step with $\tilde{A}$ in place of $\tilde{E}$. ■

Decision Criteria and Baseline Weights

All criteria were transformed to benefit form so that a higher value was preferred. Harm and treatment burden were therefore encoded as harm/burden avoidance. The baseline criterion weights summed to 1.00 (Table 2). Guideline concordance received 0.25 and the remaining 0.75 was distributed across clinical and contextual criteria. The baseline values were set by three explicit considerations rather than by elicitation. First, guideline concordance received the largest single weight (0.25) so that the recommendation component leads the ranking without dominating it: the six contextual criteria jointly retain 0.75. Second, expected disease control received the next largest weight (0.20) as the principal clinical outcome, followed by the two preference-sensitive criteria, harm and burden avoidance and patient preference alignment, weighted equally (0.15 each) to reflect the preference-sensitive character of the target decision nodes. Third, the operational criteria received the remainder

in decreasing order of clinical immediacy: feasibility and acceptability (0.10), monitoring reliability (0.10) and resource and equity fit (0.05). These are illustrative baseline settings for a methodological demonstration, not clinically validated or expert-elicited weights and the one-way and Dirichlet analyses vary them systematically.

Table 2. Decision criteria and baseline weights

Code	Criterion	Weight	Operational meaning
C1	Guideline concordance	0.25	Direction, strength, certainty and applicability
C2	Expected disease control	0.20	Expected ability to control persistence, recurrence or progression
C3	Harm and burden avoidance	0.15	Avoidance of procedural, treatment and long-term burden
C4	Patient preference alignment	0.15	Consistency with an informed patient's goals and values
C5	Feasibility and acceptability	0.10	Availability, expertise, logistics and acceptability
C6	Monitoring reliability	0.10	Ability to perform and adhere to required surveillance
C7	Resource and equity fit	0.05	Resource intensity, access and equity implications

Fuzzy TOPSIS Ranking

Let $\tilde{X} = [\tilde{x}_{ij}]$ be the fuzzy decision matrix for alternatives $i = 1, \dots, m$ and criteria $j = 1, \dots, n$. Because all ratings were already normalised to $[0,1]$, the weighted fuzzy matrix was $\tilde{v}_{ij} = w_j \tilde{x}_{ij}$, evaluated componentwise as

$$\tilde{v}_{ij} = (w_j l_{ij}, w_j m_{ij}, w_j u_{ij}) \quad (5)$$

For each benefit criterion, the fuzzy positive ideal solution (FPIS) was $\tilde{v}_j^+ = (w_j, w_j, w_j)$ and the fuzzy negative ideal solution (FNIS) was $\tilde{v}_j^- = (0, 0, 0)$. Both ideals are fixed endpoints of the weighted rating scale rather than data-derived ideal solutions. The vertex distance (Chen, 2000) between TFNs $\tilde{a}$ and $\tilde{b}$ was written over the vertex index $c \in \{l, m, u\}$ as

$$d(\tilde{a}, \tilde{b}) = \sqrt{\frac{1}{3} \sum_c (c_a - c_b)^2} \quad (6)$$

where c_a and c_b denote the corresponding vertex components of $\tilde{a}$ and $\tilde{b}$. Distances from the positive and negative ideals were summed across criteria to obtain D_i^+ and D_i^- :

$$D_i^+ = \sum_{j=1}^n d(\tilde{v}_{ij}, \tilde{v}_j^+) \quad (7)$$

$$D_i^- = \sum_{j=1}^n d(\tilde{v}_{ij}, \tilde{v}_j^-) \quad (8)$$

The closeness coefficient (CC) (Chen, 2000) was:

$$CC_i = \frac{D_i^-}{D_i^+ + D_i^-} \quad (9)$$

Alternatives were ranked from the highest to the lowest CC. CC is a relative decision index and must not be interpreted as a predicted clinical probability or treatment effect.

Proposition 2 establishes that the index is always well defined and bounded and *Proposition 3* in the *Safety Gates and Dynamic Updating* subsection establishes monotonicity and the noncompensatory character of the eligibility gates.

Proposition 2 (well-definedness and bounds of the closeness coefficient). Let the weights satisfy $w_j \geq 0$ and $\sum_{j=1}^n w_j = 1$. Then for every alternative i , $D_i^+ + D_i^- \geq 1$, so CC_i is well defined and $0 \leq CC_i \leq 1$. Moreover $CC_i = 1$ if and only if $\tilde{v}_{ij} = \tilde{v}_j^+$ for all j and $CC_i = 0$ if and only if $\tilde{v}_{ij} = \tilde{v}_j^-$ for all j .

Proof. The vertex distance in Equation (6) equals $\frac{1}{\sqrt{3}} \|\tilde{a} - \tilde{b}\|$, where $\|\cdot\|$ is the Euclidean norm applied to the vertex vectors (l, m, u) regarded as points of $\mathbb{R}^3$. It therefore inherits the three metric properties: it is nonnegative and vanishes exactly when the two *TfNs* coincide, it is symmetric and it satisfies the triangle inequality. Applying the triangle inequality to the points $\tilde{v}_j^+$, $\tilde{v}_{ij}$ and $\tilde{v}_j^-$ gives $d(\tilde{v}_j^+, \tilde{v}_j^-) \leq d(\tilde{v}_j^+, \tilde{v}_{ij}) + d(\tilde{v}_{ij}, \tilde{v}_j^-)$. Direct evaluation gives $d(\tilde{v}_j^+, \tilde{v}_j^-) = \sqrt{\frac{3w_j^2}{3}} = w_j$. Summing the triangle inequality over $j = 1, \dots, n$ yields $1 = \sum_{j=1}^n w_j \leq D_i^+ + D_i^-$, so the denominator of Equation (9) is at least 1 and CC_i is well defined. Since D_i^+ and D_i^- are sums of nonnegative terms, $0 \leq CC_i \leq 1$ follows immediately. Finally, $CC_i = 1$ holds exactly when $D_i^+ = 0$, which by the identity property of the metric holds exactly when every $\tilde{v}_{ij}$ coincides with $\tilde{v}_j^+$; the characterisation of $CC_i = 0$ is symmetric with $\tilde{v}_j^-$ in place of $\tilde{v}_j^+$. ■

Safety Gates and Dynamic Updating

Before ranking, explicit eligibility and safety conditions were checked. An ineligible alternative was removed rather than penalised. For active surveillance, simulated gate conditions included regional or distant metastases, gross extrathyroidal extension, posterior progression threatening critical structures, primary growth of at least 3 mm, aggressive features, inability to complete reliable follow-up or an informed preference for surgery. This reflects the ATA guidance on indications to pursue resection (Ringel et al., 2025). The supplied Python implementation evaluates these predicates before *TOPSIS* scoring, exports the gate decision and reason and includes eligible, ineligible and boundary-case tests. After an intervention, ATA response category, imaging, thyroglobulin-related data when applicable, adverse effects and updated preferences can be re-entered and the model rerun. These gate conditions are guideline-informed author operationalisations of Recommendation 14 rather than literal transcriptions: patient anxiety is not modelled; regional and distant metastases are merged without the new biopsy-proven qualifier; extrathyroidal extension is restricted to gross extension; posterior growth is narrowed to progression threatening critical structures; and aggressive cytological features are added as an exclusion (Ringel et al., 2025). Each predicate and two combined-failure scenarios are unit-tested in the supplementary code and every rerun re-evaluates all hard gates before any scoring.

Proposition 3 (monotonicity and noncompensatory gating). (i) Fix nonnegative weights with $\sum_{j=1}^n w_j = 1$. If two rating profiles for alternative i satisfy $\tilde{x}_{ij} \leq \tilde{x}'_{ij}$ componentwise for every j , then $CC_i \leq CC'_i$. (ii) An alternative removed by an eligibility gate is excluded from

the ranking for every attainable configuration of ratings and weights; no rating improvement can compensate for a failed gate.

Proof. (i) By Equation (5), each vertex component of $\tilde{v}_{ij}$ has the form $w_j c$ with $c \in [0,1]$. The per-criterion distance to the *FPIS* is $d(\tilde{v}_{ij}, \tilde{v}_j^+) = w_j \sqrt{\frac{1}{3} \sum_c (1 - c_{ij})^2}$ over the vertex components $c_{ij} \in \{l_{ij}, m_{ij}, u_{ij}\}$ and each term is nonincreasing in the corresponding vertex component; the distance to the *FNIS* is $w_j \sqrt{\frac{1}{3} \sum_c c_{ij}^2}$ and each term is nondecreasing. Summing over j shows that D_i^+ cannot increase and D_i^- cannot decrease when ratings improve componentwise. The map $(a, b) \mapsto \frac{b}{a+b}$ on $a, b \geq 0$ with $a+b > 0$ has partial derivatives $-\frac{b}{(a+b)^2} \leq 0$ in a and $\frac{a}{(a+b)^2} \geq 0$ in b , so it is nonincreasing in the first argument and nondecreasing in the second. Composing these monotone dependencies gives $CC_i \leq CC'_i$. (ii) The feasible set is fixed by the gate predicates, which depend only on the recorded clinical eligibility facts and, by construction, not on the criterion ratings or weights. Fuzzy *TOPSIS* is applied to the gated feasible set only, so an excluded alternative receives no rank for any values of the ratings or weights. The ranking function is therefore constant in the ratings of an excluded alternative, which is precisely the noncompensatory property required of a safety gate. ■

Showcase Generation

The showcases were deliberately constructed to span two different guideline states: a preference-sensitive, conditional/low-certainty node and a strong/moderate-certainty node. The profiles, contextual ratings, applicability values, criterion weights and numerical fuzzy encodings were author-assigned synthetic inputs selected before ranking; they were not extracted from patients or inferred from the guideline. The input-provenance summary in Table 10 and the provenance-generating code in *Supplementary Material 1* distinguish every guideline-derived input, modelling assumption and calculated output. In this manuscript the term *synthetic* therefore denotes author-constructed hypothetical patient scenarios and prespecified ratings; it does not refer to artificial observations statistically generated from empirical patient data. Each linguistic rating in Tables 3-4 was assigned by explicit scenario logic: for the guideline-anchored clinical criteria, expected disease control and harm and burden avoidance, the ordering of the ratings follows the direction of the guideline evidence summaries for the scenario in question (for example, lobectomy rated very high and active surveillance high for disease control in Showcase 1), while the contextual criteria encode the stated scenario features, such as follow-up reliability, referral pathways and stated preferences, as illustrative modelling assumptions. No rating is an empirical estimate and the provenance classification in Table 10 records the basis of every input.

Showcase 1 represented a 68-year-old adult with a unifocal 0.8-cm intrathyroidal papillary thyroid carcinoma, cT1aN0M0, no gross extrathyroidal extension or aggressive features, access to an experienced ultrasound team, high follow-up reliability and a strong preference to avoid surgery. Thermal ablation was available only through a more limited referral pathway. The alternatives were active surveillance, ultrasound-guided thermal ablation and thyroid lobectomy (Table 3). Lobectomy received the primary scope-qualified

strong/moderate/partial mapping and three alternative Recommendation 15 mappings were examined separately rather than treating the resulting crossover as a guideline-derived threshold.

Table 3. Compact input matrix for Showcase 1

<i>Alt.</i>	<i>Rec.-Cert.-App.</i>	<i>G centroid</i>	<i>DC</i>	<i>HB</i>	<i>PA</i>	<i>FA</i>	<i>MR</i>	<i>RE</i>
AS	C-L-E	0.387	H	VH	VH	VH	VH	H
TA	C-L-E	0.387	H	H	H	L	M	L
TL	S-M-P*	0.498	VH	M	L	H	H	M

Note. AS, active surveillance; TA, thermal ablation; TL, thyroid lobectomy; S, strong; C, conditional; M, moderate certainty; L, low certainty; E, exact applicability; P, partial applicability; G, guideline concordance; DC, disease control; HB, harm/burden avoidance; PA, preference alignment; FA, feasibility/acceptability; MR, monitoring reliability; RE, resource/equity fit. Linguistic ratings: L, low; M, moderate; H, high; VH, very high. The TL S-M-P entry is a scope-qualified modelling proxy because Recommendation 15 applies when resection has already been selected; alternative scope mappings are reported separately.

Showcase 2 represented a 58-year-old adult after total thyroidectomy for high-risk DTC with gross extrathyroidal extension, bulky lateral nodal disease with extranodal extension and an elevated postoperative thyroglobulin level. There was no severe renal impairment or other major contraindication to RAI and the informed patient accepted adjuvant treatment. Alternatives were no RAI, low-activity I-131 at 30-50 mCi and adjuvant I-131 at 100-150 mCi (Table 4). The lower activity was rated as a poor dose match for high-risk disease even though it retained the general advantages of receiving RAI.

Table 4. Compact input matrix for Showcase 2

<i>Alt. (mCi)</i>	<i>Rec.-Cert.-App.</i>	<i>G centroid</i>	<i>DC</i>	<i>HB</i>	<i>PA</i>	<i>FA</i>	<i>MR</i>	<i>RE</i>
No RAI	S-M-E against	0.264	L	VH	L	VH	M	VH
30-50	S-M-Poor	0.280	H	H	H	H	H	H
100-150	S-M-E	0.688	VH	M	H	H	H	M

Note. The recommendation in favour of RAI was complemented for the 'No RAI' alternative. Poor indicates a poor action/dose match. RAI, radioactive iodine; mCi, millicuries. Other abbreviations are defined below Table 3.

Evaluation Strategy

The evaluation comprised deterministic baseline ranking, model ablation, one-way sensitivity analysis, Recommendation 15 scope-mapping stress tests, conditional numerical robustness analysis and an eligibility-gate stress test. Model ablation compared a context-only model ($\lambda = 0$), the proposed baseline model ($\lambda = 0.25$) and a guideline-only model ($\lambda = 1.00$). In the one-way sensitivity analysis, λ was varied while the other six criterion weights were rescaled proportionally to retain their baseline relative proportions and approximate crossover points were identified on a 0.0001 grid. The scope of Recommendation 15 was stress-tested under four mappings: the primary scope-qualified partial proxy, full direct transfer, parent-node harmonisation and no direct parent-node credit.

Conditional numerical robustness was evaluated using a fixed-seed Monte Carlo simulation with 30,000 iterations per showcase. At each iteration, one score was drawn for every alternative-criterion TFN from a triangular distribution defined by its lower, modal and upper values. Criterion weights were sampled from a Dirichlet distribution centred on the baseline weight vector. The primary concentration parameter was 150, with sensitivity analyses at 50, 100 and 300. Sampled distances were calculated criterion by criterion and

summed in accordance with Equations (6)-(9). The proportion of iterations in which an alternative ranked first was reported as the conditional top-rank frequency; this measure quantifies numerical ranking stability under the prespecified assumptions and is not a clinical probability, a treatment-success probability or evidence of clinical superiority. In a separate gate stress test, follow-up reliability was set to unavailable in Showcase 1, causing active surveillance to be excluded before ranking. All numerical analyses were implemented in Python 3.13 using NumPy and pandas. Matplotlib was used separately for figure preparation; the complete numerical-analysis script, tests and outputs accompany the manuscript.

RESULTS

Baseline Rankings and Conditional Robustness

The baseline results and conditional numerical robustness summaries are presented in Table 5. In Showcase 1, active surveillance ranked first with CC 0.709, a deterministic margin of 0.127 over thyroid lobectomy. Active surveillance had a conditional top-rank frequency of 99.3%, whereas lobectomy and thermal ablation were first in 0.7% and 0.1% of iterations, respectively. The 5th-95th percentile interval for the active-surveillance CC was 0.674-0.807 and the mean sampled top-versus-second margin was 0.136. These conditional top-rank frequencies quantify the numerical stability of the ranking under the prespecified synthetic inputs and distributional assumptions; they are not clinical probabilities and not probabilities of treatment success.

Table 5. Baseline fuzzy *TOPSIS* and conditional robustness results

Showcase	Alternative	Rank	Baseline CC	Cond. top-rank freq.	Cond. 5th-95th CC
1	Active surveillance	1	0.709	99.3%	0.674-0.807
1	Thyroid lobectomy	2	0.582	0.7%	0.522-0.668
1	Thermal ablation	3	0.549	0.1%	0.485-0.635
2	Adjuvant RAI 100-150 mCi	1	0.686	92.7%	0.651-0.782
2	Low-activity RAI 30-50 mCi	2	0.609	7.3%	0.561-0.704
2	No RAI	3	0.483	0.0%	0.406-0.549

Note. CC, closeness coefficient; cond. top-rank freq., proportion of conditional numerical robustness iterations in which the alternative ranked first; RAI, radioactive iodine. This frequency is conditional on the specified sampling assumptions and is not a probability of treatment success or clinical superiority.

In Showcase 2, adjuvant RAI at 100-150 mCi ranked first with CC 0.686, followed by low-activity RAI at 0.609 and no RAI at 0.483. The guideline-concordant activity had a conditional top-rank frequency of 92.7%; low-activity RAI was first in 7.3% and no RAI was never first in the 30,000 primary iterations. The 5th-95th percentile interval for the 100-150 mCi CC was 0.651-0.782 and the mean sampled top-versus-second margin was 0.089. Thus, under the stated perturbation assumptions, the strong-recommendation showcase did not automatically produce a larger simulated separation than Showcase 1.

Model Ablation, Sensitivity and Safety-Gate Behaviour

Model ablation demonstrated complementary roles for the guideline and contextual components (Table 6). In Showcase 1, the context-only and proposed models selected active surveillance, while the guideline-only model selected lobectomy under the primary scope-

qualified Recommendation 15 proxy. In the corresponding one-way analysis, active surveillance remained first until λ reached approximately 0.680 (Figure 2). This value was assumption-dependent: it shifted to approximately 0.438 under full direct transfer, no rank reversal occurred below $\lambda=1.00$ under parent-node harmonisation and no crossover occurred when Recommendation 15 received no direct parent-node credit. The 0.680 value is therefore a model-specific sensitivity result, not a guideline-derived clinical threshold. All executable gate tests passed. When unreliable follow-up triggered exclusion, lobectomy became first (CC 0.582 versus 0.549 for thermal ablation), consistent with Proposition 3. The four Recommendation 15 scope mappings and the associated crossover thresholds are summarised in Table 7.

Table 6. Comparative behaviour of context-only, proposed and guideline-only models

Showcase	Context-only top	Proposed top	Guideline-only top	Crossover λ	Mean sim. margin
1	Active surveillance	Active surveillance	Thyroid lobectomy	0.680	0.136
2	Low-activity RAI	Adjuvant RAI 100-150	Adjuvant RAI 100-150	0.037	0.089

Note. λ , weight assigned to guideline concordance; sim., conditional numerical robustness simulation; RAI, radioactive iodine. The Showcase 1 crossover uses the primary scope-qualified Recommendation 15 mapping and is not guideline-derived.

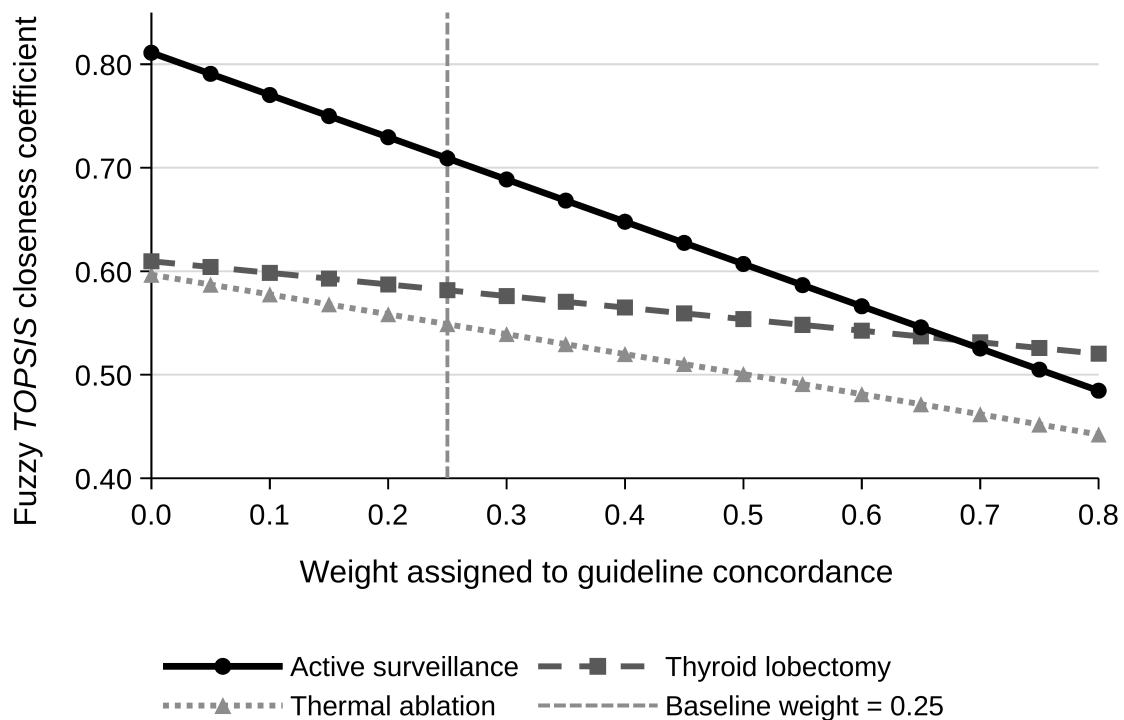


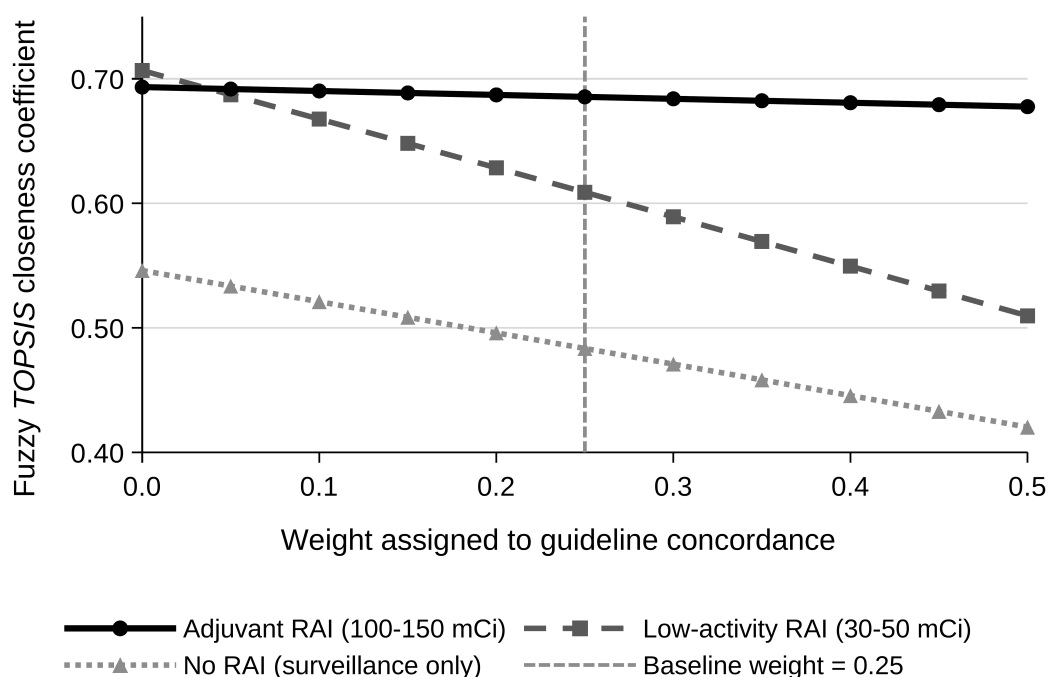
Figure 2. Showcase 1 one-way sensitivity of the fuzzy TOPSIS closeness coefficients to the guideline-concordance weight under the primary scope-qualified Recommendation 15 mapping.

Table 7. Recommendation 15 scope-mapping sensitivity

Scope mapping	Baseline leader	Lobectomy baseline CC	AS-TL crossover λ
Scope-qualified partial proxy (primary)	Active surveillance	0.5817	0.6796
Full direct transfer (stress test)	Active surveillance	0.6226	0.4378
Parent-node harmonised (conditional/low)	Active surveillance	0.5582	No reversal below 1.0
No direct Recommendation 15 credit at parent node	Active surveillance	0.4665	No crossover

Note. CC, closeness coefficient; AS, active surveillance; TL, thyroid lobectomy; λ , weight assigned to guideline concordance.

In Showcase 2, the context-only model slightly favoured 30-50 mCi because of its lower burden and stronger resource fit. Introducing a guideline-concordance weight greater than approximately 0.037 shifted the top rank to 100-150 mCi, after which the ranking was stable through the evaluated one-way range (Figure 3). Across Dirichlet concentrations of 50, 100, 150 and 300, the leading alternative remained unchanged in both showcases: the active-surveillance conditional top-rank frequency ranged from 97.9% to 99.5% and the 100-150 mCi RAI frequency ranged from 90.4% to 93.3%. The Dirichlet-concentration results are summarised in Table 8.

**Figure 3.** Showcase 2 one-way sensitivity of the fuzzy *TOPSIS* closeness coefficients to the guideline-concordance weight for the three radioactive iodine strategies.**Table 8.** Dirichlet-concentration sensitivity of the conditional top-rank frequency

Concentration	Showcase 1 leader	Frequency	Showcase 2 leader	Frequency
50	Active surveillance	0.9793	Adjuvant RAI (100-150 mCi)	0.9037
100	Active surveillance	0.9886	Adjuvant RAI (100-150 mCi)	0.9220
150	Active surveillance	0.9925	Adjuvant RAI (100-150 mCi)	0.9274
300	Active surveillance	0.9953	Adjuvant RAI (100-150 mCi)	0.9333

Note. RAI, radioactive iodine; mCi, millicuries. Frequencies are conditional top-rank frequencies over 30,000 iterations.

DISCUSSION

Principal Findings

Fuzzy-DATA behaved differently at the two guideline states for which it was designed. At the conditional/low-certainty node, contextual advantages, reliable monitoring and a strong informed preference to avoid surgery allowed active surveillance to outrank lobectomy under the primary scope-qualified mapping. At the strong/moderate-certainty node, the recommendation component prevented lower treatment burden from displacing the guideline-concordant high-risk *RAI* activity in the deterministic baseline. The conditional simulations supported the stability of both leading alternatives but did not show a monotonic relationship between recommendation strength and rank separation: active surveillance had a 99.3% conditional top-rank frequency, compared with 92.7% for 100-150 mCi *RAI*. This confirms that robustness reflects the full input matrix and perturbation assumptions, not the recommendation-strength label alone.

The hard-gate stress test was equally important. A purely compensatory model could still select active surveillance by offsetting unreliable follow-up with low treatment burden or strong patient preference. Fuzzy-DATA instead removed the alternative when a necessary condition failed (Proposition 3). The executable gate, exclusion log and passing boundary tests make this noncompensatory step reproducible rather than merely conceptual.

Methodological Contribution

The framework contributes four linked mechanisms. First, it distinguishes recommendation strength from certainty rather than collapsing both into a single ordinal code. Second, it introduces an applicability term, which is necessary because a strong recommendation may be conditional on a prior choice, population, dose or setting. Third, it preserves fuzzy information through the ranking stage and reports assumption-dependent robustness rather than presenting a single deterministic score as definitive. Fourth, it embeds the calculation in a dynamic pathway so that response assessment can alter subsequent decisions. In the illustrative fuzzy *TOPSIS* literature reviewed for this study, formal recommendation metadata, explicit applicability and executable noncompensatory guideline conditions were not identified together in one workflow (Salih et al., 2019; Palczewski and Sařabun, 2019; Alharbi et al., 2024). Table 9 summarises this comparison.

Table 9. Feature comparison with the cited fuzzy *TOPSIS* literature

Source	Recommendation metadata	Patient-level applicability	Hard gates	Uncertainty output	Clinical validation
This study (Fuzzy-DATA)	Strength, certainty and direction	Explicit applicability term	Executable, unit-tested	Conditional rank frequencies and one-way thresholds	Not claimed; synthetic showcases
Chen (2000)	None	None	None	Deterministic closeness only	Not applicable
Salih et al. (2019) survey corpus	Not reported	Rare	Not reported	Mostly deterministic	Rarely reported
Palczewski and Sařabun (2019) survey corpus	Not reported	Rare	Not reported	Mostly deterministic	Rarely reported
Alharbi et al. (2024)	None	None	None	Deterministic weights and ranks	None reported

Note. Entries for the two surveys describe the application corpora those surveys review, as reported by the survey authors. The comparison covers only the works cited in this article and is illustrative rather than systematic.

The approach also avoids treating recommendation strength as an outcome effect size. A strong recommendation can arise from multiple domains, including benefit-harm balance, certainty, values, feasibility, acceptability, resources and equity (Andrews et al., 2013; Guyatt et al., 2025; Ringel et al., 2025). Accordingly, the guideline-concordance criterion represents normative support and applicability, whereas expected disease control and harm avoidance remain separate patient-context criteria. This separation helps make the reasoning auditable.

Relevance to Practitioners, Clinical Governance and Potential Implementation

For practising clinicians and multidisciplinary teams, the framework's potential value is to make the reasoning behind a preference-sensitive recommendation explicit and auditable, subject to prior calibration and validation. Fuzzy-DATA displays which guideline statement applies, its strength and certainty, the scope and applicability judgement and the contextual criteria driving the current ranking. Under a validated implementation, a small margin or unstable conditional rank frequency could prompt closer attention to patient values and local feasibility, while a stable result could support a documented default. At the present stage, these outputs are methodological demonstrations and must not be entered into clinical notes or used for patient care as validated estimates.

A future practical implementation could be integrated into a multidisciplinary tumour board or shared-decision interface only after calibration and prospective evaluation. The system should display the underlying recommendation statement, strength, certainty, applicability judgement, criterion ratings, top-ranked option, decision margin, conditional rank frequency and sensitivity assumptions. Users should be able to revise preferences or local feasibility assumptions and immediately see whether the ranking changes. The framework should never issue autonomous treatment orders and it should not conceal disagreement behind a single score. Clinical experts must verify eligibility gates and patients must receive conventional counselling about benefits, harms, uncertainty and alternatives.

Local adaptation is also necessary. Availability of experienced ultrasonography, ablation, high-volume surgery, nuclear medicine, isolation facilities and follow-up differs across settings. The applicability component permits these factors to be represented, but local calibration should occur through a transparent multidisciplinary process rather than by undocumented alteration of guideline strength.

Limitations

This study has important limitations. First, both showcases and all performance ratings were simulated; the results demonstrate internal computational behaviour, not patient benefit, diagnostic accuracy or clinical validity. Second, *TFN* boundaries and baseline weights were prespecified rather than elicited from patients, clinicians or guideline developers. Third, multiplication of strength and certainty is a conservative modelling choice and may partly re-express information already incorporated into recommendation strength. Fourth, Recommendation 15 was represented by a scope-qualified partial proxy in a flat parent-node comparison; the crossover changed materially under alternative scope mappings. Fifth, triangular *TFN* sampling and Dirichlet weight perturbation were conditional numerical stress-test assumptions, not empirical probability models and the resulting frequencies must not be

interpreted clinically. Sixth, only two ATA decision nodes were tested. Seventh, *TOPSIS* assumes that retained criteria can be compared through a compensatory distance structure after gating; interactions among criteria were not modelled. Eighth, the resource/equity criterion was illustrative and did not constitute an economic evaluation. Finally, a ranking may create false precision if users do not attend to provenance, decision margins, scope assumptions and sensitivity outputs. The provenance and calibration status of every numerical input are summarised in Table 10. In addition, one scalar score was drawn independently for each alternative-criterion *TFN* from its fixed triangular distribution; dependence among criterion-level score draws and between those draws and the Dirichlet weight vector was not modelled. The literature comparison in Table 9 was confined to the cited examples rather than a systematic search.

Table 10. Input provenance and calibration status

Input	Value/source	Source type	Rationale	Validation status
Decision nodes; recommendation direction, strength and certainty	2025 ATA guideline	Guideline-derived	Used as reported; no patient-level data	Source-checked
Active-surveillance eligibility gate predicates	Author operationalisation informed by ATA Recommendation 14	Guideline-informed author operationalisation	Anxiety omitted; metastases merged without the new biopsy-proven qualifier; extension and posterior growth narrowed; aggressive features added (see the Safety Gates and Dynamic Updating subsection)	Requires clinical review; unit-tested in the supplementary code
Typical 100-150 mCi activity for high-risk postoperative RAI	2025 ATA guideline	Guideline-derived	Used to define the illustrative high-risk alternative	Source-checked
Strength and certainty TFNs	Table 1 numerical anchors	Author-assigned modelling parameter	Transparent ordinal-to-fuzzy encoding	Requires expert calibration
Applicability and performance-rating TFNs	Table 1 numerical anchors	Author-assigned modelling parameter	Represents exact, partial, poor-match and linguistic ratings	Requires expert calibration
Seven baseline criterion weights	0.25, 0.20, 0.15, 0.15, 0.10, 0.10, 0.05	Author-assigned synthetic input	Balanced demonstration; varied in sensitivity analyses	Not clinically validated
Showcase profiles and contextual ratings	Tables 3-4 and full fuzzy matrices	Author-constructed synthetic input	Scenario logic selected before ranking	Not patient data
Recommendation 15 partial applicability in the flat parent-node showcase	Scope-qualified partial proxy	Author-assigned scope mapping	Stress-tested against three alternative mappings	Interpret with scope sensitivity
Complement for a recommendation against an alternative	1 minus the strength-certainty product, then applicability	Author-specified modelling rule	Creates directional guideline-concordance input	Requires external calibration
Fuzzy TOPSIS distance and closeness	Summed criterion-level Chen vertex distances	Method-derived calculation	Same aggregation used in deterministic and sampled analyses	Computationally reproduced

Triangular sampling of TFNs	Bounded draws with modal emphasis	Conditional robustness assumption	Numerical stress test; not an empirical probability model	Interpret as conditional
Dirichlet weight concentration	50, 100, 150 and 300; primary value 150	Conditional robustness assumption	Controls dispersion around baseline mean weights	Sensitivity-tested
Ranks, closeness coefficients, margins and rank frequencies	Generated by the supplied script	Calculated output	Properties of the specified synthetic inputs	Reproducible; not clinical effect estimates

Note. ATA, American Thyroid Association; RAI, radioactive iodine; TFN, triangular fuzzy number; TOPSIS, Technique for Order Preference by Similarity to Ideal Solution; mCi, millicuries.

Future Validation

The next development stage should use a formal expert and patient calibration process. Multidisciplinary panels could define recommendation-to-*TFN* mappings, applicability anchors, criterion definitions and gate rules. Patient preference weights should be elicited using validated decision-aid methods. Retrospective testing should assess concordance with documented tumour board decisions and identify systematic disagreement. Prospective studies should then examine usability, decisional conflict, explanation quality, inter-rater reliability, subgroup fairness and whether rankings remain stable when real clinical uncertainty is entered. Comparison with guideline-only rules, unaided clinician judgement and alternative multi-criteria decision analysis (*MCD*A) methods would clarify incremental value.

CONCLUSION

Fuzzy-DATA provides a transparent way to incorporate recommendation strength, evidence certainty, applicability, patient context and assumption-dependent uncertainty into a single decision-support workflow without allowing preference trade-offs to override explicit eligibility constraints. In two ATA-based showcases, the model produced internally coherent baseline rankings, executable safety-gate behaviour and leading alternatives that remained stable across the tested weight-dispersion settings. Recommendation 15 scope tests also showed that a reported crossover is a modelling result rather than a guideline-derived threshold. The framework is computationally feasible and reproducible, but its numerical scales and clinical performance require multidisciplinary calibration and real-world validation before any use in patient care.

REFERENCES

- Alharbi, A., Alosaimi, W., Alyami, H., Alouffi, B., Almulihi, A., Nadeem, M., Sayeed, M. A., & Khan, R. A. (2024). Selection of data analytic techniques by using fuzzy AHP TOPSIS from a healthcare perspective. *BMC Medical Informatics and Decision Making*, 24(1), 240. <https://doi.org/10.1186/s12911-024-02651-8>
- Andrews, J. C., Schünemann, H. J., Oxman, A. D., Pottie, K., Meerpohl, J. J., Coello, P. A., Rind, D., Montori, V. M., Brito, J. P., Norris, S., Elbarbary, M., Post, P., Nasser, M., Shukla, V., Jaeschke, R., Brozek, J., Djulbegovic, B., & Guyatt, G. (2013). GRADE guidelines: 15. Going from evidence to recommendation-determinants of a recommendation's direction and strength. *Journal of Clinical Epidemiology*, 66(7), 726-735. <https://doi.org/10.1016/j.jclinepi.2013.02.003>
- Chen, C. T. (2000). Extensions of the TOPSIS for group decision-making under fuzzy environment. *Fuzzy Sets and Systems*, 114(1), 1-9. [https://doi.org/10.1016/S0165-0114\(97\)00377-1](https://doi.org/10.1016/S0165-0114(97)00377-1)
- Guyatt, G., Vandvik, P. O., Iorio, A., Agarwal, A., Yao, L., Eachempati, P., Zeng, L., Chu, D. K., D'Souza, R., Agoritsas, T., Murad, M. H., Schandelmaier, S., Rylance, J., Djulbegovic, B., Montori, V. M., Hultcrantz, M., &

- Brignardello-Petersen, R. (2025). Core GRADE 7: Principles for moving from evidence to recommendations and decisions. *BMJ*, 389, e083867. <https://doi.org/10.1136/bmj-2024-083867>
- Hwang, C. L., & Yoon, K. (1981). Multiple attribute decision making: Methods and applications: A state-of-the-art survey. Springer. <https://doi.org/10.1007/978-3-642-48318-9>
- Palczewski, K., & Sařabun, W. (2019). The fuzzy TOPSIS applications in the last decade. *Procedia Computer Science*, 159, 2294-2303. <https://doi.org/10.1016/j.procs.2019.09.404>
- Ringel, M. D., Sosa, J. A., Baloch, Z., Bischoff, L., Bloom, G., Brent, G. A., Brock, P. L., Chou, R., Flavell, R. R., Goldner, W., Grubbs, E. G., Haymart, M., Larson, S. M., Leung, A. M., Osborne, J. R., Ridge, J. A., Robinson, B., Steward, D. L., Tufano, R. P., & Wirth, L. J. (2025). 2025 American Thyroid Association management guidelines for adult patients with differentiated thyroid cancer. *Thyroid*, 35(8), 841-985. <https://doi.org/10.1177/10507256251363120>
- Salih, M. M., Zaidan, B. B., Zaidan, A. A., & Ahmed, M. A. (2019). Survey on fuzzy TOPSIS state-of-the-art between 2007 and 2017. *Computers & Operations Research*, 104, 207-227. <https://doi.org/10.1016/j.cor.2018.12.019>
- Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8(3), 338-353. [https://doi.org/10.1016/S0019-9958\(65\)90241-X](https://doi.org/10.1016/S0019-9958(65)90241-X)

Author's Contributions

This study was conducted by a single author. The author was responsible for the conception and design of the research, conducting the literature review, data collection and analysis processes, selection and implementation of the methodology, interpretation of the results, preparation of figures and tables, drafting and revising the manuscript and approving the final version of the manuscript.

Statement of Conflicts of Interest

The author declared that there is no conflict of interest.

Statement of Research and Publication Ethics

Ethics committee approval was not required because the study involved no humans or animals and used only fully illustrative scenarios. It contains no human participant data. The framework and showcase outputs are methodological illustrations and must not be used as patient-specific treatment recommendations.

Funding

This research received no specific grant, financial support or funding from any public, commercial or not-for-profit organization

Data Availability

All reported results are reproducible from the Python script provided in *Supplementary Material 1*, which is uploaded as a separate file with this article. The complete fuzzy input matrices, full baseline distances, model ablation details and conditional robustness summaries are included in *Supplementary Material 1*, while the Dirichlet-concentration, Recommendation 15 scope-mapping and input-provenance tables appear in the manuscript as Tables 7, 8 and 10.



Acknowledgements

The author sincerely thanks the anonymous reviewers for their careful reading and constructive comments.

Use of artificial intelligence (AI)

Google Gemini Pro and Grammarly (v. 26.7) assisted with translation and grammar editing. The author independently reviewed and validated all editing and accepted full responsibility for the final work. No AI tools were used to generate the scientific content.

Clinical disclaimer

The framework and its outputs must not be used as patient-specific decisions without local calibration and clinician validation.