Diagnostic Accuracy of Laryngeal Ultrasound for Evaluating Vocal Fold Movement Impairment in Children

A Prospective Validation Study

報告者：吳凱勛
指導老師：吳伯軒醫師

Journal of Pediatric Surgery 59 (2024) 109–116

Contents lists available at ScienceDirect

Journal of Pediatric Surgery

journal homepage: www.sciencedirect.com/journal/journal-of-pediatric-surgery

Diagnostic Accuracy of Laryngeal Ultrasound for Evaluating Vocal Fold Movement Impairment in Children

Shawn Izadi [a,1], Benjamin Zendejas [a,1], Jay Meisner [a], Ali Kamran [a], Somala Mohammed [a], Farokh Demehri [a], Steven Staffa [a], David Zurakowski [a], Anne Hseu [b], Michael Cunningham [b], Sukgi Choi [b,2], Carol Barnewolt [c,*,2]

[a] Department of Surgery, Boston Children's Hospital, Boston, MA, USA
[b] Department of Otolaryngology and Communication Enhancement, Boston Children's Hospital, Boston, MA, USA
[c] Department of Radiology, Boston Children's Hospital, Boston, MA, USA

Clinical Background: Vocal Fold Movement Impairment (VFMI)

Definition: Decreased or absent mobility of true vocal folds.

Airway

Stridor, respiratory distress.

Feeding

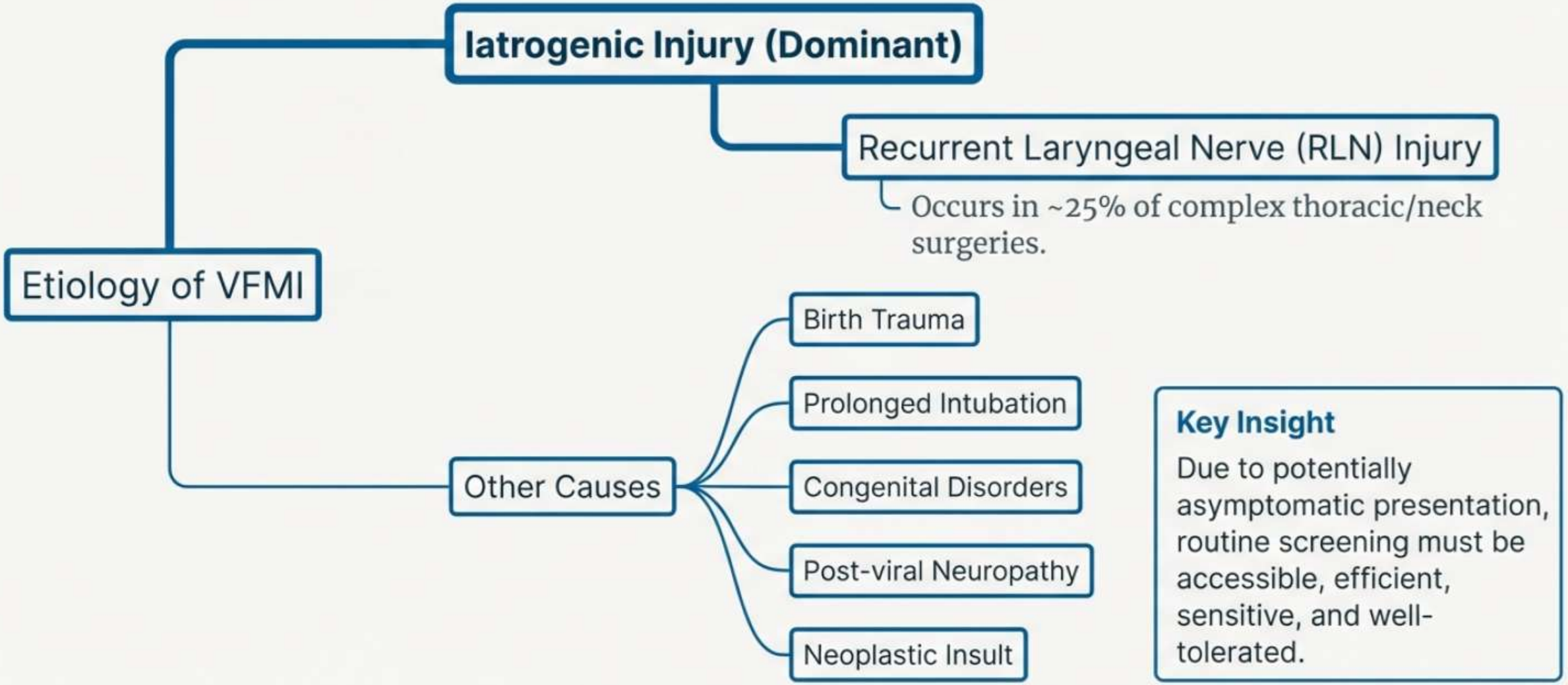
Aspiration, recurrent pneumonia.

Quality of Life

Dysphonia.

Objective: To present Laryngeal Ultrasound (LUS) as a high-accuracy, non-invasive screening alternative to Flexible Laryngoscopy (FL) for **post-operative** pediatric patients.

Etiology: Why Screening is Imperative



Key Insight

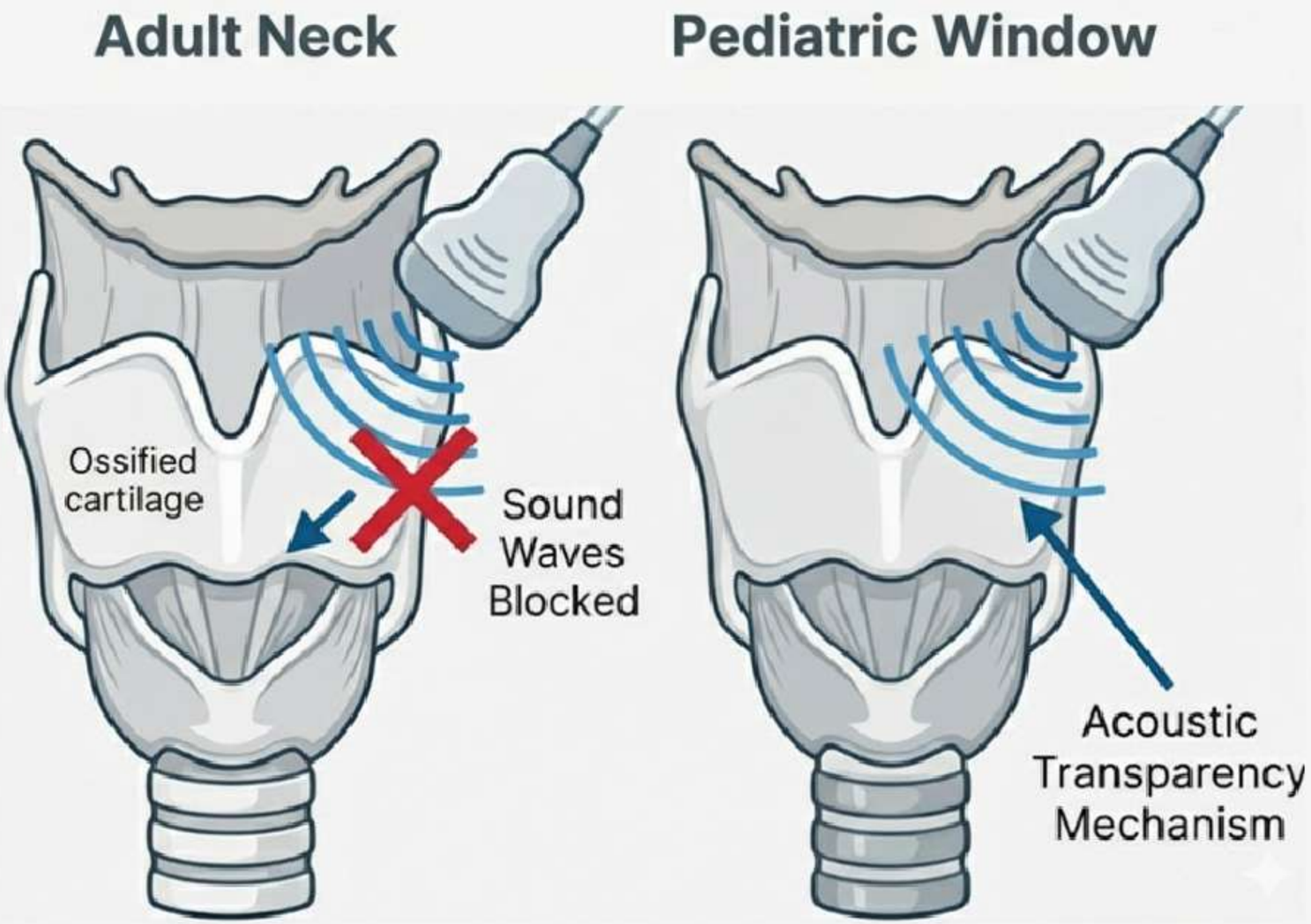
Due to potentially asymptomatic presentation, routine screening must be accessible, efficient, sensitive, and well-tolerated.

Limitations of the Current Gold Standard (Flexible Laryngoscopy)

Category	Specific Limitation
Patient Factors	Poorly tolerated in neonates/infants; risks of laryngospasm.
Clinical Risks	Hemodynamic side effects (Vagal Reflex), oxygen desaturation.
Operational	Requires specialized equipment and high provider expertise; Aerosol Generating Procedure risks.

Synthesis: While FL is the standard, its invasive nature makes it suboptimal for frequent screening in vulnerable pediatric populations.

The Solution: Laryngeal Ultrasound (LUS)



Why it works in children:

- Uncalcified laryngeal cartilages allow sound wave penetration.
- Visualizes endolaryngeal structures clearly.

Core Benefits:

- Bedside capability
- Zero radiation
- No sedation required
- High parent/patient acceptance

Study Design & Methodology Snapshot

Study Type

Prospective, single-center, single-blinded cohort study.

Timeline

2021 – 2023

Population

Children (<18 years) post-op from high-risk surgeries (Esophageal atresia repair, PDA ligation, Tracheopexy, Great Vessel surgeries).

Blinding

LUS raters and FL performers were blinded to each other’s results to prevent bias.

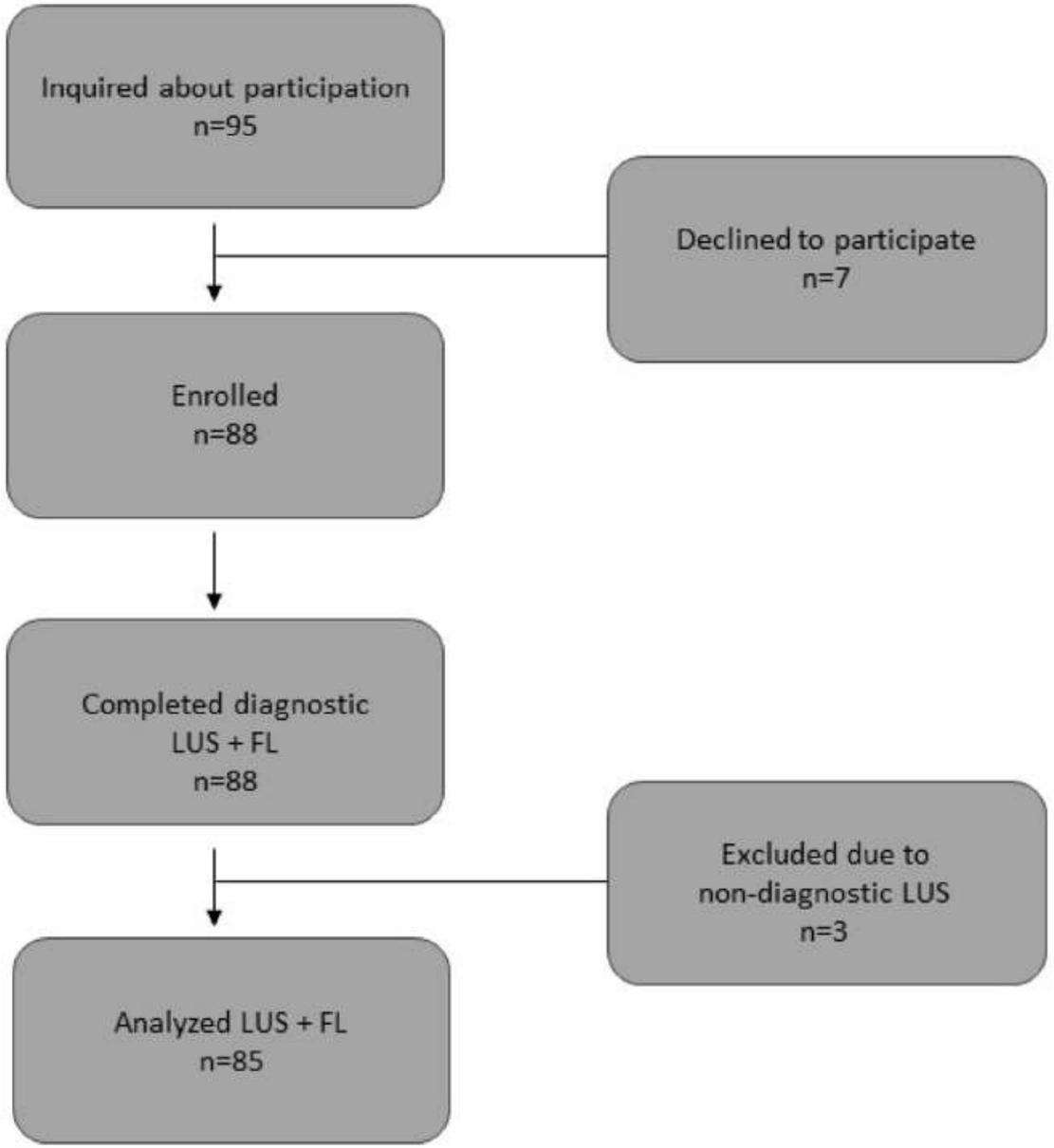
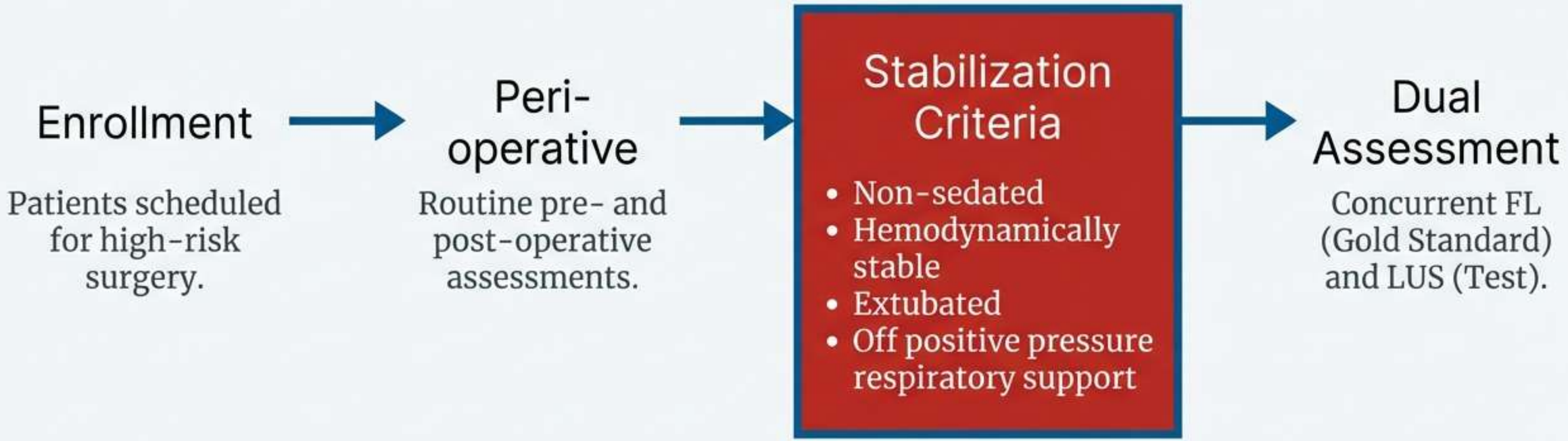


Fig. 2. Consort diagram.

Table 1
Demographics.

Sub-groups	Number of cases (%)
Age	
0–1 month	4 (4.7%)
2–12 months	42 (49.4%)
13–24 months	14 (16.5%)
25–144 months	20 (23.5%)
>144 months	5 (5.9%)
Weight quartile	
2.4 to <5.4 kg	21 (24.7%)
5.4 to <7.5 kg	22 (25.9%)
7.5 to <13.4 kg	20 (23.5%)
13.4–61.4 kg	22 (25.9%)
Sex	
Female	32 (37.7%)
Male	53 (62.3%)
Diagnosis category	
EA related	60 (70.5%)
Non-EA related	10 (11.7%)
Vascular ring/compression related	12 (14.1%)
Other	3 (3.5%)
Procedure category	n = 83
Esophageal repair only	41 (49.4%)
Airway repair only	8 (9.6%)
Vascular ring/compression repair	12 (14.5%)
Esophageal & airway repair	22 (26.5%)
Surgical approach	n = 79
Neck only	3 (3.8%)
Chest only	44 (55.7%)
Chest + neck	20 (25.3%)
Sternotomy only	7 (8.9%)
Sternotomy + neck	5 (6.3%)
Time between assessments	
<14 days	63 (74.1%)
≥14 days ≤31 days	17 (25.9%)
>31 days	5 (5.8%)

Integrated Evaluation Protocol



Quality Control: All FL procedures performed by fellowship-trained pediatric otorhinolaryngologists; exams video-recorded.

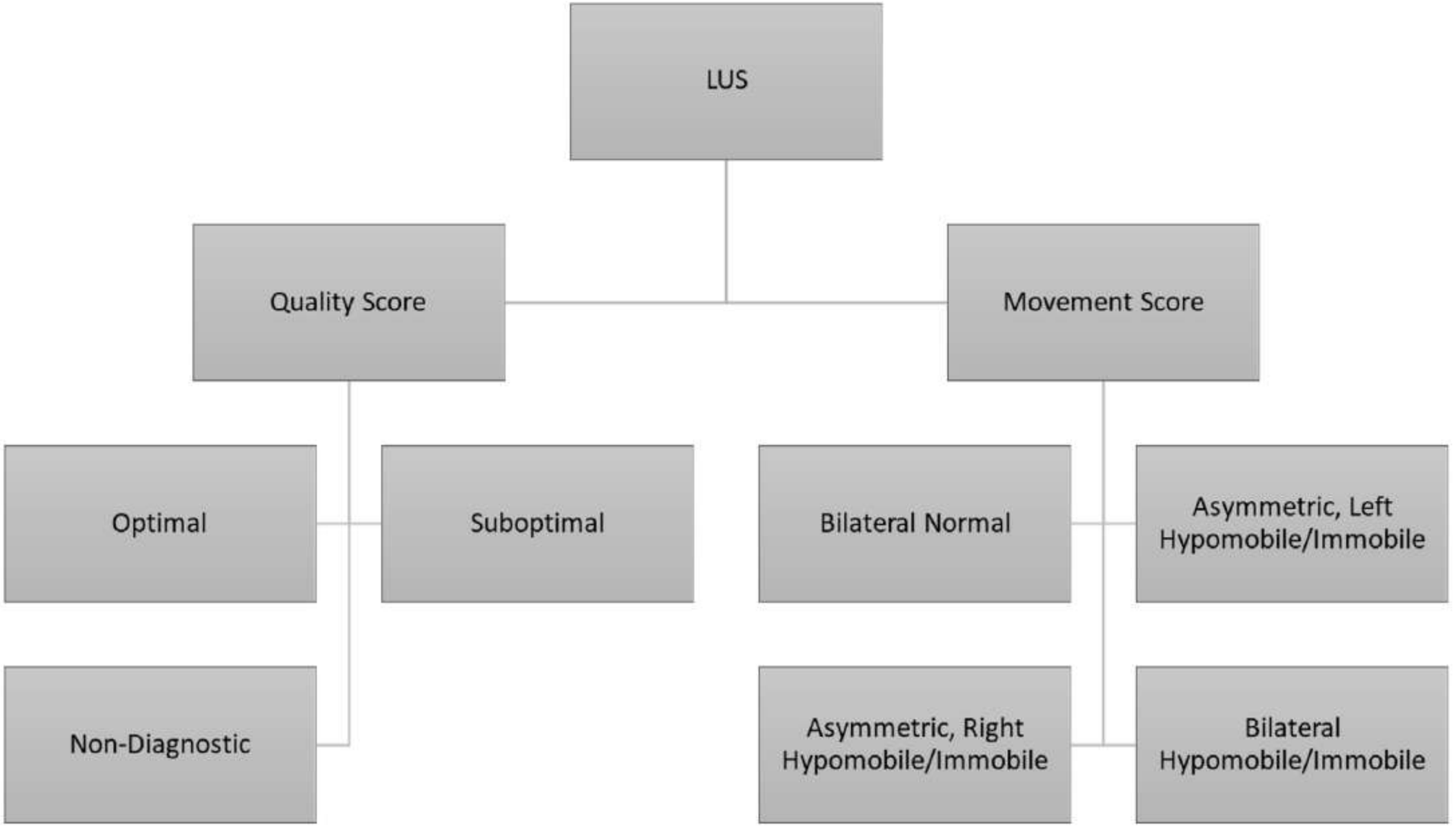


Fig. 1. Laryngeal ultrasound (LUS) assessment protocol.

Quality Scoring System & Clinical Viability

Category 1: Optimal

All landmarks and motion clearly visible.

↓

Category 2: Suboptimal

Motion assessable despite slight blurring or presence of surgical dressings.

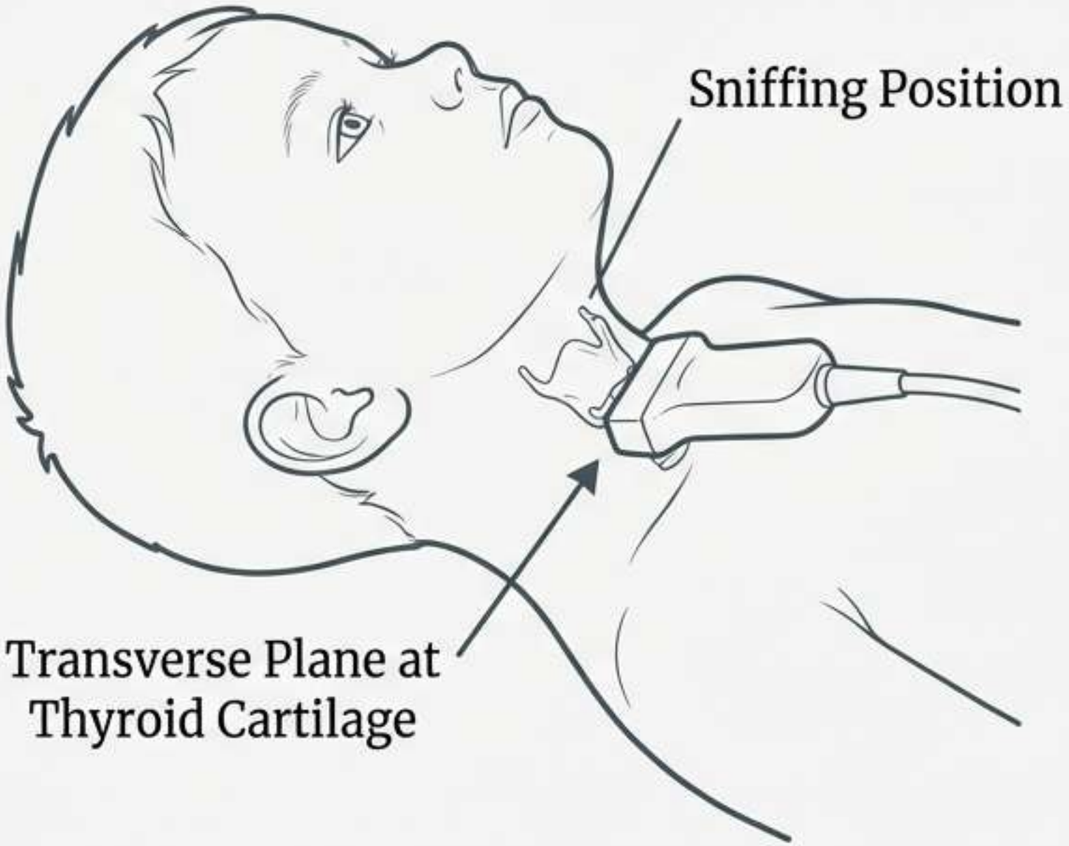
↓

Category 3: Non-diagnostic

Landmarks obscured (e.g., by air from Tracheostomy).

Study Finding: Suboptimal quality images did **NOT** significantly impact diagnostic accuracy, confirming robust clinical utility.

Technical Protocol: Equipment & Positioning



Technical Specifications:

- Equipment: High-frequency linear transducer (ML6-15 MHz).
- Position: Supine with slight neck extension.
- Scanning Plane: Transverse plane.

Data Capture Method:

- 5–15 cine clips per patient.
- Clip Duration: 3–5 seconds.
- Triggers: ‘ee’, ‘ah’, ‘Crying’, or ‘Inspiration’.

Imaging Comparison: LUS vs. CT

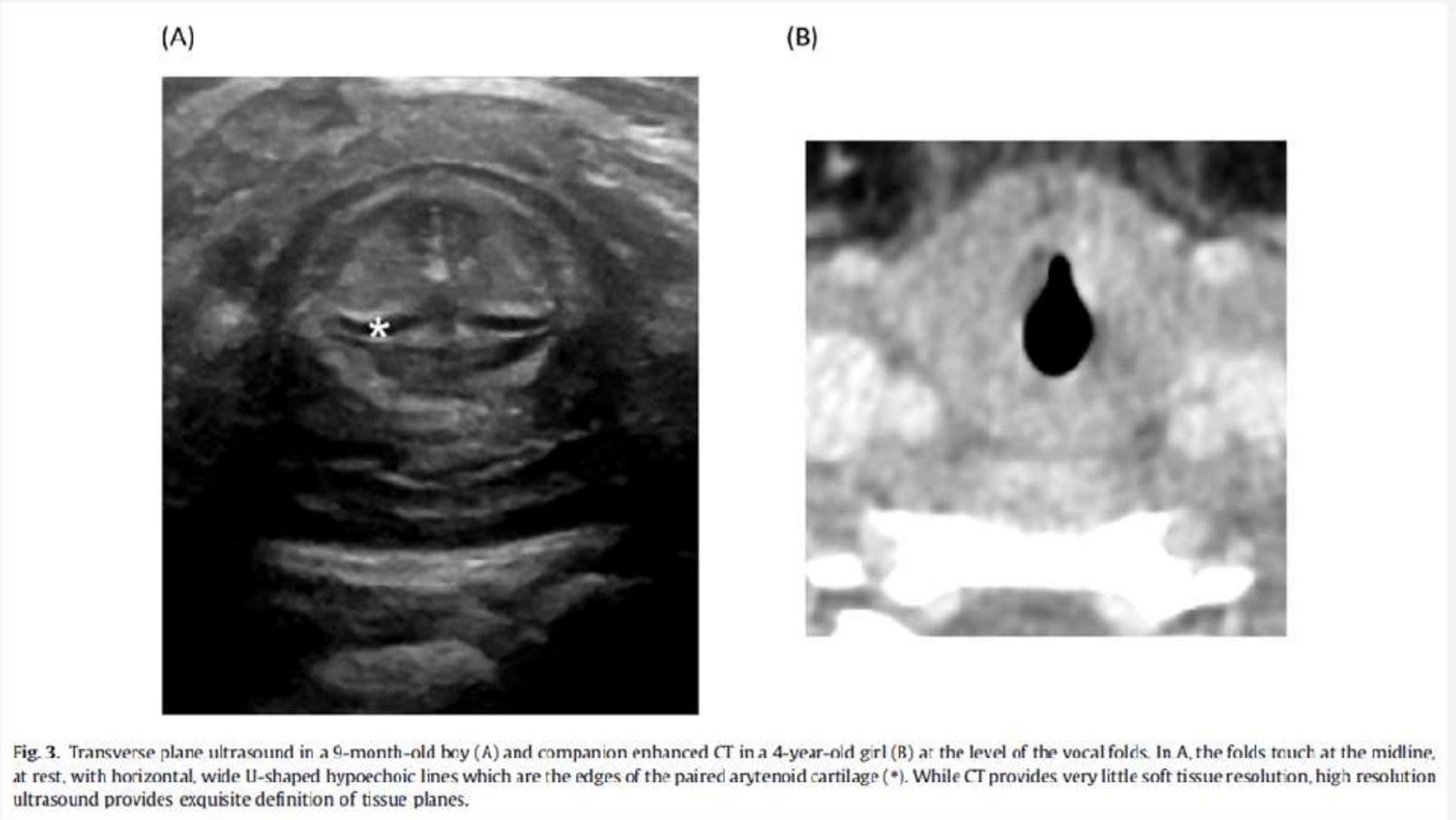
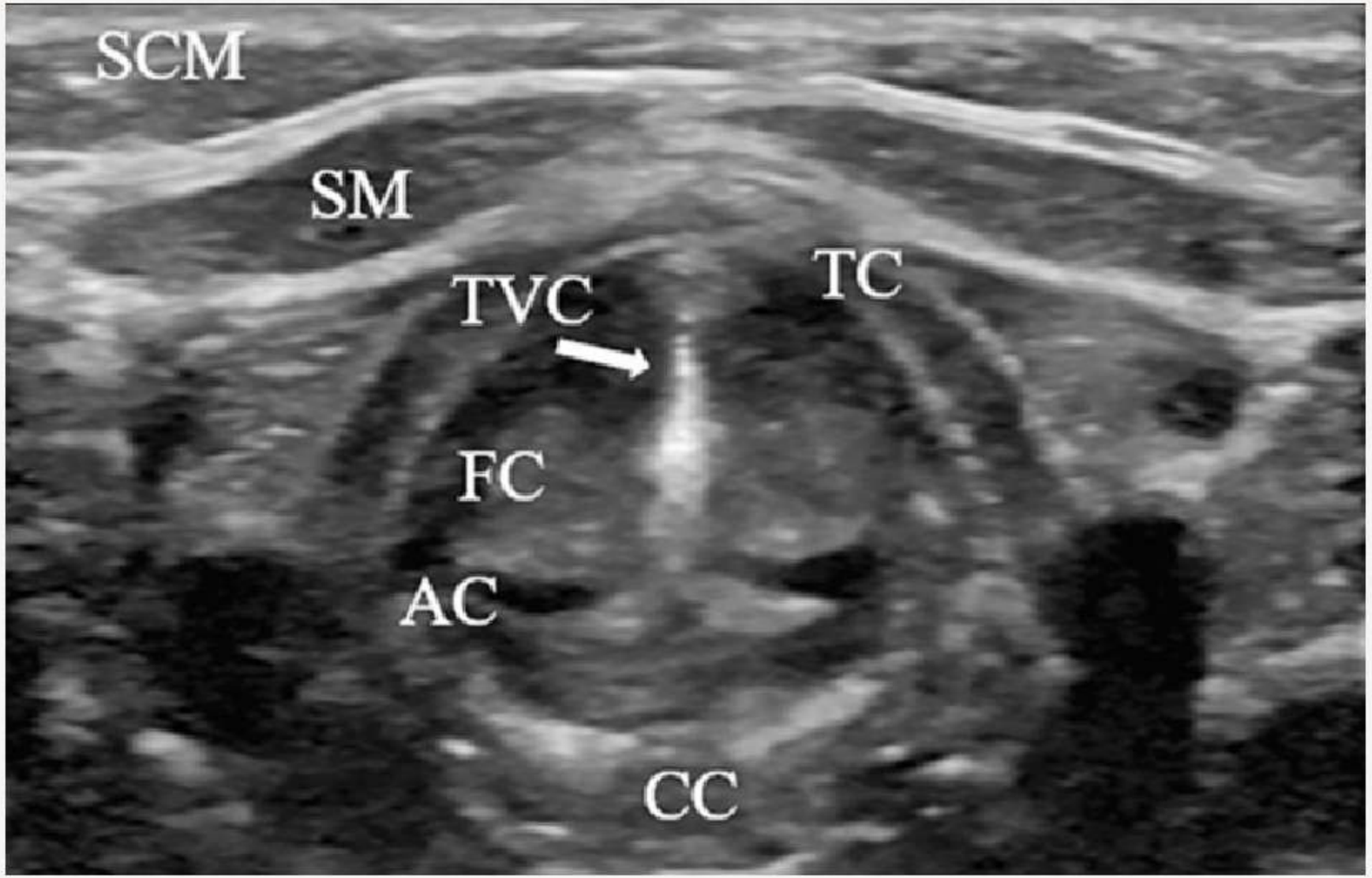


Fig. 3. Transverse plane ultrasound in a 9-month-old boy (A) and companion enhanced CT in a 4-year-old girl (B) at the level of the vocal folds. In A, the folds touch at the midline, at rest, with horizontal, wide U-shaped hypoechoic lines which are the edges of the paired arytenoid cartilage (*). While CT provides very little soft tissue resolution, high resolution ultrasound provides exquisite definition of tissue planes.

Key Takeaway: LUS clearly defines the arytenoid cartilages and vocal fold planes at the midline, whereas CT often lacks this specific soft tissue definition.

Ultrasound Anatomy: Identifying Landmarks



https://link.springer.com/article/10.1007/s00247-021-05235-0

Normal Motion Patterns (Dynamic Assessment)

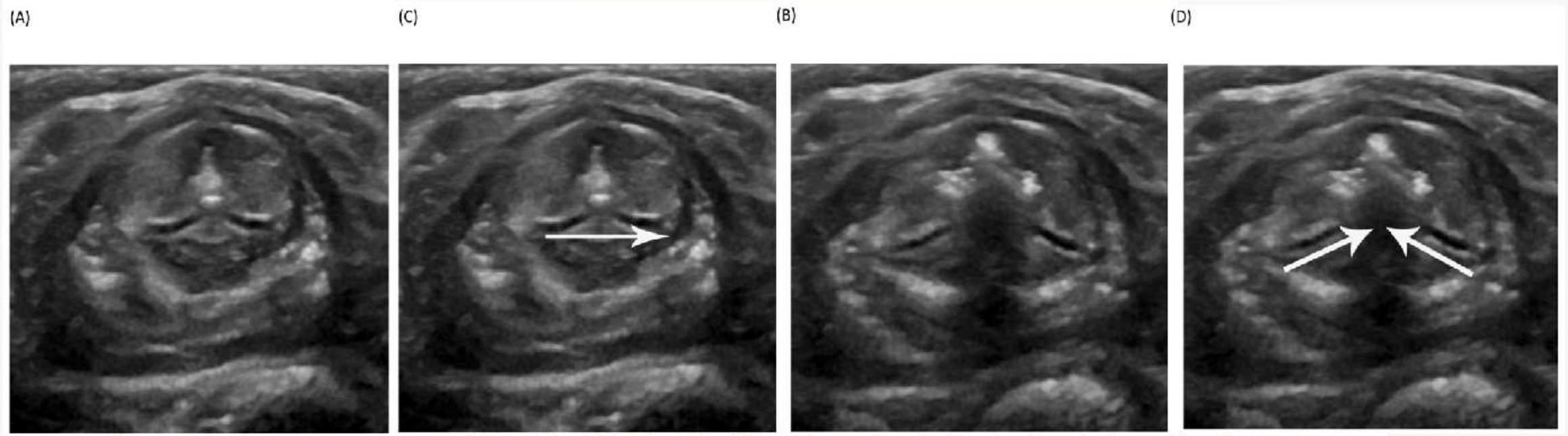


Fig. 4. Two still frames taken from a transverse laryngeal ultrasound cine clip in a 7-month-old boy with normal vocal fold motion. The paired arytenoid cartilage assumes a near horizontal plane at rest (A and C) but pulls to a symmetrical sloping plane with motion (B and D).

Core Diagnostic Indicators on LUS

1. Core Diagnostic Indicators

Symmetrical Visualization

Continuous assessment of the arytenoid cartilages to ensure bilateral symmetry during the motion cycle.

Vocal Fold Separation

Monitoring for the presence or absence of midline separation to confirm normal opening and closing of the glottis.

2. Rare Bilateral Abnormalities

Neurological vs. Random Motion

Differentiating actual neurological function from ineffective, random motion is critical in clinical diagnosis.

Two Key Focus Areas

Emphasis on identifying midline separation and evaluating the change in the angle of the arytenoid cartilage.

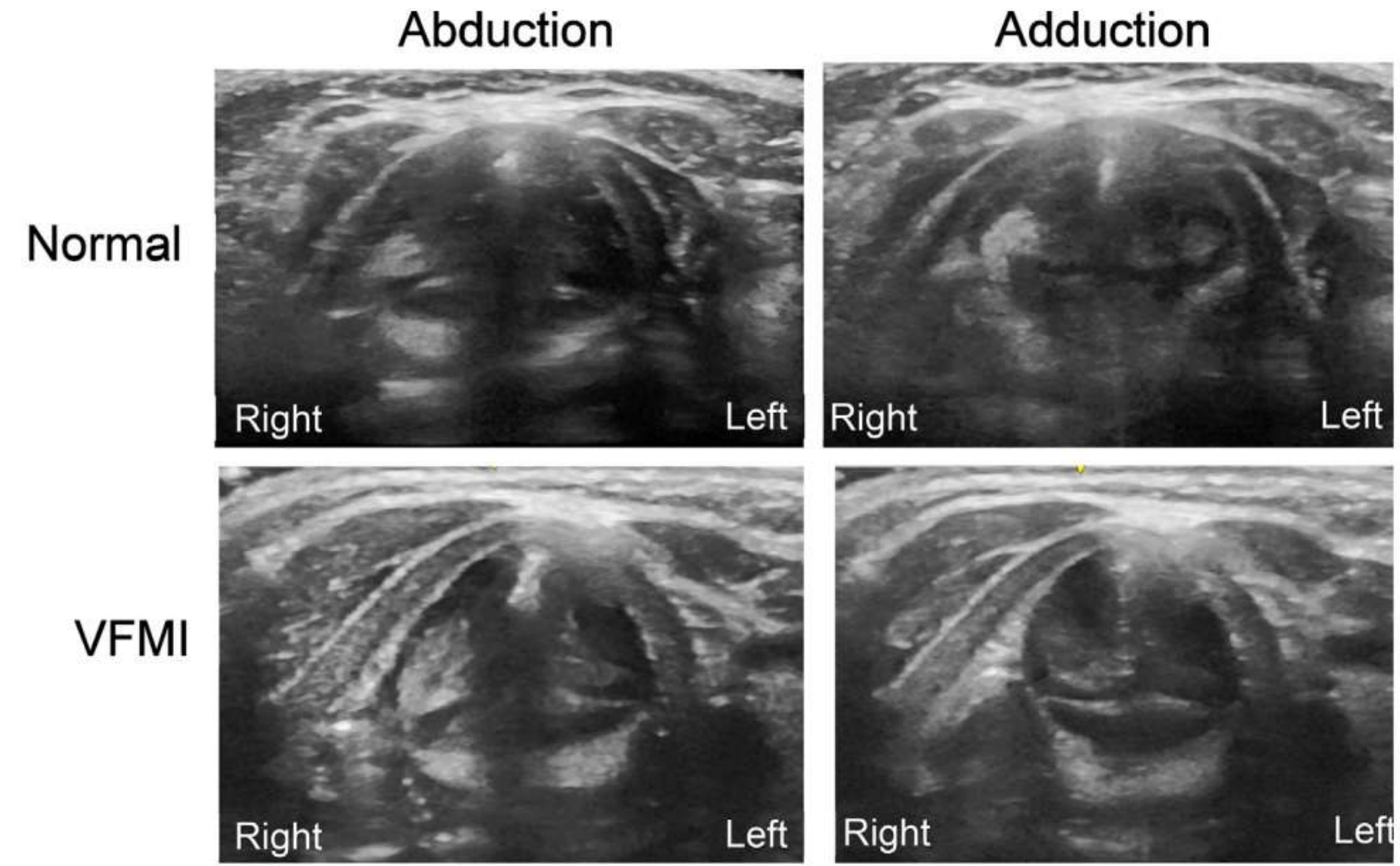


Fig. 1. Laryngeal ultrasound.

Primary Statistical Outcomes

High Sensitivity & Specificity Validated Against Fiberoptic Laryngoscopy (FL)

Study Overview

- Total Paired Evaluations: n = 85
- VFMI Prevalence: 27.1%
- Reference Standard: Fiberoptic Laryngoscopy
- Median Interval: 7 days (IQR: 2–13)

95.7%
Sensitivity

100%
Specificity

100%
Positive Predictive Value (PPV)

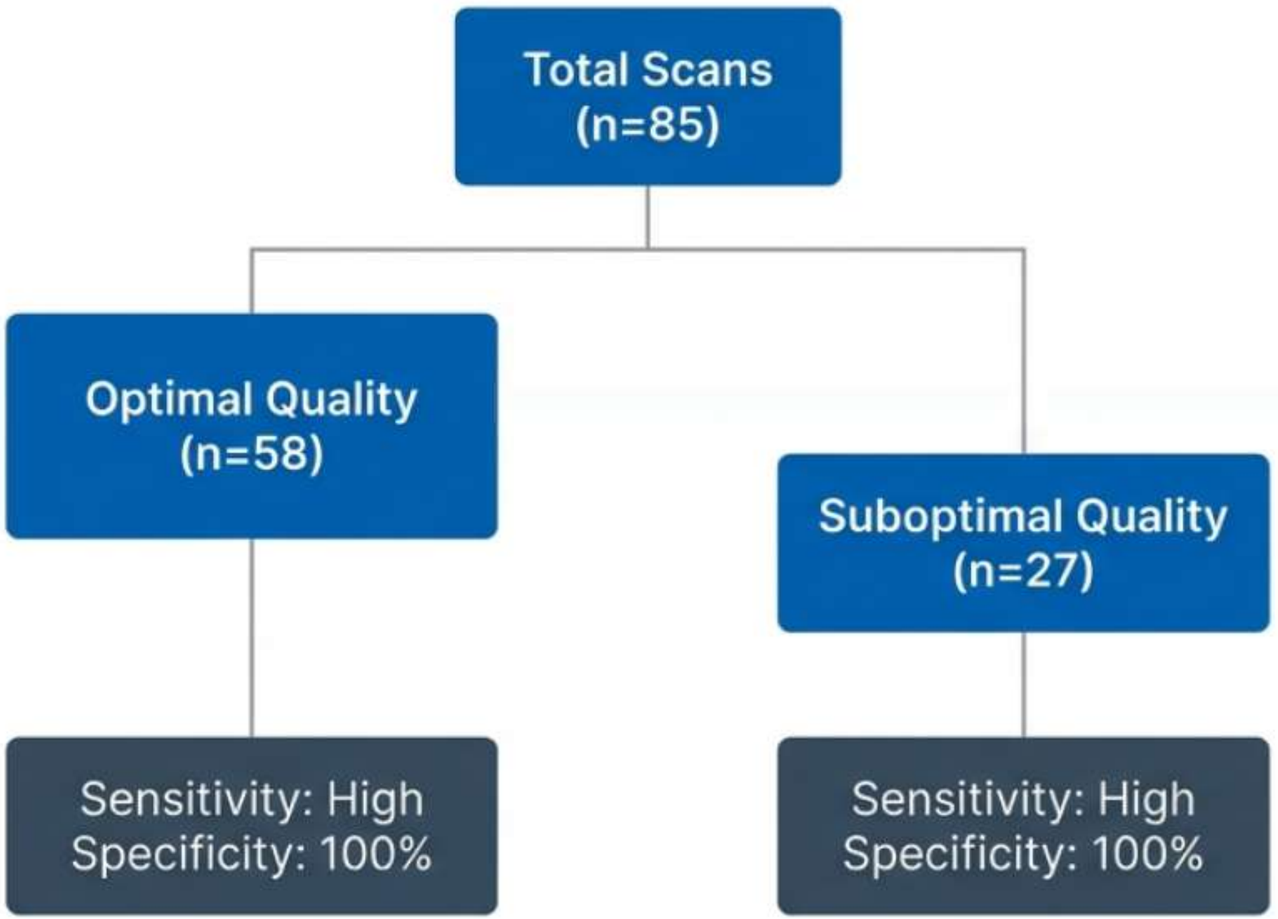
98.4%
Negative Predictive Value (NPV)

98.8%
Overall Accuracy

Statistical Validation: *Kappa = 0.97 (95% CI: 0.91–1.00), P < 0.001. 100% Specificity indicates zero False Positives.*

Clinical Editorial

Quality Stratification & False Negative Analysis



Analysis of Single False Negative

- **Incident:** 1 case
- **Setting:** Optimal Quality Group
- **LUS Finding:** Normal
- **FL Finding:** Unilateral Hypomobility
- **Conclusion:** Interpretive limitation, not quality-dependent.

Comparison of Sensitivity between quality groups showed no statistically significant difference ($P < 0.99$).

Subgroup Analysis: Validated Across Demographics

Consistent diagnostic accuracy regardless of patient age, weight, or surgical approach.

Table 3
Flexible laryngoscopy vs. laryngeal ultrasound subgroup analysis.

Sub-groups	Overall agreement	Kappa	Sensitivity	Specificity	PPV	NPV
Age						
0–1 month	100% (4/4)	[a]	[a]	[a]	[a]	[a]
2–12 months	97.6% (41/42)	0.92 (P < 0.001)	87.5% (7/8)	100% (34/34)	100%	97.2%
13–24 months	100% (14/14)	1.0 (P < 0.001)	100% (7/7)	100% (7/7)	100%	100%
25–144 months	100% (20/20)	1.0 (P < 0.001)	100% (7/7)	100% (13/13)	100%	100%
>144 months	100% (5/5)	1.0 (P = 0.013)	100% (1/1)	100% (4/4)	100%	100%
Weight quartile						
2.4 to <5.4 kg	100% (21/21)	1.0 (P < 0.001)	100% (3/3)	100% (18/18)	100%	100%
5.4 to <7.5 kg	95.5% (21/22)	0.89 (P < 0.001)	85.7% (6/7)	100% (15/15)	100%	93.8%
75 to <13.4 kg	100% (20/20)	1.0 (P < 0.001)	100% (6/6)	100% (14/14)	100%	100%
13.4–61.4 kg	100% (22/22)	1.0 (P < 0.001)	100% (7/7)	100% (15/15)	100%	100%
Sex						
Female	100% (32/32)	1.0 (P < 0.001)	100% (6/6)	100% (26/26)	100%	100%
Male	98.1% (52/53)	0.96 (P < 0.001)	94.1% (16/17)	100% (36/36)	100%	97.3%
Diagnosis category						
EA related	98.3% (59/60)	0.95 (P < 0.001)	91.7% (11/12)	100% (48/48)	100%	97.3%
Non-EA related	100% (10/10)	1.0 (P < 0.001)	100% (2/2)	100% (8/8)	100%	100%
Vascular ring/compression related	100% (12/12)	1.0 (P < 0.001)	100% (9/9)	100% (3/3)	100%	100%
Other	100% (3/3)	[a]	[a]	[a]	[a]	[a]
Procedure category						
Esophageal repair only	100% (41/41)	1.0 (P < 0.001)	100% (5/5)	100% (36/36)	100%	100%
Airway repair only	100% (8/8)	1.0 (P = 0.002)	100% (2/2)	100% (6/6)	100%	100%
Vascular ring/compression repair	100% (12/12)	1.0 (P < 0.001)	100% (9/9)	100% (3/3)	100%	100%
Esophageal & airway repair	95.5% (21/22)	0.89 (P < 0.001)	85.7% (6/7)	100% (15/15)	100%	93.8%
Surgical approach						
Neck only	100% (3/3)	1.0 (P = 0.042)	100% (1/1)	100% (2/2)	100%	100%
Chest only	100% (44/44)	1.0 (P < 0.001)	100% (7/7)	100% (37/37)	100%	100%
Chest + neck	95% (19/20)	0.89 (P < 0.001)	87.5% (7/8)	100% (12/12)	100%	92.3%
Sternotomy only	100% (7/7)	1.0 (P = 0.004)	100% (5/5)	100% (2/2)	100%	100%
Sternotomy + neck	100% (5/5)	1.0 (P = 0.013)	100% (2/2)	100% (3/3)	100%	100%
Time between studies						
<14 days	98.4% (62/63)	0.96 (P < 0.001)	95% (19/20)	100% (43/43)	100%	97.7%
≥14 days	100% (22/22)	1.0 (P < 0.001)	100% (3/3)	100% (19/19)	100%	100%

[a] Too few cases for statistical computation.

Clinical Implication:

LUS is a statistically valid diagnostic tool even for high-risk neonates (<3kg) and patients with recent cervical incisions.

Discussion – Comparison with Previous Literature

Study Era	Sample Characteristics	Outcomes/Agreement
Early Studies	Small samples (Post-cardiac/ENT)	Fluctuated 80% – 96%
Recent Data	Smaller cohorts (n=23)	Diagnostic metrics > 90%
Current Advancement (This Study)	Larger prospective validation	**Sensitivity 95.7% / Specificity 100%**

Conclusion: Findings align with established utility but offer significantly higher statistical confidence.

Clinical Utility: Safety & The Power of Exclusion

Leveraging High Negative Predictive Value (NPV) for Screening

Safety Profile

Zero Adverse Events

- Non-invasive, radiation-free, bedside procedure.

The “Rule Out” Capability

NPV = 98.4%

Probability that a negative LUS truly indicates normal movement.

Clinical Translation: A normal LUS result effectively rules out VFMI, negating the need for immediate invasive scoping in asymptomatic patients.

Clinical Utility & Implementation

The Constraint: FL Failure

- 
  Failure rates up to 20%
  (Due to crying/secretions).

- 
  65.9% of patients were aged 2–24 months
  (Difficult-to-scope demographic).

Rescue Pathway

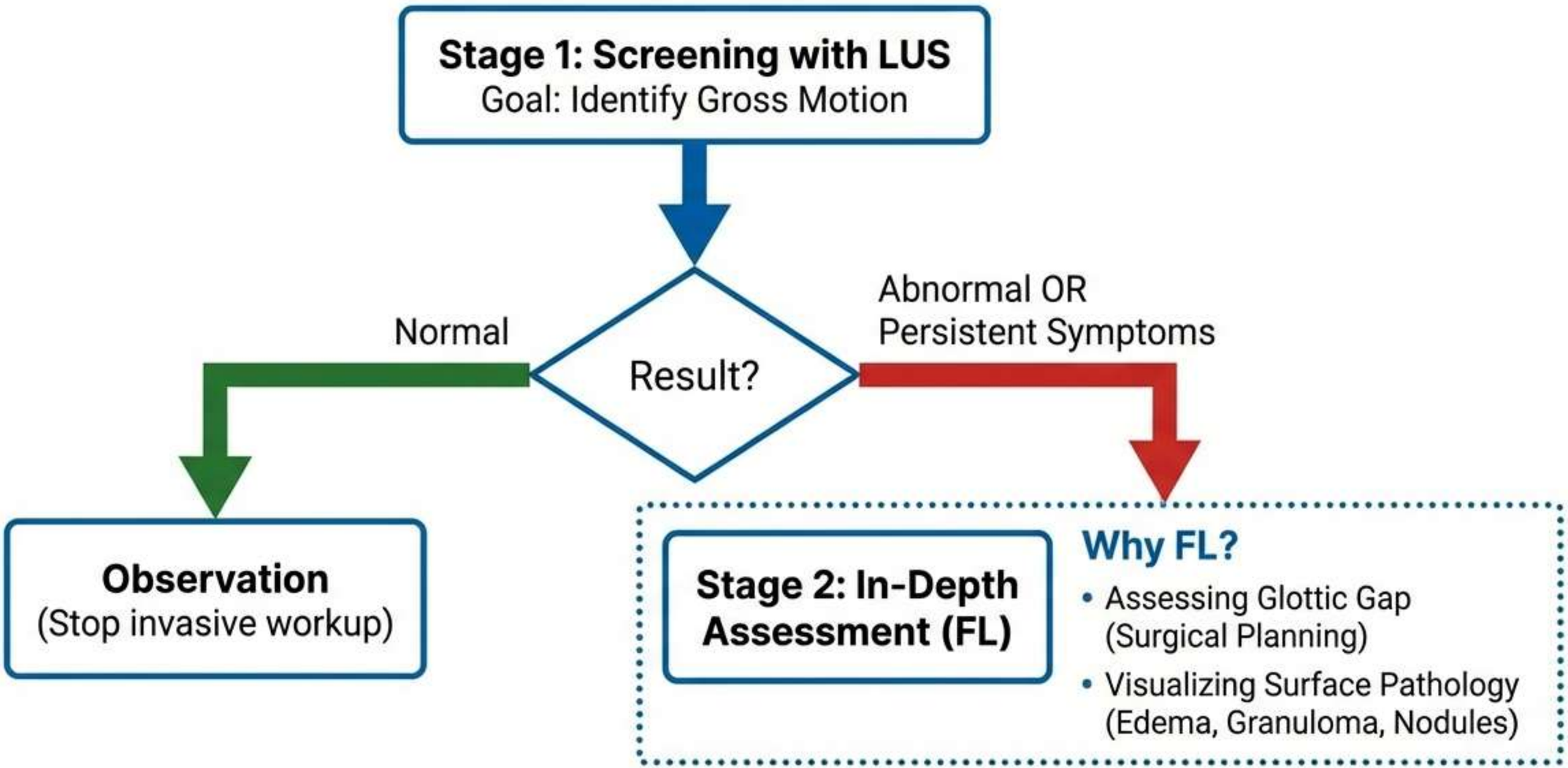
The Rescue: LUS Capability

- LUS provides a reliable alternative when FL is non-diagnostic.

Key Finding: LUS image quality did not affect sensitivity/specificity.

Robust tool for the difficult-to-scope pediatric population.

Proposed Diagnostic Workflow: The LUS-First Protocol



Study Limitations & Considerations

Operator Bias

- LUS is inherently user-dependent.
- High accuracy requires specific ultrasound training.

Physical Barriers

- Tracheostomy tubes cause air interference.
- Acoustic shadowing prevents clear visualization of the vocal folds.

Temporal Factors

- Median time gap: 7 days between LUS and FL.
- Potential for interim spontaneous recovery affecting comparison.

Conclusion & Recommendation

- **Efficacy:** LUS is a highly accurate, safe, and feasible screening tool.
- **Proposal:** Implement 'LUS-First' protocol for routine RLN monitoring.
- **Paradigm Shift:** Transition from invasive reliance to non-invasive screening.

Clinical Impact

↘ Reduces Invasive Procedures

↘ Lowers Healthcare Costs

↗ Improves Patient Safety in ICU/ENT