



DOI: 10.5281/zenodo.23438577 · UDC: 616.33-089.844-072.1

ENDOSCOPIC EVALUATION OF THE ANTIREFLUX VALVE AFTER FUNDOPLICATION: SYNTHESIS OF THE EVIDENCE AND PROPOSAL OF A FUNCTIONAL ENDOSCOPIC SCORE

EVALUAREA ENDOSCOPICĂ A VALVEI ANTIREFLUX DUPĂ FUNDOPLICATURĂ: SINTEZA DOVEZILOR ȘI PROPUNEREA UNUI SCOR ENDOSCOPIC FUNCȚIONAL

Viorel Istrate¹, Doina Fosa^{1*}, Natalia Șipitco¹, Nicolae Bodrug², Sergiu Ungureanu¹

¹ Department of Surgery No. 4, "Nicolae Testemițanu" State University of Medicine and Pharmacy, Chișinău, Republic of Moldova

² Discipline of Geriatrics and Occupational Medicine, "Nicolae Testemițanu" State University of Medicine and Pharmacy, Chișinău, Republic of Moldova

* Corresponding author: doina.fosa@usmf.md

ABSTRACT

Background. Antireflux fundoplication failure — recurrent reflux, dysphagia, gas-bloat syndrome, secondary to a herniated, slipped, too-tight, disrupted, or twisted valve — remains common, and its endoscopic diagnosis lacks a common, standardized reporting language.

Aim. To narratively review existing endoscopic instruments for postoperative antireflux valve assessment, the evidence on their interobserver reliability, and to position an original instrument — the Endoscopic Fundoplication Score (SEF) — within this landscape.

Methods. Narrative review, with a targeted search in PubMed/MEDLINE and Google Scholar (last search: 24 September 2026) using the terms “fundoplication,” “endoscopic grading,” “gastroesophageal valve,” “Hill classification,” “antireflux surgery outcome,” “interobserver reliability,” “wrap integrity,” and “hiatal hernia recurrence,” covering literature published between 1996 and 2026 in English or with an English abstract, supplemented by reference-list screening of identified articles.

Results. Three major endoscopic instruments were identified — the Hill classification (1996), the Jobe system (2004), and the American Foregut Society (AFS) nomenclature (2025) — alongside a descriptive, non-scored taxonomy of failure types (herniated, slipped, tight, disrupted, twisted). The interobserver reliability studies [1, 2] identified in this review for the postoperative setting report low agreement, and a retrospective comparison of endoscopy reports [3] documented major discrepancies between external endoscopists and operating surgeons. The AFS classification of esophagogastric junction integrity achieved substantial agreement on static images ($\kappa=0.65$) but only fair agreement under dynamic conditions ($\kappa=0.371$), and the postoperative AFS nomenclature has not yet been tested for interobserver reliability. The SEF score, an 11-item composite instrument (0–16 points) applied to a cohort of 178 patients, correlated in preliminary single-centre data with the Hill classification, Los Angeles (LA) esophagitis grade, and the Visick patient-satisfaction score (own data, manuscript submitted for publication), offering a numeric, quantifiable alternative to existing categorical systems. The SEF threshold of ≤ 10 points, defined in the same cohort, is exploratory.

Conclusions. The field of postoperative endoscopic antireflux valve assessment remains fragmented, without an interobserver-reliable consensus standard. The SEF score partially addresses this gap through a numeric structure, with experience-based weights and an exploratory threshold, but requires formal interobserver reliability testing and multicenter validation against the surgical gold standard.

Keywords: fundoplication, antireflux valve, endoscopy, Hill classification, interobserver reliability

REZUMAT

Introducere. Eșecul fundoplicaturii antireflux — recidivă de reflux, disfație, sindrom gas-bloat, secundare unei valve herniate, alunecate, prea strânse, disrupte sau torsionate — rămâne frecvent, iar diagnosticul lui endoscopic nu dispune

de un limbaj comun, standardizat, de raportare.

Obiectiv. Sinteza narativă a instrumentelor endoscopice existente pentru evaluarea valvei antireflux postoperatorii, a dovezilor privind fiabilitatea lor interobservator, și poziționarea unui instrument propriu — Scorul Endoscopic de Fundoplicatură (SEF) — în acest context.

Metodă. Revizuire narativă, cu o căutare ținută în PubMed/MEDLINE și Google Scholar (ultima căutare: 24 septembrie 2026), folosind termenii „fundoplication”, „endoscopic grading”, „gastroesophageal valve”, „Hill classification”, „antireflux surgery outcome”, „interobserver reliability”, „wrap integrity” și „hiatal hernia recurrence”, pentru literatura publicată în perioada 1996–2026 în limba engleză sau cu rezumat în limba engleză, completată prin verificarea listelor de referințe ale articolelor identificate.

Rezultate. Au fost identificate trei instrumente endoscopice majore — clasificarea Hill (1996), sistemul Jobe (2004) și nomenclatorul American Foregut Society (AFS, 2025) — alături de o taxonomie descriptivă, nescorată, a tipurilor de eșec (manșon herniat, alunecat, strâns, disrupt, torsionat). Studiile de fiabilitate interobservator identificate în această revizuire [1, 2], pentru contextul postoperator, raportează un acord scăzut, iar o comparație retrospectivă a rapoartelor endoscopice [3] a documentat discrepanțe majore între endoscopiștii externi și chirurghii operatori. Clasificarea AFS a integrității joncțiunii esogastrice a obținut un acord substanțial (substanțial) pe imagini statice ($\kappa=0,65$), dar doar acceptabil (fair) în condiții dinamice ($\kappa=0,371$), iar nomenclatorul postoperator AFS nu a fost încă testat pentru fiabilitate interobservator. Scorul SEF, un instrument compus din 11 componente cu scor de 0–16 puncte, aplicat pe o cohortă de 178 de pacienți, s-a corelat în date preliminare monocentrice cu clasificarea Hill, clasificarea Los Angeles (LA) a esofagitei și scorul Visick de satisfacție a pacienților (date proprii, manuscris trimis spre publicare), oferind o alternativă numerică, cuantificabilă, față de sistemele categoricale existente. Pragul SEF ≤ 10 puncte, definit pe aceeași cohortă, este exploratoriu.

Concluzii. Domeniul evaluării endoscopice postoperatorii a valvei antireflux rămâne fragmentat, fără un standard consensual și fiabil interobservator. Scorul SEF adresează parțial acest gol printr-o structură numerică, cu ponderi bazate pe experiență și un prag exploratoriu, dar necesită testarea formală a fiabilității interobservator și validare multicentrică față de standardul de aur chirurgical.

Cuvinte-cheie: fundoplicatură, valvă antireflux, endoscopice, clasificarea Hill, fiabilitate interobservator

Introduction

Fundoplication — Nissen (360°), Toupet (270°), or Dor (180°) — is the reference surgical treatment for gastroesophageal reflux disease (GERD) refractory to medical therapy or complicated, as well as for symptomatic hiatal hernia [4–7]. Although most patients achieve good long-term symptom control [4, 8], failure of fundoplication — recurrent reflux, persistent dysphagia, or *gas-bloat* syndrome — is reported in a significant proportion of patients during follow-up [9,10]. Failure occurs through a range of distinct anatomical mechanisms: transdiaphragmatic herniation of the valve, slippage of the fundoplication onto the stomach instead of the esophagus, excessive tightness, suture disruption, or axial torsion of the wrap [3, 11–15].

The diagnosis of these failure patterns relies largely on upper gastrointestinal endoscopy, but the interpretation of the postoperative endoscopic image remains, for the most part, descriptive and examiner-dependent. A recent pilot study, limited to Nissen fundoplication, shows that the overall accuracy in describing valve integrity and hernia is only 45.9% [1], which illustrates the absence of a common, standardized reporting language [16] — a problem with direct implications for the comparability of surgical series, for individual therapeutic decision-making (reoperation/conservative management), and for clinical research in the field.

This narrative review sets out three objectives: (1) to synthesize the existing endoscopic instruments for the assessment of the postoperative antireflux valve — the Hill classification, the Jobe system, and the nomenclature proposed by the American Foregut Society (AFS) — alongside descriptive taxonomies of failure types; (2) to synthesize the evidence on the interobserver reliability of these instruments; and (3) to position an original instrument, the Endoscopic

Fundoplication Score (SEF), in this context, including a synthesis of the preliminary results of correlation with clinical indicators of success obtained in our own cohort.

Materials and Methods

This is a narrative review, not a systematic review or a meta-analysis, and it was not prospectively registered (for example, in PROSPERO — *International Prospective Register of Systematic Reviews*); it is therefore not subject to reporting according to PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*). The aim of the search was to identify the major conceptual landmarks of the field and the most relevant recent empirical evidence on the reliability of postoperative endoscopic assessment, not an exhaustive quantification of the entire published literature.

Databases. PubMed/MEDLINE and Google Scholar were queried for literature published between 1996 and 2026 in English or with an abstract available in English; the last search was performed on 24 September 2026. Google Scholar was used as a complement, to identify additional sources, without reporting a number of results.

Search terms. The following were used, individually and in combination (Boolean operators AND/OR): “fundoplication”, “endoscopic grading”, “gastroesophageal valve”, “Hill classification”, “antireflux surgery outcome”, “interobserver reliability”, “wrap integrity”, “hiatal hernia recurrence”.

Inclusion criteria. Original or review articles that proposed, described, or evaluated an instrument or taxonomy for the endoscopic assessment of the postoperative valve/wrap after fundoplication; studies that reported interobserver reliability data for such instruments.

Exclusion criteria. Studies focused exclusively on non-

endoscopic imaging assessment (barium swallow, computed tomography) without an endoscopic component; studies focused exclusively on magnetic sphincter augmentation (LINX) without comparative reporting versus fundoplication; conference abstracts without available full text.

Additional selection. The reference lists of the identified articles were examined manually for additional relevant sources. Additional sources flagged during the review process were evaluated according to the same criteria and included when relevant. This approach, typical of a narrative review with a methodological positioning, carries an inherent risk of subjective source selection, discussed explicitly as a limitation in section 6.

Selection. The PubMed/MEDLINE search, updated on 24 September 2026, returned 60 records. Following title/abstract screening and reference-list review, 25 published sources were included. The authors' own manuscript submitted for publication [25] is cited separately and is not counted as a published source. Six published sources addressed variability or reliability of endoscopic assessment. Literature selection was performed by Viorel Istrate and Sergiu Ungureanu.

Results

Synthesis of existing endoscopic assessment instruments

The Hill classification [17] was originally designed to describe the native gastroesophageal valve (the mucosal fold seen on retroflexion), in 4 grades (I–IV), correlated with the mechanical competence of the cardia. Although it was not developed for the postoperative setting, it has been extended to it by convention, being one of the most widely used instruments in studies on the endoscopic assessment of fundoplication [18].

The system of Jobe et al. [19] was the first instrument explicitly dedicated to the postoperative endoscopic appearance of the valve. Starting from endoscopic photographs of cases considered “successful” after Nissen, Collis-Nissen, Toupet, and Dor fundoplication and Hill repair (7 images for each procedure), a group of surgeons and gastroenterologists described by consensus the ideal endoscopic appearance of each construct. Descriptive criteria were used, such as the lip and body of the valve, the anterior and posterior grooves, the lesser curvature, adherence to the endoscope, the effect of respiration, the valve type, the intra-abdominal location, and the correct position of the repair. The system does not include the Hill grade or manometric correlates. Although conceptually influential, its laborious nature and its derivation from selected cases with an “ideal” outcome limited its clinical applicability, and the system was not widely adopted [20].

The nomenclature proposed by the American Foregut Society (AFS) [20] represents the most recent attempt at standardization, with 4 components — the depth of the anteroposterior grooves, the apposition of the valve on insufflation, the length of the valve, and the position of the

esophagogastric junction (EGJ) — and specific reference grades for each type of construct (Nissen, Toupet, Dor, magnetic sphincter augmentation (LINX), combined transoral procedure (cTIF)). An AFS hiatus grade 1 (no visible aperture on retroflexion; grading adopted from the AFS classification of the integrity of the native esophagogastric junction, distinct from the 4 components above) is defined as the reference postoperative result.

In parallel with these instruments, the surgical literature uses an unscored descriptive taxonomy of fundoplication failure: transdiaphragmatic herniation, slippage of the wrap onto the stomach, excessive tightness, suture disruption, and axial torsion [3, 11–15, 21]. Reported frequencies vary substantially across retrospective series because of differences in patient selection, definitions, and denominators; therefore, pooled percentage ranges are not reproduced here. This taxonomy is clinically useful but does not provide a quantitative score comparable across studies. The instruments are summarized in Table 1.

Interobserver reliability: evidence and clinical implications

Song et al. [1], in a multicenter pilot survey-type study limited to post-Nissen anatomy, included 31 respondents: 8 esophagologists, 9 general gastroenterologists, 9 physicians in training (*fellows*) in gastroenterology, and 5 esophageal surgeons. They interpreted 20 sets of static images representing 5 anatomical variants: intact or disrupted valve, without hernia, with sliding hernia, or with paraesophageal hernia. The overall accuracy was only 45.9% (37.0% in fellows, 52.5% in esophagologists), and the interobserver agreement, measured by Krippendorff's α , was below 0.30 for all variants of valve integrity — a level described by the authors as “*low to nonexistent*”. Being limited to Nissen fundoplication and to a small number of cases, these results cannot be directly extrapolated to other types of construct (Toupet, Dor).

Fantasia et al. [2], in an Australian cross-sectional study that included 101 clinicians (64 upper gastrointestinal surgeons and 37 gastroenterologists), who evaluated 10 cases represented by static endoscopic images after fundoplication, confirmed low interobserver agreement in most domains. The total accuracy across the 10 cases was 76.0% in surgeons and 69.9% in gastroenterologists. Surgeons were significantly more accurate in describing valve integrity ($p=0.014$) and achieved significantly better results in 3 of the 10 cases ($p<0.05$). These data suggest that procedure-specific surgical expertise might influence the accuracy of interpretation, beyond general endoscopic experience.

For the Hill classification, the reliability data identified do not come from studies dedicated to the postoperative setting. Swei et al. [22] reported only acceptable agreement (*fair*) ($\kappa=0.28$) for the Hill classification, versus substantial agreement (*substantial*) for the AFS classification of EGJ integrity ($\kappa=0.65$), and a later study confirmed the superiority of the AFS classification over the Hill classification with respect to interobserver variability [23]. For the Jobe system [19], we did not identify any interobserver reliability study.

Table 1. Comparison of the endoscopic instruments for the assessment of the postoperative antireflux valve identified in the literature.

Instrument	Components	Scale / output	Validation status
Hill classification (1996)	Appearance of the valve fold on retroflexion	4 categorical grades (I–IV)	Developed for the native cardia; extended postoperatively by convention, not by dedicated validation
Jobe system (2004)	Descriptive criteria of the valve: lip, body, anterior and posterior grooves, lesser curvature, adherence to the endoscope, effect of respiration, valve type, intra-abdominal location, etc.	Qualitative description of the ideal appearance, specific to the type of construct; no numeric score	First proposal dedicated to the postoperative setting; derived by consensus from selected “successful” cases; not widely adopted
AFS nomenclature (2025)	Depth of the grooves, valve apposition, valve length, EGJ position	Categorical grades per component, specific to the type of construct	Interobserver reliability not yet tested in the postoperative setting; the related AFS classification of native EGJ integrity: acceptable (fair) agreement ($\kappa=0.371$) under dynamic conditions [20]
Descriptive taxonomy of failure (Banki & Weaver, 2017)	Herniation, slippage, tightening, disruption, torsion	Descriptive, unscored clinical categories	Established clinical framework; not a quantifiable scoring instrument
SEF score (own instrument)	11 components: position, integrity, symmetry, length, tension, typical appearance on retroflexion, hernia, migration of the wrap, slippage/telescoping of the fundoplication construct, torsion, esophagitis	Numeric score of 0–16 points, with 4 interpretation bands and an exploratory binary threshold (≤ 10 points)	Preliminary convergent validity reported in the authors’ cohort of 178 patients [25]; detailed numerical results are reserved for the original manuscript submitted separately. Interobserver reliability untested; threshold ≤ 10 exploratory and not externally validated.

AFS = American Foregut Society; EGJ = esophagogastric junction; LA = Los Angeles classification of esophagitis; SEF = Endoscopic Fundoplication Score; ρ = Spearman correlation coefficient; κ = kappa coefficient of interobserver agreement.

Relevant only indirectly to the postoperative setting, Al Lami et al. [24] included patients referred for endoscopy for the evaluation of GERD, excluding those with previous foregut surgery, and tested the AFS classification of esophagogastric junction (EGJ) integrity (grades 1–4; components: length, diameter, valve), distinct from the postoperative nomenclature described in section 3 [20]. In a prospective tandem endoscopy design in 70 patients evaluated by 5 endoscopists, the authors reported only acceptable agreement ($\kappa=0.371$) when the AFS classification of EGJ integrity was applied under dynamic, real-time examination conditions — significantly lower than the substantial agreement ($\kappa=0.65$) previously reported on static images [22]. Agreement improved to a moderate level ($\kappa=0.524$) only

when the grades were simplified into a binary form (AFS 1–2 vs. 3–4), and the “diameter” component (which describes the hiatus and is not part of the four components of the postoperative nomenclature described in section 3) showed poor agreement ($\kappa=0.077$). In 38.9% of the cases evaluated by the esophagologist (14/36), the esophagologist classified the EGJ into a higher category than the first endoscopist did, and the classification showed a weak correlation with the GERDQ questionnaire (symptom questionnaire; $\rho=0.077$).

The practical implication of these studies is direct: none of the instruments evaluated in the identified studies has demonstrated, to date, interobserver reliability sufficient for clinical or research use without specific training and without an experienced second opinion. The AFS postoperative

nomenclature [20] has not yet been formally evaluated in this respect. This finding limits the direct comparability of the results reported by different centers using the Hill classification or other descriptive systems and underscores the need for dedicated efforts to standardize the interpretation, not only the assessment criteria.

The SEF score as a response to the “identified gap”

The Endoscopic Fundoplication Score (SEF) was developed as a composite, numeric instrument applicable to postoperative upper gastrointestinal endoscopy, with the explicit aim of providing a quantifiable alternative to the existing categorical systems. The score sums 11 components, each with its own maximum weight, up to a total of 16 points: valve position (0–2), integrity (0–3), symmetry (0–1), length (0–1), tension (0–2), appearance on retroflexion (0–2), and absence of: hiatal hernia (0–1), migration (0–1), slippage/telescoping (0–1), torsion (0–1), and severe esophagitis (0–1). The total score is divided into 4 interpretation bands — functional valve (14–16 points), minor abnormalities (11–13 points), partial disruption (8–10 points), and failed fundoplication (0–7 points) — and includes an operational binary threshold (≤ 10 points) for a “SEF-positive” test. The 11 components cover three domains described in the literature: the functional appearance of the valve (position, integrity, symmetry, length, tension, appearance on retroflexion), present, in different wording, in the Hill classification [17] and in the AFS nomenclature [20]; the failure mechanisms of the descriptive taxonomy [11] (hernia, migration, slippage, torsion); and an endoscopic marker of consequence, severe esophagitis (Los Angeles classification [26]). The weights reflect the grading of each criterion: circumferential integrity, the main function of the construct, has 0–3 points; position, tension, and appearance on retroflexion, where a clinically distinct intermediate state exists, have 0–2 points each; present/absent criteria have 0–1 point. The weights were set a priori, on the basis of the authors’ endoscopic experience and the literature, not through statistical analysis of the cohort, and were not optimized on the data. They are therefore experience-based and not empirically validated. The cut-offs of the interpretation bands (14–16, 11–13, 8–10, 0–7 points) were set by the authors.

The SEF score is an original instrument, built by the authors for their own cohort, unpublished and not previously validated; the ≤ 10 point threshold was derived from the analyzed data. The binary components are scored 1 (normal appearance) or 0; the others: position of the fundoplication construct — 2 correct intra-abdominal, 1 partial proximal displacement, 0 — located entirely above the hiatus; integrity — 3 intact, 2 minor dehiscence ($< 25\%$), 1 partial dehiscence ($25\text{--}50\%$), 0 complete dehiscence/absent valve ($> 50\%$); tension — 2 adequate, 1 slightly loose or slightly tight, 0 obstructive or completely loose; retroflexion — 2 valve well conformed around the endoscope, 1 partial/irregular conformation, 0 absent or deteriorated. Severe esophagitis means Los Angeles (LA) classification grade C/D.

SEF components are not fully independent: valve position,

herniation, migration, and slippage may reflect the same anatomical abnormality. The authors explored this overlap in post-hoc sensitivity analyses, including grouping anatomical components and applying equal weights [25]. According to the authors, the overall associations with the Hill and LA classifications and Visick score remained similar. However, these exploratory analyses do not eliminate the risk of multiple penalties for one anatomical abnormality or replace validation in independent cohorts.

Severe esophagitis (LA C/D) is a consequence of reflux rather than an anatomical characteristic of fundoplication. The authors included it as a one-point endoscopic marker of valve incompetence [26]. This introduces partial circularity when comparing SEF with LA grade. An exploratory analysis excluding this item reportedly preserved the general trends [25], but this limitation remains relevant to interpretation.

SEF was preliminarily applied in a single-center cohort of 178 patients, described in the original manuscript submitted separately for publication [25]. The authors report associations with the Hill classification, Los Angeles esophagitis grade, and Visick score, as well as an association between a score ≤ 10 and resumption of proton pump inhibitor therapy. Detailed numerical results and statistical analyses of the cohort belong to the original study and are not reproduced in this narrative review. The ≤ 10 threshold was derived and assessed in the same cohort, without internal or external validation, and must be considered exploratory rather than diagnostic. PPI (proton pump inhibitor) resumption is an indirect indicator of reflux, and SEF interobserver reliability remains untested.

The central limitation, explicitly acknowledged, is precisely the absence, to date, of a dedicated interobserver reliability study for the SEF score, analogous to those conducted for the postoperative assessment of valve integrity and hernia [1, 2] or for the AFS classification of EGJ integrity compared with the Hill classification [22–24]. Without such a demonstration, the SEF score shares, at least provisionally, the reproducibility uncertainty of the whole field, even though its numeric structure makes it, at least in theory, a candidate that is easier to test and to train than the categorical alternatives.

Future directions and limitations

The need for interobserver reliability studies for the SEF score

The next necessary methodological step is a dedicated interobserver reliability study for the SEF score, following the design of studies already published for other instruments — either on sets of standardized static images (replicating the study of Song et al.) or, ideally, through a real-time tandem endoscopy design (replicating the study of Al Lami et al.), in order to capture the real variability of dynamic examination, not only that of the interpretation of fixed images. Prior training of the raters and an illustrated scoring manual, with reference examples for each grade of each component, could improve agreement, as suggested by the observed difference between agreement on static and dynamic images for the AFS classification of EGJ integrity [24].

Validation against the surgical standard

Formal diagnostic validation of the SEF score (sensitivity, specificity, predictive values, area under the Receiver Operating Characteristic (ROC) curve) against intraoperative findings at surgical reoperation — the only true objective standard for fundoplication failure — remains an essential but unachieved objective, owing to the insufficient number of reoperated cases in our own cohort [25]. Systematic prospective collection of these data, including through multicenter collaboration, is necessary to establish the true diagnostic properties of the score, beyond the preliminary convergent validity.

Multicenter studies and standardization

The SEF score has been developed and applied, to date, in a single center. By contrast, the AFS nomenclature [20] was developed through the consensus of a multi-institutional working group of a professional society, but it has not yet been prospectively validated in multicenter studies. Extension to a multicenter setting — with raters from different institutions, varied levels of experience, and, ideally, a common standardized photographic/video-endoscopy platform — is the logical next step both for testing interobserver reliability and for generalizing the convergent validity results reported here. This direction is consistent with the research priorities explicitly formulated by the American Foregut Society [20]: the correlation between the postoperative endoscopic appearance and effective reflux control, the short-/long-term predictive capacity, and the impact on quality of life.

Limitations of this review

Being a narrative review and not a prospectively registered systematic review, this work does not follow a PRISMA protocol, does not report an exhaustive, independently reproducible search strategy, and does not formally assess the risk of bias of the included studies. The selection of sources, although guided by the search terms declared in section 2, retains an inherent degree of subjectivity, typical of this type

of positioning article. The interpretation of the results must take into account the language restriction to literature in English or with an English abstract. In addition, the authors are the developers of the SEF score evaluated in the paper, and the data on the own cohort of 178 patients are preliminary and come from an unpublished dataset, presented here on the basis of the manuscript submitted for publication. Although the manuscript reports in section 2 the number of records identified and of sources included, the selection was not carried out through a complete systematic flow, so the risk of omitting relevant studies cannot be excluded. The SEF score has experience-based weights that are not empirically validated, partially overlapping components, partial circularity between the “severe esophagitis” item and LA grade, and an exploratory threshold defined in the same cohort without external validation.

Conclusions

Postoperative endoscopic evaluation of the antireflux valve after fundoplication remains a fragmented field, with few instruments, none of which has demonstrated, in the identified studies, satisfactory interobserver reliability; the postoperative nomenclature of the American Foregut Society, the most recently proposed, has not yet been tested in this respect, and the related classification of the native EGJ achieved only “fair” agreement under dynamic conditions. The SEF score, through its numeric structure and through the explicit quantification of failure mechanisms rarely formalized in other systems, offers a potential alternative, experience-based weights, an exploratory threshold, and preliminary convergent validity data obtained in a single-center cohort of 178 patients. Formal testing of interobserver reliability and multicenter validation against the surgical gold standard remain the necessary steps before any broader adoption.

Received – 07.09.2026, accepted for publication – 16.10.2026, published online – 17.10.2026

Author contributions statement: Viorel Istrate – study conception and design, manuscript writing, investigation, formal analysis, statistical analysis of the cohort data, supervision; Doina Fosa – data collection, manuscript writing; Natalia Șipitco – data collection, manuscript writing; Nicolae Bodrug – critical review and editing of the manuscript; Sergiu Ungureanu – study conception and design, manuscript writing, investigation, formal analysis, statistical analysis of the cohort data, supervision, critical review and editing of the manuscript. All authors have read and approved the final version.

Funding statement: This research received no external funding.

Conflict of Interest: The authors are the developers of the SEF endoscopic score discussed in the paper; they have no financial interests related to it.

Ethics approval: This manuscript is a narrative review; however, it references preliminary data from the authors' own cohort of 178 patients. The underlying clinical data collection is part of a doctoral research project concerning patients with gastroesophageal reflux disease and hiatal hernias, which was approved by the Research Ethics Committee of the “Nicolae Testemițanu” State University of Medicine and Pharmacy (Project No. 84, approved on 21.12.2022). The sharing of the anonymized dataset is also covered by this ethics approval.

Data availability: The anonymized dataset of the analyzed cohort is available from the corresponding author upon reasoned request; its sharing is also covered by the ethics approval of the research project that included these patients.

Citation: Istrate V, Fosa D, Șipitco N, Bodrug N, Ungureanu S. Endoscopic evaluation of the antireflux valve after fundoplication: synthesis of the evidence and proposal of a functional endoscopic score [Evaluarea endoscopică a valvei antireflux după fundoplicatură: sinteza dovezilor și propunerea unui scor endoscopic funcțional]. *Arta Medica*. 2026;101(4):e2026407. doi: 10.5281/zenodo.23438577

Author Information: Viorel Istrate - <https://orcid.org/0000-0002-1243-0716>; Doina Fosa - <https://orcid.org/0009-0009-5389-788X>; Natalia Șipitco - <https://orcid.org/0009-0003-6733-809X>; Nicolae Bodrug: <https://orcid.org/0000-0003-0295-1574>; Sergiu Ungureanu - <https://orcid.org/0000-0002-3707-4738>

Bibliography

1. Song EJ, Yadlapati R, Chen JW, et al. Variability in endoscopic assessment of Nissen fundoplication wrap integrity and hiatus herniation. *Dis Esophagus*. 2022;35(5):doabo78. doi:10.1093/dote/doabo78
2. Fantasia JJ, Cock C, Watson DI, Bright T, Thompson SK. Assessment of laparoscopic fundoplication with endoscopy: room for improvement. *Surg Endosc*. 2024;38(2):713-719. doi:10.1007/s00464-023-10570-4
3. Latorre-Rodríguez AR, Kim P, Mittal SK. Endoscopic assessment of failed fundoplications differs between endoscopists. *Surg Endosc*. 2024;38(11):6839-6845. doi:10.1007/s00464-024-11107-z
4. DeMeester SR. Laparoscopic hernia repair and fundoplication for gastroesophageal reflux disease. *Gastrointest Endosc Clin N Am*. 2020;30(2):309-324. doi:10.1016/j.giec.2019.12.007
5. Kasalický M, Koblihová E. Chirurgie hiátové kýly a refluxní choroby jícnu, Nissen, nebo Toupet? [Surgery of the hiatal hernia and gastroesophageal reflux disease, Nissen or Toupet?]. *Rozhl Chir*. 2015;94(12):510-515. Czech.
6. Bonavina L. Individualizing the choice of surgical therapy for gastroesophageal reflux disease. *Curr Opin Gastroenterol*. 2025;41(4):245-250. doi:10.1097/mog.0000000000001082
7. Patrzyk M, Hummel R, Kersting S. Operative Strategie bei Hiatushernien [Surgical strategy for hiatal hernias]. *Chirurgie (Heidelb)*. 2024;95(4):336-344. German. doi:10.1007/s00104-024-02054-5
8. Zhu T, Gong D. Clinical outcomes of 3D laparoscopic hiatal hernia repair either combined with Toupet fundoplication or Nissen fundoplication: a comparative analysis. *Ann Ital Chir*. 2025;96(12):1706-1714. doi:10.62713/aic.4375
9. Siemssen B, Hentschel F, Ibach MJ. Long-term results after laparoscopic revision fundoplication: a retrospective, single-center analysis in 194 patients with recurrent hiatal hernia. *Esophagus*. 2024;21(3):390-396. doi:10.1007/s10388-024-01060-0
10. Aiolfi A, Bona D, Wang Q, Bonavina L. Reflections on surgery for hiatal hernia. *Updates Surg*. 2026;78(1):1-8. doi:10.1007/s13304-025-02490-w
11. Banki F, Weaver M. Failed fundoplication and complications of antireflux surgery: radiographic, endoscopic and laparoscopic views. *JSM Gen Surg Cases Images*. 2017;2(1):1021. doi:10.47739/2573-1564/1020
12. Hunter JG, Smith CD, Branum GD, et al. Laparoscopic fundoplication failures: patterns of failure and response to fundoplication revision. *Ann Surg*. 1999;230(4):595-604. doi:10.1097/00000658-199910000-00015
13. Awais O, Luketich JD, Schuchert MJ, et al. Reoperative antireflux surgery for failed fundoplication: an analysis of outcomes in 275 patients. *Ann Thorac Surg*. 2011;92(3):1083-1090. doi:10.1016/j.athoracsur.2011.02.088
14. Furnée EJ, Draaisma WA, Broeders IA, Gooszen HG. Surgical reintervention after failed antireflux surgery: a systematic review of the literature. *J Gastrointest Surg*. 2009;13(8):1539-1549. doi:10.1007/s11605-009-0873-z
15. Khajanchee YS, O'Rourke R, Cassera MA, et al. Laparoscopic reintervention for failed antireflux surgery: subjective and objective outcomes in 176 consecutive patients. *Arch Surg*. 2007;142(8):785-792. doi:10.1001/archsurg.142.8.785
16. Boys JA, Azadgoli B, Martinez M, Oh DS, Hagen JA, DeMeester SR. Adequacy of EGD reporting: a review of 100 reports from 100 endoscopists. *J Gastrointest Surg*. 2021;25(5):1117-1123. doi:10.1007/s11605-020-04634-2
17. Hill LD, Kozarek RA, Kraemer SJ, et al. The gastroesophageal flap valve: in vitro and in vivo observations. *Gastrointest Endosc*. 1996;44(5):541-547. doi:10.1016/S0016-5107(96)70006-8
18. Fantasia JJ, Thompson SK. Assessment of anti-reflux surgery with endoscopy: a narrative review. *Ann Esophagus*. 2023;6:43. doi:10.21037/aoe-22-11
19. Jobe BA, Kahrilas PJ, Vernon AH, et al. Endoscopic appraisal of the gastroesophageal valve after antireflux surgery. *Am J Gastroenterol*. 2004;99(2):233-243. doi:10.1111/j.1572-0241.2004.04042.x
20. Nguyen NT, Bell R, Abu Dayyeh B, et al. American Foregut Society white paper on the endoscopic assessment of the gastroesophageal valve after anti-reflux surgery. *Foregut*. 2025;5(2):95-106. doi:10.1177/26345161241300480
21. Banki F, Kaushik C, Roife D, et al. Laparoscopic reoperative antireflux surgery: a safe procedure with high patient satisfaction and low morbidity. *Am J Surg*. 2016;212(6):1115-1120. doi:10.1016/j.amjsurg.2016.09.011
22. Swee E, Almario JA, Abou Diwan E, Ashhab AA, Canto MI. American Foregut Classification is superior to Hill classification in evaluation of the gastroesophageal junction. *Foregut*. 2024;4(4):404-410. doi:10.1177/26345161241293522
23. Siboni S, Sozzi M, Lovece A, et al. The AFS endoscopic classification of esophago-gastric junction integrity is superior to the Hill classification in terms of interobserver variability. *J Clin Gastroenterol*. 2025. doi:10.1097/MCG.0000000000002288
24. Al Lami Z, Jaber F, Ayoub FW, et al. Interobserver agreement of the American Foregut Society esophagogastric junction classification: a prospective observational study. *Foregut*. 2026;6(1):22-30. doi:10.1177/26345161251393104
25. Istrate V, Fosa D, Şipitco N, Bodrug N, Ungureanu S. Clinico-endoscopic correlation in gastroesophageal reflux disease and hiatal hernias: the Endoscopic Score of antireflux Fundoplication versus postoperative symptomatic status. *Chirurgia (Bucureşti)*. Submitted for publication.
26. Lundell LR, Dent J, Bennett JR, et al. Endoscopic assessment of oesophagitis: clinical and functional correlates and further validation of the Los Angeles classification. *Gut*. 1999;45(2):172-180. doi:10.1136/gut.45.2.172