The Amsterdam Classification of Completeness of Endoscopic Sinus Surgery (ACCESS): a new CT-based scoring system grading the extent of surgery*

S. Reitsma, G.F.J.P.M. Adriaensen, M.E. Cornet, R.M. van Haastert, M.H. Raftopulos, W.J. Fokkens

Department of Otorhinolaryngology, Amsterdam University Medical Centers, location AMC, Amsterdam, The Netherlands

Rhinology 58: 6, 538 - 543, 2020
https://doi.org/10.4193/Rhin20.165

***Received for publication:**
April 17, 2020
Accepted: May 13, 2020

Abstract

Background: A debate is ongoing on the role of the extent of sinus surgery in disease control in chronic rhinosinusitis (CRS). The newly developed Amsterdam Classification on Completeness of Endoscopic Sinus Surgery (ACCESS) score provides a way to quantify extent of surgery. This study aimed to validate the ACCESS scoring system and to report its interrater agreement compared to the widely used Lund-Mackay (LM) scoring system.

Methodology: Forty hand-picked anonymized computed tomography scans of sinuses of patients with varying pathology and degree of previous sinus surgery were independently scored by six rhinologists. Interrater agreement was determined by the intraclass correlation (ICC) statistic.

Results: The interrater agreement of the ACCESS score was excellent, comparable to the LM score. The ACCESS interrater agreement was not influenced by degree of opacification or diagnosis. The ACCESS score reliably measured predicted differences induced by sinus surgery.

Conclusions: the ACCESS score is an easy-to-use valid tool to assess extent of sinus surgery with an excellent interrater agreement. Further validation in a random group of CRS cases is required.

Key words: chronic rhinosinusitis, CT scan, endoscopic sinus surgery, extent of surgery, interrater agreement

Introduction

Over the past decades, a debate is ongoing on the role of the extent of sinus surgery in chronic rhinosinusitis (CRS). Many argue that more extensive surgery would allow better access of the sinus system to local medication, thus improving disease control. Many aspects of the extent of surgery have been published upon, such as the approach to the maxillary sinus[1], the frontal sinus[2], and the use of 'reboot' techniques aiming to completely remove all diseased sinus mucosa. This latter idea was originally published by Jankowski in 1997[3] and is now again gaining attention over the past years[4]. However, most papers in this area suffer from methodological flaws, such as small groups, poor or no control groups, blinding issues, etc. As extensively discussed in the EPOS2020 guidelines, there are no conclusive data on this issue yet and the 'impact of extent of surgery' is therefore highlighted as a research need for the coming period[5]. However, a validated tool to assess the extent of surgery is currently lacking.

Already at the introduction of the now widely accepted Lund-Mackay (LM) scoring system, a subsection was written on the extent of surgery[6]. For each anatomical site (e.g., posterior ethmoid, ostiomeatal complex, etc.) a 0 or 1 indicated whether surgery had been performed here. However, these scores were not accounted for in the total LM score which was based on the opacification scores per site alone. As such, common use of the LM system now entails only scoring opacification as a marker of disease severity. Moreover, one can often find cases where a previous surgeon performed incomplete surgery (e.g., opening

the ethmoidal bulla but leaving the entire rest of the anterior ethmoid in place) which would trigger a discussion on how to grade on a 2-point scale.

An imaging-based scoring system is only valid and useful if it is easy to apply and has acceptable interrater agreement. Based on an elaborate comparison of various scoring systems for disease extent, Metson et al. concluded in 1997 that despite a low interrater agreement (based on kappa statistics), the LM system was preferable because of its low complexity[7]. Therefore, we adapted this system into the Amsterdam Classification of Completeness of Endoscopic Sinus Surgery (ACCESS) scoring system, aiming to give a quantitative measure of the extent of endoscopic sinus surgery. The main paradigm is the question whether a sinus is functionally opened, i.e. are the bony boundaries addressed in such a way that local medication (e.g., saline irrigation) is able to enter the sinus without the need of extended surgical procedures / drilling. By following the LM structure, the ACCESS score is easy to learn and apply. The current study was performed to establish the interrater agreement of the ACCESS score and to validate its sensitivity to (additionally) performed sinus surgery.

Materials and methods

The ACCESS scoring system

The structure of the scoring system is very similar to the Lund-Mackay (LM) system; per side, 6 anatomical sites are graded with a 0, 1, or 2. Contrary to the LM score, the ACCESS does not focus on sinus opacification, but only on access to the sinus based on bony boundaries. The sites are the same as the LM score: frontal sinus, anterior ethmoid, posterior ethmoid, sphenoid sinus, maxillary sinus, and ostiomeatal complex (OMC).

A score of 0 means that no additional surgery is needed to warrant access to this site ('functionally opened'). A score of 1 means that previous surgery did address this site, but was inadequate to open it fully ('touched but not functional'). A score of 2 means that no previous surgery was performed to this sinus/site. In analogy to the LM score, the ostiomeatal complex can only be scored with a 0 or 2; also, a non-developed sinus is scored 0. This way, 6 sites are graded per side, which are added up to give the total ACCESS score. Thus, the maximum ACCESS score is 24 (as the LM score). Higher ACCESS scores suggest a lower extent of previous surgery, while lower scores indicate adequate surgery was performed to address most sinuses. This might be counter-intuitive as one could argue that larger scores should indicate larger procedures performed. However, we choose the current setup for a better fit with the LMS. Either way, the direction of the scoring (from 0 to 2 or from 2 to 0) does not influence the way the score is obtained / the mechanics of the ACCESS 'reasoning'.

All authors (experienced rhinologists) were given written instructions including some examples (supplementary material). No further training was given.

Selection of scans and scoring procedure

Forty sinus computed tomography (CT) scans were collected by the first author (SR). These were all scanned according to the same protocol including multiplanar reconstructions in three planes (axial, coronal, sagittal) with a slice thickness of 1.5 mm. Based on the patient records (e.g., surgical reports and/or findings from nasal endoscopy), an estimate was made of the extent of previous surgeries (if any). This way, a case mix was constructed with the full range of expected ACCESS scores. To evaluate the influence of sinus opacification on the reliability of the ACCESS scores, a fair distribution was made between scans with significant sinus opacification (n=22) and those with no (significant) opacification (n=18). To test the sensitivity of the ACCESS scoring system to additional surgery, seven scans were selected that were made after additional surgery (with the preoperative scans also in the case mix). All scans were anonymized and made available in three planes (axial, coronal, sagittal). The scans were scored by each author separately. Both the ACCESS and the LM scores were obtained per scan. The online database (Castor EDC, Amsterdam, the Netherlands) was set up in such a way that authors were blinded to the scores of others.

Statistical analysis

After completion of the scoring of the scans by all authors, the database was locked and data were exported to the SPSS software package (IBM SPSS Statistics version 26, IBM Corporation). Per scan, the ACCESS and LM scores per site were added up to give the total ACCESS and LM scores. Interrater agreement was determined using the intraclass correlation (ICC) for total scores with a two-way mixed, consistency-based, average-measures model.

ICC results were interpreted as: <0.50: poor; 0.50 – <0.75: moderate; 0.75 – <0.90: good; 0.90 and above: excellent(8). The interrater agreement of the generally accepted LM scoring system was used as further reference. The level of statistical significance was set at an alpha of 0.05.

Sub-analyses were performed based on:

1. the level of opacification (low: median LM score <5 (n=18); high: median LM score ≥5 (n=22));
2. diagnosis (CRS with nasal polyps (n=10), any CRS (n=18), and others (such as inverted papilloma, osteoma, non-sinogenic headache etc.; n=22);
3. a large variation between raters in the ACCESS scores. For scans with an interquartile range >5, the scans and scores were analyzed per site to find possible causes for disagree-

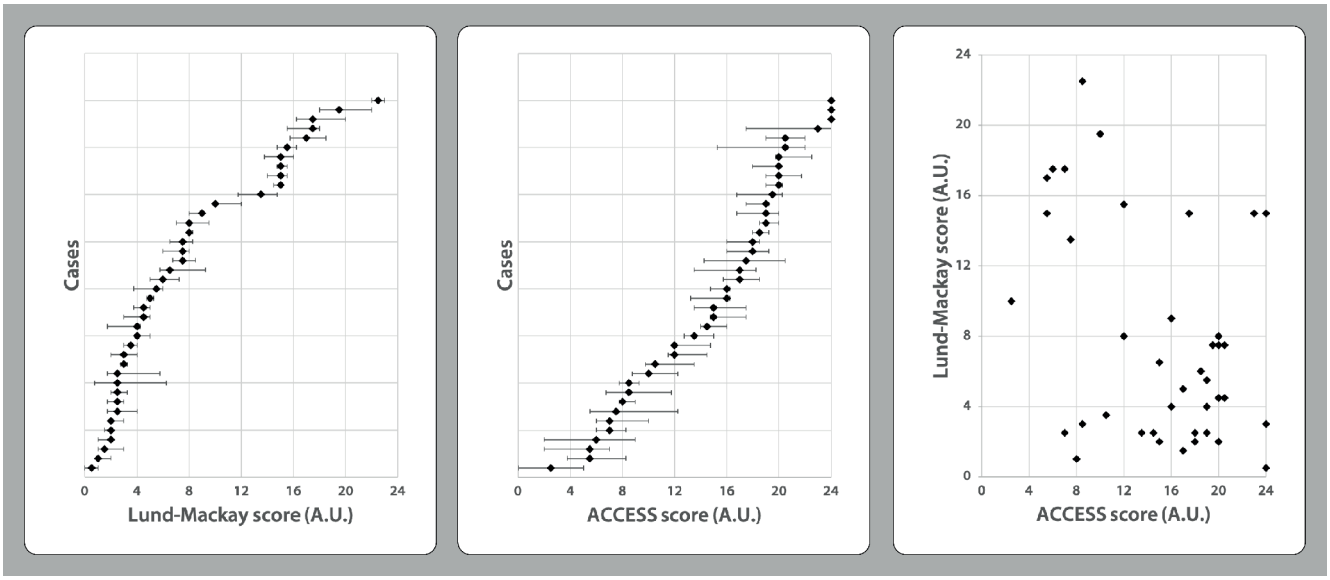

Figure 1. Distribution of ACCESS and Lund-Mackay scores over the case-mix. Left and middle panel: median Lund-Mackay and ACCESS scores with interquartile ranges for the 40 cases. Note that for both scoring systems, the full range of scores was incorporated in the case mix. The position of the various cases does not correspond between these two panels. Right panel: combined median Lund-Mackay and ACCESS scores per case.

Table 1. Interrater agreement measures for Lund-Mackay and ACCESS scores.

	Total group (n=40)	Low opacification (n=18)	High opacification (n=22)	CRSwNP (n=10)	Any CRS (n=18)	Non-CRS (n=22)
	Mean ICC (95% CI)	Mean ICC (95% CI)	Mean ICC (95% CI)	Mean ICC (95% CI)	Mean ICC (95% CI)	Mean ICC (95% CI)
Lund-Mackay	0.992 (0.988-0.995)	0.816 (0.644-0.922)	0.988 (0.978-0.994)	0.962 (0.909-0.989)	0.991 (0.983-0.996)	0.980 (0.963-0.991)
ACCESS	0.977 (0.964-0.986)	0.985 (0.971-0.994)	0.973 (0.950-0.987)	0.973 (0.936-0.992)	0.976 (0.954-0.990)	0.977 (0.959-0.989)

ICC: intraclass correlation; CI: confidence interval; CRS(wNP): chronic rhinosinusitis (with nasal polyps).

ment between raters;

4. the effect of surgery by including consecutive scans before and after surgeries (n=7). Raters were unaware that these scans were included.

Results

Overall interrater agreement

The distribution of the LM and ACCESS scores over 40 CT scans as rated by six observers is depicted in Figure 1, showing a well-balanced case mix. The interrater agreement statistics are listed in Table 1. The intraclass correlation data show that both scoring systems had an excellent interrater agreement.

Interrater agreement and degree of opacification and diagnosis

From Table 1, it is clear that the ACCESS score had an excellent interrater agreement in both low and high opacification groups, as well as in the various diagnosis groups. The same is true for the LM score. The ICC of the latter was only slightly reduced in the low opacification group ('good' interrater agreement).

Analyses of cases with large variation in ACCESS scores

In order to define 'problematic' scans for the ACCESS score, those with an ACCESS interquartile range of 5 or more (n=8) were analyzed separately. All had significant opacification (median LM scores of 7.5 – 22.5) and (thus) almost all were from CRS cases (n=7). Interestingly, these scans had either low or high ACCESS scores. For the high ACCESS scores (around 20), discrepancies arose from sites that were opened according to some, while others deemed them untouched. The level of opacification probably hampered a comparable judgement (Figure 2, left upper panel). For the low ACCESS scores (around 5), most discrepancies arose from the ethmoid and frontal sinuses, especially in the presence of osteitis, an underdeveloped frontal sinus and/or an anterior ethmoidal artery in mesentery (Figure 2, other panels).

Sensitivity of the ACCESS scoring system to surgery

Figure 3 shows the results of the blinded rating of pre- and postoperative scans, compared to the expected result. Cases 1

Figure 2. Examples of reasons for low interrater agreement for ACCESS scores. Left upper panel: coronal view. In this dense image, raters disagreed on whether the maxillary sinus / ostiomeatal complex had been addressed bilaterally or not, easily giving a scoring difference of several points. Right upper panel: sagittal view of previously operated sinuses; the anterior ethmoidal artery is in mesentery with a relatively large bony remnant still attached to it (arrow). Furthermore, the frontal sinus is underdeveloped (or, according to some raters: not present). Again, a scoring difference of several points is easily induced. Lower panels: coronal and sagittal views from a single case showing extensive osteitis (and an anterior ethmoidal artery in mesentery). Raters disagreed largely on the extent of adequate surgery to the ethmoids and frontal sinus.

and 2 have an expected ACCESS score difference of 0 as they encompass cases with a previously adequately opened maxillary sinus undergoing either removal of scar tissue at the infundibulum or a medial maxillectomy. Other cases represent increasing degrees of surgery, from a unilateral sphenoidotomy to a Draf type III in a previously unoperated patient. Despite the (small) variations in the rating of the pre-operative scans, the six raters where quite able to score the difference induced by surgery according to the expected difference. In other words, the system proves to be sensitive and accurate when indicating differences induced by surgery.

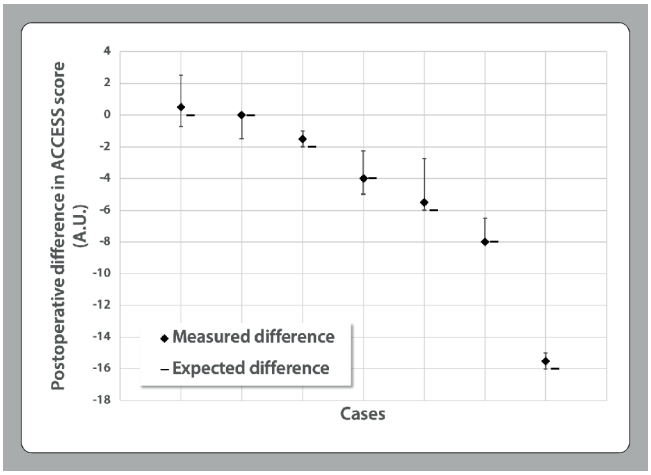

Figure 3. Sensibility of ACCESS to postoperative changes.Median change in ACCESS score per rater (including interquartile ranges) for seven cases that were blindly scored pre- and postoperatively. The measured differences (diamonds) comply with the expected differences (dashes) to a great extent.

Discussion

In a non-random selection of forty cases, we show that the ACCESS score is an easy-to-use and sensitive tool to assess the extent of endoscopic sinus surgery. It can be mastered quickly by rhinologists through written instructions without the need for additional training. Scans can be assessed for both ACCESS and LM scores within minutes. When looking at ICC statistics, both scoring systems have an excellent interrater agreement.

Kappa statistics have often been used in this kind of studies to determine interrater agreement. As such, it is a helpful measure to be able to compare current outcomes to existing literature. However, for measurements with many categories (24 in this case), kappa statistics will automatically give poor results as they require an exact match. But a poor kappa statistic does not automatically mean a large or relevant disagreement between raters; one could debate, for example, how clinically relevant the difference between 17 and 18 really is for the Lund-Mackay score. A possible solution would be to limit the number of categories by scoring subdivisions instead of total LM or ACCESS scores. However, to really overcome the influence of exact matches (or the lack thereof), determining the ICC is needed as it also incorporates the magnitude of disagreement between raters. As a result, the outcome will not deteriorate much when the absolute difference between rated scores is small, whereas larger absolute differences will give lower outcomes[9]. Therefore, the presented ICC results are much more informative on the extent of interrater agreement. As expected, the kappa values for the LM scores of the current study fit well into existing literature (data not shown)[7,10,11].

In the current setup, the ACCESS scores were determined by experienced rhinologists. It is likely that – in contrast to the LM score – non-surgically experienced observers might show greater interrater variability. It requires (some) surgical insight especially for the ethmoid and frontal sinuses to be graded adequately. Another drawback compared to the LM score is that the ACCESS score cannot be graded accurately on other imaging modalities, such as magnetic resonance imaging, as it depends heavily on the presence or absence of (thin) bony boundaries. Furthermore, these thin boundaries dictate that a reliable ACCESS score can only be obtained from a sinus CT scan with adequate settings/slice thickness.

As indicated previously (introduction and written instructions (online material)), the central paradigm for obtaining the ACCESS score is whether a sinus/site is deemed to be functionally opened. This 'functional viewpoint' can give rise to differences in interpretation. To minimize these differences, one could translate this paradigm into two questions:

- "Has any surgery been performed in this site/sinus?" (Score 2 if "no")
- If "yes": "Do any more bony ridges/cells/lamellae need to be removed to further open up this site/sinus so that it can be reached by local medication / rinsing without using extended procedures / drilling?" (Score 1 if "yes" and 0 if "no").

It is especially important to answer these questions without looking at opacification (if any); the ACCESS score should quantify surgical extent, whereas the LM score is a marker of disease severity. The fact that the interrater agreement data of the ACCESS score did not differ between scans with and without opacification, shows that this is feasible. However, this functional viewpoint remains a possible limitation of the current setup. A correctly executed Draf type IIa would lead to a score of 0 (functionally opened / complete surgery) for the frontal sinus, even in case of a prominent nasal beak. One could argue that in selected cases only a Draf type III would be adequate enough to open the frontal sinuses and make them accessible for local medication[12–15]. This is one of the very components of the debate on extent of surgery. Should sound data arise showing that in CRS, a Draf type III gives better control of (frontal sinus) disease, the ACCESS scoring system might need to be adjusted accordingly in the future.

Obviously, the ACCESS score is only an aid to quantify extent of surgery. It is not a goal in itself, so we would not advice to ask for CT scans only to obtain this score, but rather to follow guidelines for requesting CT scans and/or to score scans that are already available. Similarly, the surgeon should tailor the approach of his surgical intervention to the anatomical distribution of disease; a high ACCESS score does not indicate surgery per se.

The scans used for validation of the ACCESS score were obtained from patients with a variety of sinonasal diseases (if any). It is currently most likely that grading the extent of surgery will mainly be applied to patients with CRS. When setting up this study, we were primarily wondering whether the surgical extent of previous procedures could be scored reliably independent of the presence of opacification. We therefore choose not to use CRS cases only, which is a limitation of the study when extrapolating the data to a CRS-only cohort. As the interrater agreement did not differ significantly between scans with high and low opacification, nor between diagnoses, it is likely that the ACCESS score can be applied reliably to this patient group as well. However, most variation of scores arose in the ethmoid and frontal sinuses, especially in the presence of osteitis which is mainly seen in (severe) CRS patients. It will therefore still require further studies to establish the clinical and scientific value of the ACCESS score in CRS, preferably from a random sample of scans.

Conclusion

The ACCESS score is an easy-to-use imaging-based tool to reliably quantify extent of sinus surgery. As such, it might aid clinical decision-making processes, and present researchers the opportunity to uniformly address the issue of extent of sinus surgery.

Acknowledgements

None

Authorship contribution

SR and WF developed the ACCESS scoring system; SR collected all scans, constructed the database, collected and analyzed the data and wrote the manuscript. All other authors graded the scans and reviewed the manuscript.

Conflict of interest

There was no conflict of interest.

References

1. Myller J, Dastidar P, Torkkeli T, Rautiainen M, Toppila-Salmi S. Computed tomography findings after endoscopic sinus surgery with preserving or enlarging maxillary sinus ostium surgery. Rhinology. 2011;49:438-444.
2. Silverman JB, Prasittivatechakool K, Busaba NY. An evidence-based review of endoscopic frontal sinus surgery. Am J Rhinol Allergy. 2009:23:e59-e62.
3. Jankowski R, Pigret D, Decroocq F. Comparison of functional results after ethmoidectomy and nasalization for diffuse and severe nasal polyposis. Acta Otolaryngol. 1997;117:601-608.
4. Alsharif S, Jonstam K, van Zele T, Gevaert P, Holtappels G, Bachert C. Endoscopic Sinus Surgery for Type-2 CRS wNP: An Endotype-Based Retrospective Study. Laryngoscope. 2019;129:1286-1292.
5. Fokkens WJ, Lund VJ, Hopkins C, et al. European Position Paper on Rhinosinusitis and Nasal Polyps 2020. Rhinology. 2020;58(Suppl S29):1-464.
6. Lund VJ, Mackay IS. Staging in rhinosinusitus. Rhinology. 1993;117:S35-S40.
7. Metson R, Gliklich RE, Stankiewicz JA, et al. Comparison of sinus computed tomography staging systems. Otolaryngol - Head Neck Surg. 1997;117:372-379.
8. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med. 2016;15:155-163.
9. Hallgren KA. Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. Tutor Quant Methods Psychol. 2012;8:23-34.
10. Oluwole M, Russell N, Tan L, Gardiner Q, White P. A comparison of computerized tomographic staging systems in chronic sinusitis. Clin Otolaryngol Allied Sci. 1996;21:91-95.
11. Julkunen A, Terna E, Numminen J, et al. Inter-observer agreement of paranasal sinus computed tomography scans. Acta Otolaryngol. 2017;137:611-617.
12. Harvey RJ, Goddard JC, Wise SK, Schlosser RJ. Effects of endoscopic sinus surgery and delivery device on cadaver sinus irrigation. Otolaryngol - Head Neck Surg. 2008;139:137-142.
13. Snidvongs K, Kalish L, Sacks R, Sivasubramaniam R, Cope D, Harvey RJ. Sinus surgery and delivery method influence the effectiveness of topical corticosteroids for chronic rhinosinusitis: Systematic review and meta-analysis. Am J Rhinol Allergy. 2013;27:221-233.
14. Zhao, K; Kim, K; Craig, JR; Palmer J. Using 3D printed sinonasal models to visualize and optimize personalized sinonasal sinus irrigation strategies. Rhinology. 2020; Jun 1;58(3):266-272. .
15. Zhang, L; Zhang, Y; Gao, K; Wang, K; Lou, H; Meng, Y; Wang C. Long-term outcomes of different endoscopic sinus surgery in recurrent chronic rhinosinusitis with nasal polyps and asthma. Rhinology. 2020;58:126-135.

Sietze Reitsma, MD, PhD
Afdeling KNO
Meibergdreef 9
1105 AZ Amsterdam
the NEtherlands

T: +31-20-5663789
E: s.reitsma@amsterdamumc.nl