




# International assessment of inter- and intrarater reliability of the International Frontal Sinus Anatomy Classification system

Ricardo Villarreal, MD<sup>1</sup>, Bozena B. Wrobel, MD<sup>2</sup>, Luis F. Macias-Valle, MD<sup>3</sup> , Greg E. Davis, MD, MPH<sup>4</sup>, Thomas J. Prihoda, PhD<sup>5</sup>, Amber U. Luong, MD, PhD<sup>6</sup>, K. Christopher McMains, MD<sup>7</sup>, Erik K. Weitzel, MD<sup>8</sup>, William C. Yao, MD<sup>6</sup> , Joe Brunworth, MD<sup>9</sup>, David W. Clark, MD<sup>10</sup>, Salil Nair, MBChB, MD<sup>11</sup>, Constanza J. Valdés, MD<sup>12</sup>, Ashleigh Halderman, MD<sup>13</sup>, David W. Jang, MD<sup>14</sup>, Rahuram Sivasubramaniam, BSc, MBBS, MS<sup>15</sup>, Zhipeng Zhang, MD, MPH<sup>2</sup> and Philip G. Chen, MD<sup>1</sup> 

**Background:** Inconsistencies in the nomenclature of structures of the frontal sinus have impeded the development of a validated “reference standard” classification system that surgeons can reliably agree upon. The International Frontal Sinus Anatomy Classification (IFAC) system was developed as a consensus document, based on expert opinion, attempting to address this issue. The purposes of this study are to: establish the reliability of the IFAC as a tool for classifying cells in the frontal recess among an international group of rhinologists; and improve communication and teaching of frontal endoscopic sinus surgery (ESS).

**Methods:** Forty-two computed tomography (CT) scans, each with a marked frontal cell, were reviewed by 15 international fellowship-trained rhinologists. Each marked cell was classified into 1 of 7 categories described in the IFAC, on 2 occasions separated by 2 weeks. Inter- and intrarater reliability were evaluated using Light’s kappa ( $\kappa$ ), the inter-class correlation coefficient (ICC), and simple proportion of agreement.

**Results:** Interrater reliability showed pairwise  $\kappa$  values ranging from 0.7248 to 1.0, with a mean of 0.9162 (SD,

0.0537). The ICC was 0.98. Intrarater reliability showed  $\kappa$  values ranging from 0.8613 to 1.0, with a mean of 0.9407 (SD, 0.0376). The within-rater ICC was 0.98.

**Conclusion:** Among a diverse sample of rhinologists (raters), there was substantial to almost perfect agreement between raters, and among individual raters at different timepoints. The IFAC is a reliable tool for classification of cells in the frontal sinus. Further outcome studies are still needed to determine the validity of the IFAC. Published 2018. This article is a U.S. government work and is in the public domain in the USA.

### Key Words:

sinus anatomy; endoscopic sinus surgery; paranasal sinuses; computed tomography; statistics

### How to Cite this Article:

Villarreal R, Wrobel BB, Macias-Valle LF, et al. International assessment of inter- and intrarater reliability of the International Frontal Sinus Anatomy Classification system. *Int Forum Allergy Rhinol.* 2018;00:1-7.

The frontal sinus is considered the most challenging sinus to address surgically,<sup>1,2</sup> and inadequate

<sup>1</sup>Department of Otolaryngology–Head and Neck Surgery, University of Texas Health San Antonio, San Antonio, TX; <sup>2</sup>Caruso Department of Otolaryngology Head and Neck Surgery, University of Southern California, Keck School of Medicine, Los Angeles, CA; <sup>3</sup>Department of Otolaryngology Head and Neck Surgery, Hospital Español de México, Mexico City, México; <sup>4</sup>Department of Otolaryngology–Head and Neck Surgery, University of Washington, Seattle, WA; <sup>5</sup>Department of Pathology, University of Texas Health San Antonio, Graduate School of Biomedical Sciences, San Antonio, TX; <sup>6</sup>Department of Otorhinolaryngology–Head and Neck Surgery, University of Texas Health Science Center at Houston, Houston, TX; <sup>7</sup>Department of Surgery, South Texas Veterans Health Care System, San Antonio, TX; <sup>8</sup>United States Army Institute of Surgical Research Fort Sam Houston, TX; <sup>9</sup>Department of Otolaryngology–Head and Neck Surgery, Saint Louis University School of Medicine, St Louis, MO; <sup>10</sup>Department of Otolaryngology, Texas A&M University College of Medicine/Baylor

Scott & White Health, Temple TX; <sup>11</sup>Department of Otolaryngology, Auckland Hospitals and University of Auckland, Auckland, New Zealand; <sup>12</sup>Department of Otolaryngology–Head and Neck Surgery, Hospital del Salvador, Universidad de Chile, Santiago, Chile; <sup>13</sup>Department of Otolaryngology Head & Neck Surgery, University of Texas Southwestern Medical Center, Dallas, TX; <sup>14</sup>Division of Head and Neck Surgery & Communication Sciences, Duke University Medical Center, Durham, NC; <sup>15</sup>Department of Otolaryngology, Auckland City Hospital, Auckland, Auckland, New Zealand

Correspondence to: Philip G. Chen, MD, Department of Otolaryngology–Head and Neck Surgery, University of Texas Health San Antonio, 7703 Floyd Curl Drive, MC 7777, San Antonio, TX 78229-3900; e-mail: p\_g\_chen@hotmail.com

Funding sources for the studies: Synedgen (to L.M.-V.) and Olympus (to D.J.).

clearance of the frontal sinus drainage pathway (FSDP) is a frequent cause of failure of endoscopic sinus surgery (ESS) for chronic rhinosinusitis.<sup>3</sup> Frontal sinus surgery requires a detailed understanding of the cellular structure and FSDP<sup>4</sup> unique to each patient, making high-resolution computed tomography (CT) scans an indispensable tool to assess the difficulty of planned sinus surgery.<sup>2</sup> Using multiplanar CT scans, surgeons can develop a 3-dimensional (3D) understanding of the anatomy of the frontal recess and FSDP as well as a surgical plan for each cell.<sup>5</sup> Such an understanding allows for safer and more complete dissection, which will help minimize mucosal trauma, and improve healing in the frontal recess and frontal ostium.<sup>4</sup>

The acceptance of frontal cells as a cause of frontal sinusitis has led to the development of multiple classification systems.<sup>3,6–11</sup> However, the lack of a reference standard for frontal recess anatomy impedes communication among surgeons and trainees.<sup>9</sup> Until recently, the modified Kuhn classification (MKC) has been considered the most user-friendly and comprehensive classification of the anatomic structures of the frontal recess,<sup>4</sup> but others have identified redundancies in the MKC, and additional classification systems have since been proposed.<sup>3</sup> Furthermore, the only study to assess reliability using the MKC to identify frontal cells reported only “reasonable agreement” for identification of the presence of any frontal cells ( $\kappa = 0.392$ ), and an even lower agreement (low-moderate) when attempting to identify individual types of frontal cells ( $\kappa = 0.249$ – $0.3$ ).<sup>12</sup>

The International Frontal Sinus Anatomy Classification (IFAC) was developed as a consensus document based on the expert opinions of leading rhinologists.<sup>11</sup> Yet, some argue that expert opinion can also be fallacious<sup>13</sup> due to selection of the panel, lack of generalizability,<sup>14</sup> and the susceptibility to cognitive biases. For these reasons, it is important to assess the reliability of any proposed new reference standard or classification system.

Our study aimed to evaluate the rate of agreement between a group of international rhinologists (raters) when asked to classify frontal cells on CT scans using the IFAC (intrater reliability). We then evaluated the reproducibility of their responses to assess each rater’s self-consistency when classifying the same scans at a different timepoint (intrater reliability). We also hoped to provide a framework for future outcome studies to truly validate the IFAC as a reference standard for classifying frontal sinus cells.

## Materials and methods

This study was granted institutional review board approval by the University of Texas Health San Antonio (#HSC20170615H). A baseline power analysis (described in what follows) was performed to establish the minimum number of CT scans per cell type, and raters. The repository of CT scans was built by contributions from fellowship-trained rhinologists, including the principal investigator (P.G.C.). Other inclusion criteria were that all participants were >18 years of age and had a high-resolution sinus CT scan (1–2 mm slices) showing minimal mucosal thickening, clearly defined cell borders, minimal opacification, and windowing. Rhinologists from various countries were asked to submit 7 de-identified CT scans that they believed represented each cell type in the IFAC. One hundred CT scans were accumulated. Each scan was initially screened by the principle investigator (PI) to ensure a baseline level of quality defined by the inclusion criteria. Forty-one CT scans were excluded due to evidence of previous sinus surgery, excessive mucosal thickening, or lack of clearly defined cell borders. After this initial screening, all further handling of the remaining 59 CT scans was delegated to a separate member of the team (R.V.). The 59 scans were grouped into their respective IFAC categories and the cell type with the least number of CT scans (ie, 6 scans) defined the maximum number of CT scans per cell type, so that all 7 cell types were equally represented in the final sample. Thus, of the eligible images, 6 scans per cell type were selected, which resulted in the exclusion of 17 more scans, leading to a total of 42 scans. The 17 scans were excluded strategically to ensure that images from all contributors were included in the final sample, as confirmed by one who did not participate in the survey (R.V.). A limit of 4 scans per rhinologist, each of a different cell type, was included in the questionnaire, to avoid overrepresentation of any individual rhinologist. A mean number of  $2.8 \pm 1.1$  scans per rhinologist (range, 1–4) were included in the final sample.

All 42 selected scans were processed by R.V. using MS Paint (Microsoft Corporation, Redmond, WA) to achieve a uniform appearance of all images. This process involved cropping the images to more clearly display the cell of interest, and removing any digital information (ie, time stamps, date, distance marker) found in certain scans. The images were all arranged in the same order (coronal, parasagittal, axial), and the cell of interest was marked with a yellow asterisk (Fig. 1). The scans were not augmented or enhanced in any other way. The labeled images were presented individually in separate pages of a web-based Research Electronic Data Capture (REDCap) survey.<sup>15</sup> The same 15 rhinologists who submitted the images, including the PI, were asked to classify the labeled cell into 1 of the 7 IFAC categories on the standardized images. The scans were presented in an order determined by a random sequence generator. The same scans in a different random order were automatically resent to the raters 14 days after their initial submission. Survey responses were collected

Potential conflict of interest: L.M.-V.: Stryker, consultant. P.C.: Latera Spirox, speaker. G.D.: Intersect ENT and Medtronic, consultant. A.L.: Medtronic, 480 Biomedical, Aerin Medical, and Arrinex, consultant; and ENTvantage, advisory board.

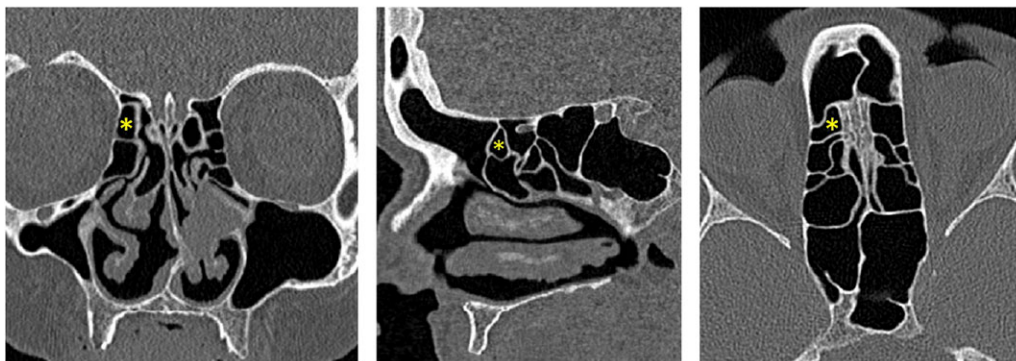
Disclaimer: The views expressed herein are those of the authors and do not reflect the official policy or position of the Brooke Army Medical Center, the United States Army Medical Department, the U.S. Army Office of the Surgeon General, the Department of the Army, the Department of Defense, the Department of Veterans Affairs, or the U.S. Government.

Received: 8 May 2018; Revised: 5 July 2018; Accepted: 31 July 2018

DOI: 10.1002/alr.22200

View this article online at [wileyonlinelibrary.com](http://wileyonlinelibrary.com).

Using the international frontal sinus anatomy classification, how would you classify the cell marked with the yellow asterisk?



- Agger nasi cell
- Supra agger cell
- Supra agger frontal cell
- Supra bulla cell
- Supra bulla frontal cell
- Supraorbital ethmoid cell
- Frontal septal cell

FIGURE 1. Example of how a frontal cell was marked, formatted, and presented in the survey.

and managed in a REDCap database hosted at UT Health San Antonio.

### Power analysis

A baseline power analysis was performed. For the study to have 80% power to detect a pairwise  $\kappa$  of at least 0.5 (moderate strength of agreement)<sup>16</sup> with 4 subjects (CT scans) for each of the 7 IFAC categories, a minimum of 28 total CT scans would be required. Gwet stated that “the coverage rates of the two unconditional variance estimators are all reasonably close to their nominal values of 95% when there are 7 or more raters.”<sup>17</sup> Therefore, a minimum goal of 28 CT scans (4 per category), and 7 raters, was justified.

### Statistics

The inter- and intrarater reliabilities were evaluated using Light’s  $\kappa$ ,<sup>18</sup> which is calculated by computing the mean Cohen’s  $\kappa$  for all possible rater pairs.<sup>18–20</sup> The strength of agreement was interpreted using guidelines established by Landis and Koch,<sup>21</sup> in which  $\kappa$  values <0 had a poor strength of agreement,  $\kappa$  0–0.2 had a slight strength of agreement, and  $\kappa$  0.21–0.4 fair, 0.41–0.6 moderate, 0.61–0.8 substantial, 0.81–1 almost perfect. In addition, intra-class correlation coefficient (ICC)<sup>22</sup> and simple proportion of agreement were calculated. All data analyses were performed using SAS version 9.4 for Windows (SAS Institute, Cary, NC).

## Results

### Interrater reliability

All 15 raters classified all 42 CT scans (fully crossed study design) into 1 of 7 IFAC categories, at 2 timepoints (T1 and T2). The mean time interval between survey submissions was  $17.1 \pm 3.2$  (range, 14–23) days. For T1, the 105 pairwise Cohen’s  $\kappa$  values ranged from 0.7248 to 1.0, with a median of 0.9168 and mean of 0.9162 (Light’s  $\kappa$ ) with a standard deviation (SD) of 0.0537. This indicated very high and significant agreement. The ICC measure of reliability for these 15 raters was 0.98, also very high and very reliable. The tests of symmetry of the little disagreement that occurred pairwise were all nonsignificant, indicating the disagreements that occurred were random and no rater was biased toward any of the 7 categories (1—agger nasi cell, 2—supra agger cell, 3—supra agger frontal cell, 4—supra bulla cell, 5—supra bulla frontal cell, 6—supraorbital ethmoid cell, 7—frontal septal cell). The simple proportion of agreement of the 15 raters for each category was:

1. 5 of 6 100%, 1 of 6 93.3%.
2. 3 of 6 100%, 2 of 6 93.3%, 1 of 6 86.7%.
3. 4 of 6 100%, 1 of 6 93.3%, 1 of 6 80%.
4. 3 of 6 100%, 2 of 6 93.3%, 1 of 6 60%.
5. 4 of 6 100%, 1 of 6 86.7%, 1 of 6 80%.
6. 4 of 6 100%, 2 of 6 86.7%.
7. 6 of 6 100%.

Overall, in 29 of 42 items all 15 raters agreed. In 6 of 42 items, 93.3% of the raters agreed; in 4 of 42 items, 86.7% of the raters agreed; in 2 of 42 items, 80% of the raters agreed; and in 1 of 42 items, 60% of the raters agreed.

### Intrarater reliability

The  $\kappa$  value ranged from 0.8613 to 1.0 with a median of 0.9444 and mean of 0.9407 (SD, 0.0376), indicating strong reproducibility and agreement for all raters. The ICC within raters was 0.98, indicating strong reliability of raters over time for identifying the different cell types.

### Discussion

Endoscopic sinus surgery requires detailed knowledge of the anatomy of the internal nose and paranasal sinuses. Schaeffer first developed the term “frontal cell” to describe accessory frontal sinus cells.<sup>6</sup> Van Alyea defined “frontal cells” as air cells that encroach on the frontal recess or extend into the frontal sinus.<sup>23</sup> The configuration of these air cells within the frontal sinus and the varying attachments of the uncinate process affect the FSDP.<sup>10</sup> These frontal cells can also lead to obstruction and clinically significant frontal sinusitis.<sup>6</sup>

Historically, there has been considerable confusion concerning the classification of cells in the frontal recess. The clinical importance of these cells was often underestimated, or their presence was overlooked completely.<sup>23</sup> The widespread use of endoscopes and high-resolution CT scans, and the acceptance of frontal cells as a potential cause of frontal sinusitis,<sup>6</sup> led to the development of many classification systems.<sup>3,6-11</sup> Bent and Kuhn distinguished frontal cells from other anterior ethmoid cells and classified them into 4 frontal cell types (types 1–4 or K1-K4).<sup>6</sup> Kuhn further described pneumatization patterns found in the frontal recess, which led to the Kuhn classification of frontoethmoid cells.<sup>7</sup> Wormald and Chan proposed the modified Kuhn classification (MKC), which distinguishes Kuhn type 3 cells (K3) from type 4 cells (K4) (Table 1). This distinction between K3 and K4 cells in the MKC was clinically important because it was found that K3 cells can usually be addressed endoscopically from below, whereas K4 cells typically require a combined approach, or an endoscopic modified Lathrop procedure.<sup>4,8</sup> Using high-resolution CT scans with triplanar scrolling, Lee et al adjusted the criteria for frontal recess pneumatization patterns to better suit the needs of contemporary ESS. Lee et al also characterized the supra bulla cell (SBC) and frontal bullar cell (FBC), which had been poorly described.<sup>9</sup> Pianta et al argued that the MKC was redundant and proposed the agger bullar classification (ABC) of the FSDP, based on a compartmental rather than morphologic criterion.<sup>3</sup> Wormald et al proposed the IFAC (Table 2) with the goal of providing a more precise nomenclature based on the position of cells. This would allow a more complete picture of the anatomy to be established, easier communication between surgeons, and more accurate teaching of frontal ESS.<sup>11</sup>

To our knowledge, this is the first study of this magnitude to assess both the inter- and intrarater reliability of the IFAC as a tool for classifying frontal cells by a diverse group of rhinologists, using various reliability statistical methods.

In a recent publication, Choby et al evaluated 100 scans among a small group of raters to establish an ICC for the IFAC system.<sup>24</sup> The results showed a favorable interrater reliability (median ICC, 0.80). However, the raters who evaluated the scans in the study consisted of only 3 rhinologists, all from the same center. In the present study, an international group of 15 rhinologists reviewed the images. In addition, our methodology of marking each individual cell of interest ensures each rater is referring to the same structure, which is important in a study of reliability.

We found a very high level of agreement using various statistical parameters among a diverse group of fellowship-trained rhinologists, and very high reproducibility of the results. These results provide compelling evidence that the IFAC is a reliable tool for rhinologists to use for classifying cells in the frontal sinus and frontal recess. In addition, by including a diverse group of international rhinologists to both contribute and rate images, we aimed to increase the true variance, which would theoretically increase the reliability.<sup>25</sup> We did not have raters in our study classify frontal cells using the MKC, so we did not measure the interrater reliability of the MKC. Instead, we conducted a literature review to find any existing data about interrater reliability using the MKC. We only found 1 study, by Langille et al, that assessed interrater reliability using the MKC to identify frontal cells. They reported only “reasonable agreement” for identification of the presence of any frontal cells ( $\kappa = 0.392$ ), and even lower agreement (low-moderate) when attempting to identify individual types of frontal cells ( $\kappa = 0.249-0.3$ ).<sup>12</sup> If we compare the results obtained from our study, raters who used the IFAC demonstrated superior agreement over raters who used the MKC to classify frontal cells ( $\kappa = 0.7248-1.0$  vs  $\kappa = 0.249-0.3$ ). However, direct comparison between our reliability study and the reliability study performed by Langille et al may be misleading as raters were not fellowship-trained rhinologists.

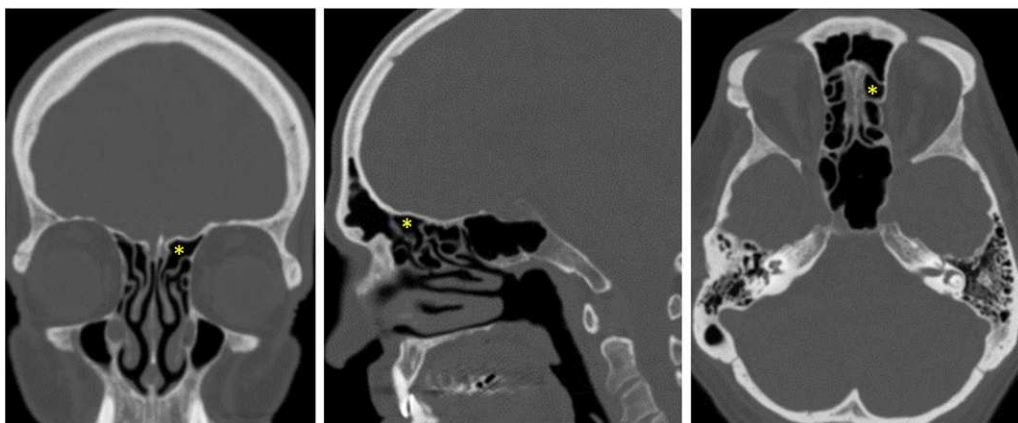
In the vast majority of images (41 of 42),  $\geq 80\%$  of raters agreed on the classification of the marked cell. However, in 1 CT scan with a marked SBC (Fig. 2), only 60% of raters (9 of 15) classified the cell as an SBC. The other 6 raters classified it as a supra bulla frontal cell (SBFC). When the identical image was presented in the second survey, 66.7% of raters (10 of 15) classified it as an SBC, 4 raters classified it as an SBFC, and 1 as a supra agger frontal cell (SAFC). The images of both surveys were reexamined to exclude an error in the creation of the survey. The fact that the agreement was substantially higher when classifying the other 5 examples of SBCs suggests that the low agreement was likely due to a property of the selected image, rather than a problem with the proposed definition in the IFAC. The SBC is classically described as being superior to the bulla ethmoidalis, with its superior wall being the base of the skull, and its anterior border not crossing the frontal ostium into the frontal sinus.<sup>9,11</sup> The SBFC is essentially an SBC that has pneumatized through the frontal ostium into the frontal sinus, and now the skull base forms its posterior

**TABLE 1.** Modified Kuhn classification of frontal ethmoidal cells

ANC (Agger nasi cell)—Usually a single cell anterior to the middle turbinate  
 SOEC (supraorbital ethmoid cell)—SBC protruding into orbital roof  
 Frontoethmoidal cells (Bent and Kuhn frontal cells)—cells close or within frontal process of maxillary bone and above the ANC  
 K1—Single frontal recess cell above ANC  
 K2—Tier of cells in frontal recess above ANC  
 K3—Single massive cell pneumatizing cephalad into frontal sinus occupying <50% of vertical sinus height  
 K4—Single isolated cell within the frontal sinus occupying >50% of vertical sinus height  
 SBC (supra bulla cell) —A cell or cells above the bulla ethmoidalis  
 Frontal bulla cell (FBC)—SBC protruding into frontal sinus  
 Interfrontal sinus septal cell (ISSC)—Cell resulting from pneumatization of interfrontal sinus septum pushing FSDP laterally and narrowing the frontal ostium

**TABLE 2.** International Frontal Sinus Anatomy Classification

	Abbreviation	Cell name	Definition
Anterior cells—push FSDP medial, posterior, or posteromedially	ANC	Agger nasi cell	Anterior to origin of middle turbinate or above most anterior insertion of middle turbinate into lateral nasal wall.
	SAC	Supra agger cell	Anterior-lateral ethmoidal cell above ANC that does not pneumatize into frontal sinus.
	SAFC	Supra agger frontal cell	SAC that extends into frontal sinus (wide range of sizes from floor up to roof of frontal sinus).
Posterior cells—push FSDP anteriorly	SBC	Supra bulla cell	Above bulla ethmoidalis; does not enter frontal sinus.
	SBFC	Supra bulla frontal cell	SBC that pneumatizes along skull base to posterior frontal sinus. Skull base forms posterior wall of cell.
	SOEC	Supraorbital ethmoid cell	Anterior ethmoid cell pneumatizing around, anterior, or posterior to anterior ethmoidal artery over roof of orbit. Often forms part of posterior wall of extensively pneumatized frontal sinus separated only by a bony septation.
Medial cells—push FSDP laterally	FSC	Frontal septal cell	Medially based cell of anterior ethmoid or inferior frontal sinus; attached or located in interfrontal sinus septum; associated with medial aspect of frontal sinus outflow tract, pushing FSDP laterally and frequently posteriorly.



**FIGURE 2.** Computed tomography scan with the lowest percentage of agreement among raters.

wall.<sup>11</sup> In the parasagittal view of the CT scan (Fig. 2), the anterior wall of the marked cell appears to extend exactly to the level of the frontal ostium, but it is unclear whether it actually crosses the frontal ostium into the frontal sinus.


It should be emphasized that, although our study demonstrates reliability of the IFAC, it does not assess its validity. A comprehensive validation process includes both internal and external validation strategies that can evaluate the accuracy, generalizability, reproducibility, and ideally the clinical effectiveness of a new reference standard.<sup>26</sup> Because there is no true “gold standard” established for the classification of frontal cells, validation would require a significant effort, and it may not even be possible to demonstrate clinical effectiveness.<sup>26</sup> The intended purpose of the IFAC, according to the original authors, was to establish a more complete picture of the anatomy, easier communication between surgeons, and more accurate teaching of the necessary surgical steps in frontal ESS.<sup>11</sup> Radiologic classification systems in other specialties have been validated by comparing the predicted classification to intraoperative findings (ie, the gold standard). Because the frontal recess is such a small area, it is unclear whether intraoperative identification of cells would be feasible as a gold standard. For this reason, we propose that validation of the IFAC could be better studied by comparing outcomes of ESS using the IFAC vs other classification systems. Another goal of the IFAC was to provide more accurate teaching of ESS, so trainee proficiency evaluations could, theoretically, also be useful for assessing validity.

Generalizability is a measure of the degree to which a sample truly represents the target population of study. The diversity in training, location, individual practices of the raters, and the variety of patients they treat contributed to a fair representation of a random population of rhinologists. Having such a diverse group of rhinologists also helped to eliminate the selection bias that would arise from having all scans chosen by 1 person.

The most glaring limitation of our study is the possibility that a rater recognized CT scans they submitted, which

would falsely elevate the level of agreement. Several measures were taken to minimize this factor, including formatting of CT scans by a different team member (R.V., as described earlier) to provide a uniform appearance of all images, and minimizing overrepresentation of CT scans from any individual rhinologist (also described earlier). In addition, all raters were blinded to the number of CT scans per cell type in the questionnaire, which of the CT scans they submitted would be used in the questionnaire, and how the images were processed. Because the PI was involved in initial screening of all images to ensure a baseline level of quality, and also a subsequent rater, all further handling of images (ie, processing, selection of images included in the survey, development of the survey) was delegated to another team member (R.V.) to minimize the chance that the PI would recognize a CT scan, which would also falsely elevate the level of agreement. There was also a difference in the time interval between survey submissions (range, 14–23 days). However, we do not believe the time interval difference between submissions significantly influenced rater responses. Finally, given our static images, the raters were not able to scroll through the images in a triplanar manner, as they would in real-world conditions. However, this ensured all raters were viewing the exact same image and were referring to the same cell.

## Conclusion

A detailed understanding of the 3-dimensional anatomy of the frontal recess is extremely important when planning frontal ESS. The IFAC was developed as a consensus document based on the expert opinion of rhinologists from around the world. This study has indicated that, among a diverse sample of rhinologists, there was substantial to almost perfect agreement between different raters, and among individual raters at different timepoints. The IFAC is a very reliable tool for identifying frontal cells, but further outcome studies are still needed to assess its validity. 

## References

1. Wormald PJ. Surgery of the frontal recess and frontal sinus. *Rhinology*. 2005;43:82–85.
2. Wormald PJ, Bassiouni A, Callejas CA, et al. The international classification of the radiological complexity (ICC) of frontal recess and frontal sinus. *Int Forum Allergy Rhinol*. 2017;7:332–337.
3. Pianta L, Ferrari M, Schreiber A, et al. Agger-bullar classification (ABC) of the frontal sinus drainage pathway: validation in a preclinical setting. *Int Forum Allergy Rhinol*. 2016;6:981–989.
4. Wormald PJ. Three dimensional building block approach to understanding the anatomy of the frontal recess and frontal sinus. *Oper Techn Otolaryngol Head Neck Surg*. 2006;17:2–5.
5. Chen PG, McMains KC, Tewfik MA, et al. Teaching frontal sinus anatomy using the frontal sinus masterclass 3-D conceptualization model. *Laryngoscope*. 2018;128:1294–1298.
6. Bent JP, Cuijly-Siller C, Kuhn FA. The frontal cell as a cause of frontal sinus obstruction. *Am J Rhinol*. 1994;8:185–191.
7. Kuhn FA. Chronic frontal sinusitis: the endoscopic frontal recess approach. *Oper Techn Otolaryngol Head Neck Surg*. 1996;7:222–229.
8. Wormald PJ, Chan SZ. Surgical techniques for the removal of frontal recess cells obstructing the frontal ostium. *Am J Rhinol*. 2003;17:221–226.
9. Lee WT, Kuhn FA, Citardi MJ. 3D computed tomographic analysis of frontal recess anatomy in patients without frontal sinusitis. *Otolaryngol Head Neck Surg*. 2004;131:164–173.
10. Lund VJ, Stammberger H, Fokkens WJ et al. European Position Paper on the Anatomical Terminology of the Internal Nose and Paranasal Sinuses. *Rhinology*. 2014;24:1–34.
11. Wormald PJ, Hoseman W, Callejas C et al. The International Frontal Sinus Anatomy Classification (IFAC) and Classification of the Extent of Endoscopic Frontal Sinus Surgery (EFSS). *Int Forum Allergy Rhinol*. 2016;6:677–696.
12. Langille M, Walters E, Dziegielewski PT, et al. Frontal sinus cells: identification, prevalence, and association with frontal sinus mucosal thickening. *Am J Rhinol Allergy*. 2012;26:e107–10.
13. Walton D, Koszowy M. Arguments from authority and expert opinion in computational argumentation systems. *AI Soc*. 2017;32:483–496.
14. James D, Warren-Forward H. Research methods for formal consensus development. *Nurse Res*. 2015;22:35–40.
15. Harris PA, Taylor R, Thielke R, et al. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42:377–381.
16. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther*. 2005;85:257–268.
17. Gwet KL. Variance estimation of nominal-scale interrater reliability with random selection of raters. *Psychometrika*. 2008;73:407.
18. Light RJ. Measures of response agreement for qualitative data: some generalizations and alternatives. *Psychol Bull*. 1971;76:365–377.
19. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20:37–46.

20. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. *Tutor Quant Methods Psychol.* 2012;8:23.
21. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;1:159–174.
22. Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educ Psychol Meas.* 1973;33:613–619.
23. Van Alyea OE. Frontal cells: an anatomic study of these cells with consideration of their clinical significance. *Arch Otolaryngol.* 1941;34:11–23.
24. Choby G, Thamboo A, Won TB, et al. Computed tomography analysis of frontal cell prevalence according to the International Frontal Sinus Anatomy Classification. *Int Forum Allergy Rhinol.* 2018;00:1–6.
25. Streiner DL, Norman GR. "Precision" and "accuracy": two terms that are neither. *J Clin Epidemiol.* 2006;59:327–330.
26. Gold R, Reichman M, Greenberg E, et al. Developing a new reference standard: is validation necessary? *Acad Radiol.* 2010;17:1079–1082.