A comparative study of two grading systems for epistaxis in hereditary haemorrhagic telangiectasia*

O.J. Jorgensen1, J. Steineger1, G. Bachmann-Harildstad2, S. Dheyauldeen1

1 Department of Otohinolaryngology & Head and Neck Surgery, Oslo University Hospital, Medical Faculty, University of Oslo, Oslo, Norway
2 Department of Otohinolaryngology, Akershus University Hospital and Institute of Clinical Medicine, Akershus University Hospital, Oslo University, Lørenskog, Norway

Abstract

Background: Different institutions use different grading systems for hereditary haemorrhagic telangiectasia (HHT)-associated epistaxis. It is important to have a universal, standardized grading system to compare and evaluate the effectiveness of different treatment options. We introduced the “Intensity, Frequency and need for Blood Transfusion” (IFT) grading system for HHT-associated epistaxis in 2008. Hoag et al. proposed the “Epistaxis Severity Score” (ESS) for the International HHT foundation in 2010. This study aimed to evaluate the potential correlation between the ESS and IFT grading systems.

Methods: The study included 354 simultaneous reports using the IFT and ESS from 106 patients. The correlation between the ESS, IFT and haemoglobin levels was measured using Pearson’s correlation coefficient. The ESS and IFT were scored simultaneously by the patient and doctor in 48 cases to evaluate if there was a discrepancy in the scoring applied by either set of responders.

Results: The measured correlation between the two grading systems was good (0.75). The grade of epistaxis reported by patients and doctors respectively showed no significant difference. Both the IFT and ESS grading systems correlate significantly to the haemoglobin level.

Conclusions: Both the IFT and ESS scores correlate to each other, and their results are comparable. Whether the IFT or ESS scoring was performed by the patient or doctor had no significant impact.

Key words: Epistaxis grading, epistaxis severity score (ESS), hereditary haemorrhagic telangiectasia, epistaxis intensity, frequency and need for blood transfusion score (IFT)

Introduction

Hereditary haemorrhagic telangiectasia (HHT, also known as Rendu-Osler-Weber syndrome) is a rare autosomal dominant inherited vascular disorder (prevalence of 1:5000 to 8000) (1,2) that results in vascular malformations in the form of mucocutaneous telangiectasias, and occasionally arteriovenous malformations (AVMs) in visceral organs (e.g., pulmonary, hepatic, cerebral, gastrointestinal, or spinal AVMs). The presence of these vascular malformations can cause a variety of clinical signs and symptoms of which epistaxis is the most common (3).

A diagnosis of HHT is based on the presence of four criteria, known as the “Curaçao criteria”: recurrent spontaneous epistaxis, mucocutaneous telangiectasias, visceral AVM(s) and a first-degree relative with the HHT diagnosis. The diagnosis of HHT is definite if three or four criteria are present, possible if two criteria are present and unlikely if only one criterion or none is present (4). Genetic tests are available for individuals who do not fully meet the clinical diagnostic criteria. However, genetic testing should be offered to all patients, even if they meet the clinical diagnostic criteria, as it will allow for predictive testing in their immediate and potentially extended families.

The most common symptom reported in patients with HHT is spontaneous and recurrent epistaxis due to fragile telangiectasias in the nasal mucosal membranes (5). Eighty to 100% of HHT patients eventually suffer from epistaxis (5).
The frequency and severity of epistaxis in HHT patients varies considerably, even within the same affected family. It ranges from mild and infrequent with little impact on daily life to multiple and severe episodes of epistaxis with debilitating consequences. Some patients require repeated blood transfusions, frequent surgical interventions, and hospitalisation. Patients with HHT report epistaxis to be the most debilitating symptom, with the greatest impact on quality of life (QoL).

Many treatment options are described for HHT-associated epistaxis. Management of epistaxis associated with HHT varies depending on the grade of severity. Multiple grading systems have been used to evaluate the severity of epistaxis in patients with HHT. This makes it difficult to assess and compare the effectiveness of different treatment options. In 2008 a new grading system was published as a proposal for a commonly accepted system. This is the Epistaxis Intensity, Frequency, and need for Blood Transfusion score (IFT). It is mainly based on the opinions of experts who have published research on epistaxis in HHT. In 2010 another system was published. This was the Epistaxis Severity Score (ESS). ESS was created based on patient feedback about independently associated risk factors for self-reported epistaxis severity. The two grading systems differ regarding the specific period of observation (three months in the case of ESS and four weeks in the IFT classification). While ESS focuses on the duration of the nosebleed episode, IFT focuses on the amount of blood loss. Another aspect is the method used to calculate the results. It is quite hard to calculate ESS manually, and a computer is needed for this purpose. In IFT, the results can easily be calculated manually using a simple calculation table.

In our previous studies with fewer patients, we observed a close relationship between the IFT and ESS. However, the correlation between the IFT and ESS grading systems has not been systematically examined. Both grading systems are used in HHT research but comparing published research results is a challenge due to the unknown correlation. For this reason, we often perform both IFT and ESS scoring concurrently to reflect the epistaxis severity more accurately and to compare the results. However, this practice is time-consuming and burdens the HHT patient with multiple questions.

This study aimed to evaluate the correlation between the IFT and ESS grading systems. A strong correlation would help to compare and evaluate the effectiveness of interventions for HHT-associated epistaxis, regardless of the chosen grading system. Furthermore, we investigated which grading system correlates best with haemoglobin (Hgb) levels. In our clinic, both the IFT and ESS are often scored by the doctor. Thus, in the current study we also evaluated the potential difference in scoring from results provided by both the doctor and patient.

**Methods**

All patients included in this study are part of a research database for HHT at Oslo University Hospital, Rikshospitalet. This database includes HHT patients, diagnosed clinically according to Curacao criteria 3-4, and/or by additional genetic testing in the case of Curacao 1-2. The patients included in this study were previously included in other studies concerning HHT-associated epistaxis and we collected IFT and ESS scores simultaneously. In these previous studies, the IFT and ESS scoring was performed either by the patients themselves or by the responsible doctor. However, in forty-eight cases, the doctors and the patients registered IFT and ESS simultaneously. The collection of the Hgb samples was performed in conjunction with outpatient controls and treatment.

Epistaxis intensity, frequency and need for blood transfusion score (IFT)

The IFT was proposed by Al-deen (Dheuaydeen) and Bachmann-Harildstad in 2008, as a grading system for HHT-associated epistaxis. The IFT is composed of three scales.

This multi-scale system focuses on a definite period of observation, which is one month (four weeks). The system gives a quick overview of the different aspects of epistaxis in the HHT patient. It can easily be converted into a single-scale system in order to compare with other single-scale grading systems, such as the ESS. The IFT system uses the abbreviation (I) for the intensity of bleeding, (F) for frequency, and (T) for blood transfusion. Our experience is that patients can usually only remember the two most frequent bleeding intensities in a given period, and this is why we only include the two most frequent bleeding intensities, instead of more than two. The digits corresponding to the appropriate intensity and frequency are added. For example, a patient has experienced epistaxis every other day during the last month (e.g., about 15 times) with some slight stains on the handkerchief. In addition, the patient had two episodes of severe epistaxis, filling a bowl, and needed a blood transfusion on one occasion. The epistaxis grading for this patient would be ISF4T1 (I1+4F3+1T1) (Table 1).

In order to compare the IFT score with the ESS scale, one must convert the IFT to a single-scale system. This is done by multiplying the digit representing intensity (I) by the corresponding digit for frequency (F) and adding them together, and then adding the digit for blood transfusion (T). This gives a grading scale from zero to 30. Zero is referred to as “no bleeding,” 1-5 as “mild bleeding,” 6-10 as “moderate,” 11-15 as “severe,” and 16-30 as “intractable bleeding.” Applying this to the above-mentioned example, the single-scale score for the patient grading is 8 and calculated as follows: (1x3) + (4x1) +1 = 8.
Epistaxis severity score (ESS)

In 2010, Hoag et al. (11) published a proposal for a standardised system to measure epistaxis severity with the Epistaxis Severity Score (ESS). The purpose was to determine factors associated with patient-reported epistaxis severity in order to develop a severity score (11). An issue raised by the authors was that the former proposed grading systems were not statistically validated.

ESS was created based on patient feedback in the form of an electronic survey where the patients were asked about independently associated risk factors for self-reported epistaxis severity (11). Six factors were identified as having the highest correlation to epistaxis severity: intensity, frequency, duration, need for medical attention, anaemia, and need for transfusion. The ESS was developed based on these factors (Table 3).

Computer software is used to calculate the results. Each response is weighted by their respective coefficient (two dismal digits). These results are then added to create a raw ESS, which is then divided by the range of the raw score and multiplied by 10 to give the normalised ESS within a range of 0 (no epistaxis) to 10 (most severe epistaxis (mild 1-4, moderate 5-7 and severe 8-10)) (Table 4) (11).

Statistical analysis

All statistical calculations were performed using SPSS-22. The T-test was used to compare means and test for discrepancy. Data were analysed using Pearson’s correlation coefficient (PCC), P-value, and Cohen’s kappa coefficient. Kappa coefficient is recommended to determine relative agreement between evaluators for nominal and categorical data, as it eliminates the effect of expected agreement at random. Agreement level is graded as follows (Altman) (17): poor (K ≤ 0.20), moderate (K ≤ 0.20 to 0.40), fair (K ≤ 0.40 to 0.60), good (K ≤ 0.60 to 0.80) or very good (K ≤ 0.80 to 1.00).
Results
354 simultaneous reports using the IFT and ESS from 106 patients were included in the study. A total of 316 Hgb samples were collected. Table 5 shows the patient characteristics, mean scores and range of the ESS, IFT, and Hgb levels.

Table 5. Demographic characters of the included patients, mean and range.

<table>
<thead>
<tr>
<th>Total number of patients</th>
<th>106</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Females: 48 (45.3%)</td>
<td>Males: 58 (54.7%)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>56.2 (±12.4) years</td>
<td>21-85 years</td>
</tr>
<tr>
<td>ESS</td>
<td></td>
</tr>
<tr>
<td>4.15 (±1.93)</td>
<td>0-9.09</td>
</tr>
<tr>
<td>IFT</td>
<td></td>
</tr>
<tr>
<td>8.04 (±4.65)</td>
<td>0-28.00</td>
</tr>
<tr>
<td>Hgb</td>
<td></td>
</tr>
<tr>
<td>12.60 (±1.87)</td>
<td></td>
</tr>
</tbody>
</table>

ESS= Epistaxis Severity Score; IFT= Epistaxis intensity, frequency and need for blood transfusion score.

Table 6. Pearson’s correlation coefficient between IFT, ESS and Hgb.

<table>
<thead>
<tr>
<th>IFT</th>
<th>ESS</th>
<th>Hb</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.000</td>
<td>0.750**</td>
<td>-0.288**</td>
</tr>
<tr>
<td>0.750**</td>
<td>1.000</td>
<td>-0.358**</td>
</tr>
<tr>
<td>-0.288**</td>
<td>-0.358**</td>
<td>1000</td>
</tr>
</tbody>
</table>

ESS= Epistaxis Severity Score; IFT= Epistaxis intensity, frequency and need for blood transfusion score. ** Correlation is significant at the 0.01 level. IFT and ESS are adequate to strongly correlated with a correlation coefficient of 0.750 and p-value ≤0.001. Both IFT and ESS are correlated to the Hgb level with a correlation coefficient of -0.288 and -0.358 respectively and p-value ≤0.001.

Discussion
The purpose of this study was to establish the potential correlation between the ESS and IFT. Until a universal, standardised grading system for epistaxis is accepted, knowing that the ESS and IFT grading systems correlate will help researchers and care providers to compare and evaluate the effectiveness of interventions for HHT-associated epistaxis, regardless of the chosen grading system.

Interpretation of correlation coefficients is not straightforward, and for health research, values <0.60 may indicate inadequate agreement [13]. The results of our study show a good correlation (0.75) between the ESS and IFT. This finding leads us to conclude...
that research results regarding the severity of HHT-associated epistaxis measured with the ESS and IFT are comparable. The low correlation for questions regarding blood transfusion (ESS 6 and IFT 5) may be explained by the fact that the question in ESS 6 focuses on transfusions specifically given in relation to nose bleedings. The question in the IFT score incorporates all transfusions given, regardless of the site of bleeding.

The measured correlation between the ESS and IFT with Hgb levels was moderate but statistically significant. We observed that as the number of patient evaluations and Hgb samples increased, their correlation decreased. Hgb levels could be affected by many factors other than epistaxis alone, e.g., blood transfusion, iron supplementation, and gastrointestinal bleeding. Hgb levels may not, for these reasons, be a good measure of epistaxis severity.

In the usual clinical setting, doctors grade patient epistaxis severity. However, in certain conditions (for instance, in the case of post-therapy follow-up), it is necessary that the patient report the epistaxis grade. It has been a subject of discussion whether a discrepancy occurs between the reporting of the patient and that of the doctor. In this study, simultaneous registration by the patient and doctor using both the IFT and ESS grading systems was performed in 48 cases. The results showed no significant difference between the score registered by the patient or doctor. This leads us to conclude that both the doctor and patient have a very similar perception of both the IFT and ESS systems, and that their reports are reliable.

The successful treatment of HHT-associated epistaxis is dependent on the ability to evaluate and score the epistaxis severity accurately. The severity of epistaxis in HHT patients varies considerably from patient to patient. Although a patient may initially respond to a specific treatment, the condition may progress, and new evaluations are warranted. An epistaxis score is also essential in therapeutic research, as an outcome measure. As new treatment modalities are introduced, proper epistaxis grading is required to assess their effectiveness and treatment response.

The past decades have seen several proposed grading systems for HHT-associated epistaxis. Some of the earliest systems were too subjective or focused only on one or a few aspects of the bleeding episodes, e.g., only bleeding duration, and were not sufficient to measure treatment response or be used in clinical research. The IFT and ESS grading systems consider multiple aspects of the bleeding episodes and the need for transfusion within a certain period of time. Because different institutions use different grading systems, comparing and evaluating research results as well as the effectiveness of treatment is challenging, if not impossible. In our clinic, we have chosen to record the epistaxis severity using both the IFT and ESS grading systems. This allowed us to compare the results from each grading system and calculate any correlation.

Although the ESS and IFT are now commonly used systems to grade epistaxis severity in patients with HHT, critics have raised some important concerns. In the ESS, the bleeding data is self-reported and may be subject to bias due to patient exaggeration or under-reporting of the frequency or severity of symptoms. The patient population who participated in the survey consisted mainly of North American Caucasians. Their subjective perception of epistaxis episodes may not necessarily be representative of other ethnicities or countries, where the respective healthcare systems may lack the ability to provide similar levels of care. Also, the period observed in ESS (three months) may be too long, possibly resulting in recall bias and perhaps lacking the required sensitivity to capture small, but potentially important changes in HHT-related epistaxis following treatment. ESS was, however, found to be a significant predictor for invasiveness as patients with higher ESS score had a much greater risk of requiring surgical treatment for their epistaxis. In addition,

---

Table 7. Pearson’s correlation coefficient between the I, F, and T components of the IFT scale with ESS1, ESS2, and ESS3 components of the ESS scale.

<table>
<thead>
<tr>
<th>I</th>
<th>F</th>
<th>T</th>
<th>ESS1</th>
<th>ESS2</th>
<th>ESS3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.364**</td>
<td>0.293**</td>
<td>0.311**</td>
<td>0.389**</td>
<td>0.188**</td>
</tr>
</tbody>
</table>
| **Correlation is significant at the 0.01 level. * Correlation is significant at the 0.05 level.**

<table>
<thead>
<tr>
<th>Patient</th>
<th>Doctor</th>
<th>Patient</th>
<th>Doctor</th>
</tr>
</thead>
<tbody>
<tr>
<td>IFT</td>
<td>4.74 (±1.79)</td>
<td>10.04 (±5.86)</td>
<td>9.38 (±4.76)</td>
</tr>
<tr>
<td>ESS</td>
<td>4.71 (±1.69)</td>
<td>10.04 (±5.86)</td>
<td>9.38 (±4.76)</td>
</tr>
</tbody>
</table>

No significant difference between the grade of epistaxis scored by the patient and the doctor simultaneously.

ESS= Epistaxis Severity Score; IFT= Epistaxis intensity, frequency and need for blood transfusion score.
the minimal important difference of the ESS was demonstrated to be 0.71 in a previous study [23].

The IFT score focuses on a shorter observation period, four weeks, which may reduce the possibility of recall bias [24]. The authors of both the ESS and IFT [10,11] have noted that the ease of calculating the grading results is an important issue. IFT is, in this regard, an easier grading system to use. However, a limitation of the IFT is the preliminary lack of systemic validation. Another possible weak point in the IFT is that T (transfusion) is not limited to transfusions received because of anaemia due to epistaxis alone. However, T is calculated by addition, not multiplication, and therefore is not as heavily weighted as I and F. As noted in a recent study by Pagella et al. comparing the Frequency, Intensity and Duration score (FID) with the ESS [25], intravenous iron supplementation is now more often used in treating HHT-related anaemia than blood transfusion, a method that was more common around the time the IFT was created. We acknowledge this important fact.

During the period the data in our study were collected, there was no limit to the observation period for the ESS; in a later revision it was limited to 3 months. The observation period for the IFT was limited to one month from the start. This is a possible source of error as patients may have reported receiving blood transfusions in the ESS outside the later revised observation period of 3 months.

Patients with HHT-related epistaxis often suffer from different intensities of nosebleed during an observation period. For example, a patient may experience a strong but short-lasting nosebleed, a weak but long-lasting nosebleed and a weak but short-lasting nosebleed in the same observation period. The ESS and FID only allows for the most frequent type of bleeding episode to be registered. The IFT allows for the two most frequent types of nosebleeds to be documented and this may increase the precision of the epistaxis grading. None of the existing grading systems considers the patient’s own attempt(s) at stopping the nosebleeds, which in some patients may be quite effective and have an impact on the bleeding duration. The recorded duration is of little use if the patient stops the nosebleed “prematurely”. Patients also have difficulty in noting the exact time of onset and cessation of the nosebleeds. These are the main reasons why we chose to use bleeding intensity instead of duration in the IFT. We found, however, that there was a fair correlation between ESS2 (duration) and I (intensity).

We believe that by further supplementing a grading system with factors such as QoL, and potential biomarkers, a more precise measure of epistaxis severity may be achieved. This could be a subject for further inquiry.

Conclusion
The IFT and ESS scores correlate with each other. Whether the IFT or ESS scoring was performed by the doctor or patient had no significant influence. Further prospective studies, with a larger sample size and including other parameters such as QoL, are required.

Acknowledgement
None. No mentionable funding has been provided to support this study.

Authorship contribution
OJJ: Contributions to the conception and design of the work. Interpretation of data for the work. Drafting the majority of the version to be published. JS Substantial contributions to the conception/design of the work and the acquisition, analysis, and interpretation of data for the work. Revising the work critically for important intellectual content. Final approval of the version to be published. GBH: Revising the work critically for important intellectual content. Final approval of the version to be published. SD Substantial contributions to the conception/design of the work and the acquisition, analysis, and interpretation of data for the work. Revising the work critically for important intellectual content. Final approval of the version to be published.

Conflict of interest
None of the authors have any conflicts of interest to disclaim.

References
7. Lennox PA, Hitchings AE, Lund VJ, Howard DJ. The SF-36 health status questionnaire in assessing patients with epistaxis secondary