Pediatric gastroesophageal reflux and upper gastrointestinal tract motility: the use of multichannel intraluminal impedance and high resolution manometry

van Wijk, M.P.

Citation for published version (APA):
van Wijk, M. P. (2010). Pediatric gastroesophageal reflux and upper gastrointestinal tract motility: the use of multichannel intraluminal impedance and high resolution manometry

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
chapter 2

Optimization of the reflux-symptom association statistics for use in infants being investigated by 24 hour pH-impedance

Taher Omari
Andre Schwarzer
Michiel van Wijk
Lisa McCall
Stamatiki Kritas
Sibylle Koletzko
Geoffrey Davidson

Submitted for publication
ABSTRACT

Introduction: pH-impedance monitoring is used to diagnose symptomatic gastroesophageal reflux (GER) based upon symptom association probability (SAP). Current criteria for calculation of SAP are optimized for heartburn in adults. Infants, however, demonstrate a different symptom profile. The aim of this study was to optimize criteria for calculation of SAP in infants with GER disease.

Methods: Ten infants referred for investigation of symptomatic reflux were enrolled. GER episodes were recorded using a pH-impedance probe which remained in place for 48 hours. During the test, cough, crying and regurgitation were marked. Impedance recordings were analyzed for the occurrence of bolus reflux episodes. SAP was separately calculated for Day 1 and Day 2 using automated reporting software which enabled the time window used for SAP calculations to be modified from 15-600 seconds. Day to day agreement of SAP was assessed by calculating the 95% limits of agreement (mean difference ± 1.96 standard deviation of differences) and their confidence intervals.

Results: The number of bolus GER and symptom episodes reported did not differ from day to day. The best agreement in SAP between the two days was found using time intervals of two minutes for cough, five minutes for crying and 15 seconds and/or 2-5 minutes for regurgitation.

Conclusions: We conclude that the standard two minute time interval is appropriate for the investigation of cough and regurgitation symptoms. The day to day agreement of SAP for crying was poor using standard criteria and our results suggest increasing the reflux-symptom association time interval to five minutes.
INTRODUCTION

Twenty four hour multichannel intraluminal pH-impedance (pH-MII) allows detection of all bolus gastroesophageal reflux (GER) including gas, mixed, liquid, acid, weakly acidic and weakly alkaline GER. The enhanced detection of bolus GER provided by pH-MII significantly increases the potential for identifying infant GER as a cause of symptoms such as excessive irritability and crying, feed refusal, cough, apnea, choking and gagging. Many of these symptoms are not specific to GER disease, and can be due to other causes, such as food allergies/intolerances, infections or functional gastrointestinal disorders such as infantile colic or constipation. With evidence now emerging that empirical prescription of acid suppression therapy to infants is largely ineffective and potentially harmful, more precise diagnostic testing offers the potential for anti-reflux therapy to be better targeted at patients in whom symptoms can be demonstrated to be due to GER.

Studies in infants, children and adults have characterized the impact of GER detection by pH-MII monitoring on diagnosis of GER-symptom associations. When compared with pH-metry alone, pH-MII has been consistently shown to increase the yield of patients in whom a positive GER-symptom association can be demonstrated. The degree of GER-symptom association is best defined using symptom association probability (SAP) which is derived from the statistical probability ($P$) that GER episodes and symptoms are temporally related using a Fisher exact test ($SAP = [1-P]*100$).

The standard GER-symptom association interval is 2 minutes and was originally based on investigations of heartburn symptoms in adults. The applicability of these criteria to infants, who typically demonstrate a different symptom profile, and in whom symptoms are reported second hand through the observation of a parent or guardian, is unknown. Furthermore, the day to day reproducibility of SAP calculations has not been examined in infants. These are important issues which need to be addressed in order for this diagnostic approach to gain wider acceptance. We therefore evaluated the effect of modifying the time window for SAP calculations and examined the day to day agreement of SAP in infants with typical GER disease symptoms.

METHODS

Patients

Ten infants, 1.6-7.7 months old and 4.6-6.7 kg in weight (6 male), were referred for ambulatory 24hour pH-MII monitoring to investigate the association of symptoms suggestive of GER disease (crying, coughing and/or regurgitation) with GER. Patients were off all anti-reflux medication for 48h prior to investigation. The study was carried out at the Children, Youth and Women’s Health Service (CYWHS) and ethical approval
of the study protocol was obtained from the Human Research Ethics Committee of the Women’s and Children’s Hospital, North Adelaide.

**Equipment**
A multi-channel intraluminal impedance ambulatory data logger (Sleuth™, Sandhill Scientific, Colorado, USA) was used to perform esophageal pH-MII monitoring studies. After calibration and intubation, correct position (between vertebral T6 and T8) of the pH-MII catheter (ComforTech Infant catheter) was confirmed by a lateral chest X-ray. Parents/guardians were instructed to maintain normal daily routines. Meal times and symptoms of ‘cough’, ‘crying’ and ‘regurgitation’ (the latter defined as both observed vomiting and behavior considered by the parent/guardian to be consistent with reflux entering the mouth) were recorded by the parent/guardian(s) using pre-set marker buttons on the data logger or a symptom diary.

**pH-impedance analysis**
The tracings were divided into the first and second 24h period and manually analyzed using semi automated impedance analysis software (Sandhill Scientific Bioview) to determine the occurrence of liquid and/or gas reflux. Liquid GER episodes were identified by a decrease in impedance of at least 50% from baseline and >2 seconds in duration traveling orally in the esophageal body. The bolus clearance time was defined as the time from onset to recovery of the impedance signal recorded on the most distal impedance channel. The proximal extent of reflux was defined by the most proximal impedance channel demonstrating an impedance drop of >50% from baseline. Gas GER episodes were scored when there was a ≥50% increase in impedance from baseline observed over any two impedance segments, either simultaneously or sequentially orally, with at least one channel showing a change in impedance of >5000Ω. Impedance detected GER episodes that exhibited both liquid and gas characteristics were characterized as mixed GER. For each impedance detected GER episode, the pH of the refluxate was determined by the esophageal pH sensor. Reflux episodes were defined according to esophageal pH as acidic (pH <4), weakly acidic (4 ≤ pH <7), or weakly alkaline (pH ≥7). Acid reflux index (% time pH<4) was determined using automated analysis of the pH tracing (Sandhill Scientific, GERD Check).

**Symptom association analysis**
For the purposes of this study, all pH-MII detected bolus GER events (gas, mixed and liquid GER combined) were used in the SAP calculation. ‘pH only’ reflux episodes, which do not exhibit the typical liquid and/or gas reflux impedance signature and therefore do not appear to be associated with refluxate of sufficient volume to fill the esophagus, were not
included in the symptom association analysis as inclusion of these events has been shown to reduce diagnostic yield.\textsuperscript{4} SAP calculations were performed using analysis software (Sandhill, Bioview) which was modified by the manufacturer to allow the duration of the association time window to be manipulated. Association time windows of 15 seconds, 30 seconds, 60 seconds, 120 seconds (standard), 300 seconds and 600 seconds were evaluated. The following algorithm was used for SAP calculation (personal communication J. Mabary, Sandhill Scientific): The entire procedure was divided into 2 minute segments. Segment \( n \) was tested for reflux and compared with segment \( n+1 \) which was tested for the existence of a symptom. Any segment intersecting with a reflux measurement was considered to exhibit reflux. Any segment intersecting with a symptom was considered to exhibit that symptom. Segments with multiple occurrences of symptoms of the same type were considered as one. Segments with multiple occurrences of reflux were considered as one. An association matrix was generated for each symptom type based on the totals representing all four Boolean combinations derived (\( R^+ / S^+; R^+ / S^-; R^- / S^+; R^- / S^- \)). The SAP for each symptom type, was computed using the Two-Tailed Fisher Exact Test probability \[ \text{SAP} = (1 - P) \times 100\% \]. The higher the SAP, the less likely that an association is coincidental.

### Statistical analysis

The degree of day to day agreement of SAP calculations was assessed by calculating the 95\% limits of agreement (LOA) using the method of Bland and Altmann.\textsuperscript{6} For each patient, the day to day difference in SAP was determined. The upper and lower LOA was calculated as the mean differences \( \pm 1.96 \) standard deviations and their respective confidence intervals (\textit{Figure 1}). This approach provided an interval within which 95\% of the differences between the two measurement periods are expected to lie. By calculating the LOA of SAP for the six different association time intervals examined, the ‘optimal’ time window for SAP calculation was defined by the time interval that produced the least variation in the 95\% limits of agreement (i.e. the smallest difference between the upper and lower LOA). This approach was used to define optimal SAP criteria based upon both the duration of the association time interval and the minimum number of symptom episodes used for calculating the SAP.

### RESULTS

The average number of reflux episodes and symptom events did not differ significantly from day to day (\textit{table 1}). All patients demonstrated symptoms of cough and crying on both days, however, symptoms of regurgitation were only recorded on both days in six patients.
Figure 1. The limits of agreement calculated from SAPs for symptoms of coughing (A), crying (B) and regurgitation (C). These data are based on patient studies with a minimum of five coughing, five crying and/or two regurgitation episodes per study day. The upper and lower limits of agreement (LOA) are shown (bold lines) with their respective confidence intervals (dotted lines respectively). Note that the LOA convergence indicates better day to day agreement. D Shows the same data for the three symptom types presented as LOAD (difference between upper and lower LOA).

Figure 1 shows the upper and lower LOA based on SAP calculated for the different reflux-symptom association time intervals. The LOA for cough and crying symptoms clearly converge around association intervals of 120 seconds and 300 seconds respectively (Figure 1 A and B). The LOA for regurgitation demonstrated a biphasic pattern, converging at 15 seconds and from 120 – 300 seconds (Figure 1C). Figure 1D plots the difference

Table 1. Average number of GER episodes and symptom events recorded.

<table>
<thead>
<tr>
<th>Event Type</th>
<th>Day 1</th>
<th>Day 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acid bolus</td>
<td>32.1 ± 4.7</td>
<td>30.6 ± 5.4</td>
</tr>
<tr>
<td>All bolus</td>
<td>72.9 ± 9.4</td>
<td>68.6 ± 10.6</td>
</tr>
<tr>
<td>Reflux Index (%)</td>
<td>6.7 ± 1.8</td>
<td>4.6 ± 0.9</td>
</tr>
<tr>
<td>No. Symptom Events</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cough</td>
<td>20.8 ± 3.2</td>
<td>14.8 ± 3.3</td>
</tr>
<tr>
<td>Crying</td>
<td>38.3 ± 7.6</td>
<td>32.8 ± 7.9</td>
</tr>
<tr>
<td>Regurgitation</td>
<td>8.8 ± 2.8</td>
<td>8.0 ± 2.5</td>
</tr>
</tbody>
</table>

The LOA for cough and crying symptoms clearly converge around association intervals of 120 seconds and 300 seconds respectively (Figure 1 A and B). The LOA for regurgitation demonstrated a biphasic pattern, converging at 15 seconds and from 120 – 300 seconds (Figure 1C). Figure 1D plots the difference
Figure 2. Contour plots showing the effect of both association time interval and symptom number on the LOAD calculated for coughing (A) and crying (B) symptoms. Note darker contour regions signify a lower difference between upper and lower limit of agreement (LOAD) and therefore a better day to day agreement. The 'x' demarcates the region of optimal agreement.
between upper and lower LOA (LOAD) for the three symptom types. In this case, the least LOAD indicates the interval of optimal convergence and therefore optimal agreement. Whilst the same pattern of convergence is demonstrated, Figure 1D also shows that the degree of agreement for crying symptoms was less overall (minimum LOAD 83.1) when compared to coughing symptoms (minimum LOAD 29.8) and regurgitation symptoms (minimum LOAD 19.2).

Figure 2 shows the further impact of symptom frequency on the LOAD by contour plotting the LOAD based upon both the association time interval and the minimum number of symptoms required for inclusion of a patient in the calculation. Data are shown for crying and coughing symptoms (regurgitation not shown due to insufficient data). The level of day to day agreement was improved by increasing the number of symptoms. As shown in Figure 2, the time interval criteria can be further optimized by applying additional criteria of a minimum number of five cough symptoms and 20 crying symptoms in order to support a reliable SAP result.

**DISCUSSION**

The additional diagnostic yield of pH-MII monitoring for determining GER-symptom associations has been previously reported in infants and children, however despite what appears to be an improvement in diagnostic methodology, there are no outcome studies showing that a high SAP is indeed predictive of symptomatic improvement with anti-reflux therapy. In the absence of these data, it is important that a SAP-based finding can be reproduced from day to day. The current study investigated this in patients undergoing 48h pH-impedance studies. The day to day agreement of SAP calculations was assessed by calculating the limits of agreement. This statistical approached allowed us to determine optimal criteria for the calculation of SAP based on symptom type, symptom frequency and association time interval.

In this study we exclusively used SAP to assess the probability of a causal association between bolus GER and symptoms. Symptom index (SI) and symptom sensitivity index (SSI) were not used because the effect of time interval on these measures was entirely predictable by the way these measures are calculated; any increase in the association time window will result in an increase in the number of observed associations between reflux and symptoms and therefore an increase in both the SI and SSI.

Our results clearly show that the time interval used in the SAP calculation determines the level of SAP agreement from day to day. Importantly the optimal time interval varies based on symptom type. A 2 minute time interval for SAP calculation is the current standard and our results support this interval for regurgitation and for cough symptoms. Whilst, for regurgitation in particular, a range of different intervals were optimal, our findings do not support any other interval being better than the 2 minute standard. In contrast our results
clearly support a change in criteria for a symptomatic association of GER with crying. Our analysis suggests that the time interval for SAP in relation to crying symptoms should be increased from 2 minutes to 5 minutes. Crying, and its range of related behaviors (irritability, fussing etc) are common symptoms in infants that are often considered due to GER disease when all other possible causes have been excluded. Crying infants are often prescribed PPI therapy empirically which has been shown to be ineffective compared to placebo.3

There are no published criteria with respect to the minimum number of symptoms that should be recorded in order to consider an SAP calculation reliable. In our own centre we have applied criteria of a minimum of five symptom episodes in order to support a positive SAP. We analyzed the impact of symptom frequency thresholds on the day to day agreement of the SAP. A positive SAP to regurgitation was found to be highly reproducible even with inclusion of a minimum of two symptom episodes. This is perhaps not surprising as true regurgitation must always follow GER. Our results suggest that parents/guardians are able to recognize this symptom well, even when it does not manifest in overt vomiting. However, as regurgitation is a relatively common occurrence in normal infants, the overall frequency of regurgitation may be more clinically relevant than findings of a positive SAP. For coughing, the day to day agreement of SAP markedly improved with the inclusion of patients with five or more symptom episodes. In contrast, the day to day agreement of SAP for crying was poor with five symptom episodes, but improved markedly with the inclusion of patients with 20 or more symptom episodes.

From this examination of the impact of association time interval and symptom frequency on the day to day agreement of SAP calculations, we conclude that the standard two minute interval is appropriate for the investigation of regurgitation and cough symptoms. We would further recommend that positive SAP findings be based on the observation of least five symptomatic GER episodes during the period of reflux monitoring. The day to day agreement of SAP for crying symptoms was poor using the standard two minutes criteria. Our results suggest an increase in the reflux-symptom association time interval from two to five minutes and, in addition, that positive SAP findings are based upon the observation of least 20 symptomatic episodes during the recording period. These findings are consistent with the fact that crying in infants is very common and can be normal behavior. It is known that, even when considered a symptom, crying is unlikely to respond to anti-reflux therapy when prescribed empirically.3

We conclude that the standard two minute time interval is appropriate for the investigation of cough and regurgitation symptoms. The day to day agreement of SAP for crying was poor using standard criteria and our results suggest increasing the reflux-symptom association time interval to five minutes and recording a minimum of 20 episodes to ensure more reliable findings.
REFERENCES


