The development and validation of a novel canine mucosal endoscopic scoring system applied to dogs with inflammatory bowel disease

by

Jennifer E. Slovak

A thesis submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Veterinary Clinical Science

Program of Study Committee:
Albert E. Jergens, Major Professor
Krysta Deitz
Heather Flaherty
JoAnn Morrison

Iowa State University
Ames, Iowa
2014

Copyright © Jennifer E. Slovak, 2014. All rights reserved.
DEDICATION

I dedicate this thesis to my husband, Dr. Robert G. Dyke, a fellow veterinarian and clinician scientist. He is an exceptionally patient, understanding and supportive spouse, for whom I can never thank enough.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>NOMENCLATURE</td>
<td>vii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>viii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>ix</td>
</tr>
<tr>
<td>CHAPTER 1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Background</td>
<td>1</td>
</tr>
<tr>
<td>Thesis formatting</td>
<td>6</td>
</tr>
<tr>
<td>References</td>
<td>7</td>
</tr>
<tr>
<td>CHAPTER 2 LITERATURE REVIEW</td>
<td>9</td>
</tr>
<tr>
<td>Chapter 2 Human endoscopic indices</td>
<td>9</td>
</tr>
<tr>
<td>Chapter 2 Veterinary endoscopic indices</td>
<td>12</td>
</tr>
<tr>
<td>Chapter 2 References</td>
<td>14</td>
</tr>
<tr>
<td>CHAPTER 3 INTER-OBSERVER AGREEMENT IN THE DUODENAL ENDOSCOPIC ASSESSMENT OF CANINE INFLAMMATORY BOWEL DISEASE</td>
<td>16</td>
</tr>
<tr>
<td>Chapter 3 Abstract</td>
<td>16</td>
</tr>
<tr>
<td>Chapter 3 Introduction</td>
<td>17</td>
</tr>
<tr>
<td>Chapter 3 Materials &amp; Methods</td>
<td>18</td>
</tr>
<tr>
<td>Chapter 3 Results</td>
<td>20</td>
</tr>
<tr>
<td>Chapter 3 Discussion</td>
<td>21</td>
</tr>
<tr>
<td>Chapter 3 References</td>
<td>26</td>
</tr>
<tr>
<td>CHAPTER 4 DEVELOPMENT AND VALIDIFICATION OF A NOVEL ENDOSCOPIC SCORING INDEX FOR CANINE INFLAMMATORY BOWEL DISEASE</td>
<td>32</td>
</tr>
<tr>
<td>Chapter 4 Abstract</td>
<td>32</td>
</tr>
<tr>
<td>Chapter 4 Introduction</td>
<td>32</td>
</tr>
<tr>
<td>Figure</td>
<td>Chapter</td>
</tr>
<tr>
<td>--------</td>
<td>---------</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1</td>
<td>Chapter 3: Description of endoscopic variables</td>
<td>29</td>
</tr>
<tr>
<td>Table 2</td>
<td>Chapter 3: Clinical characteristics of IBD dogs</td>
<td>31</td>
</tr>
<tr>
<td>Table 1</td>
<td>Chapter 4: Qualitative weighted kappa results</td>
<td>47</td>
</tr>
<tr>
<td>Table 2</td>
<td>Chapter 4: Quantitative weighted kappa results</td>
<td>48</td>
</tr>
</tbody>
</table>
### NOMENCLATURE

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IBD</td>
<td>Inflammatory Bowel Disease</td>
</tr>
<tr>
<td>GI</td>
<td>Gastrointestinal</td>
</tr>
<tr>
<td>ARD</td>
<td>Antibiotic Responsive Diarrhea</td>
</tr>
<tr>
<td>DRD</td>
<td>Diet responsive diarrhea</td>
</tr>
<tr>
<td>GALT</td>
<td>Gut Associated Lymphoid Tissue</td>
</tr>
<tr>
<td>APC</td>
<td>Antigen Presenting Cell</td>
</tr>
<tr>
<td>PRR</td>
<td>Pattern Recognition Receptor</td>
</tr>
<tr>
<td>WSAVA</td>
<td>World Small Animal Veterinary Association</td>
</tr>
<tr>
<td>CD</td>
<td>Crohn’s Disease</td>
</tr>
<tr>
<td>UC</td>
<td>Ulcerative Colitis</td>
</tr>
<tr>
<td>CDEIS</td>
<td>Crohn’s Disease Endoscopic Index of Severity</td>
</tr>
<tr>
<td>SES-CD</td>
<td>Simplified Endoscopic Activity Score for Crohn’s disease</td>
</tr>
<tr>
<td>CIBDAI</td>
<td>Canine Inflammatory Bowel Disease Activity Index</td>
</tr>
<tr>
<td>CCECAI</td>
<td>Canine Chronic Enteropathy Activity Index</td>
</tr>
<tr>
<td>pANCA</td>
<td>Perinuclear antineutrophilic cytoplasmic antibodies</td>
</tr>
<tr>
<td>CRP</td>
<td>C-reactive protein</td>
</tr>
<tr>
<td>LPE</td>
<td>Lymphocytic-plasmacytic enteritis</td>
</tr>
<tr>
<td>IL</td>
<td>Intestinal lymphangiectasia</td>
</tr>
</tbody>
</table>
ACKNOWLEDGEMENTS

I would like to thank my committee chair, Dr. Al Jergens, and my committee members, Dr. Krysta Deitz, Dr. Heather Flaherty and Dr. JoAnn Morrison, for their guidance and support throughout the course of this research, and also to Dr. Chong Wang for his statistical expertise and guidance.

In addition, I would also like to thank my friends, colleagues, the department faculty and staff for making my time at Iowa State University a wonderful experience. I want to also offer my appreciation to those who were willing to participate in my surveys and observations, without whom, this thesis would not have been possible.

Finally, thanks to my family for their encouragement and to my husband for his hours of patience, respect and love.
ABSTRACT

Measures of disease activity are necessary when diagnosing and monitoring response to therapy in canine inflammatory bowel disease (IBD). Endoscopy has been described as the gold standard to evaluate the mucosal surface and obtain biopsies for a histopathologic diagnosis. Previous veterinary studies have failed to definitively determine the benefit of endoscopic gastrointestinal mucosal evaluation in defining disease severity in dogs with IBD.

The aim of the following thesis was to evaluate the inter-observer agreement between trainee and expert endoscopists in the assessment of mucosal lesions in dogs with IBD, and to evaluate if trained operators can identify and agree upon most endoscopic lesions of mucosal inflammation using the proposed simplified endoscopic scoring index.

Archived images from endoscopic procedures performed in dogs diagnosed with IBD at the Iowa State Lloyd Veterinary Medical Center from 2002-2012 were reviewed. In total, 95 images of inflammatory and normal mucosa from dogs with IBD were displayed to 3 expert and 5 trainee endoscopists (initial test). Each picture was assessed independently by the endoscopist for inflammatory changes using established indices or interpreted as normal mucosa from multiple areas of the GI tract (ie. stomach, duodenum, and colon). Agreement was measured between the trainee and expert endoscopists for each organ. The developed index was then applied to a prospective independent group of dogs (23 total) diagnosed with inflammatory bowel disease for a validation study. Comparisons were made between 2 expert endoscopists (JES & AEJ) to measure mucosal assessment agreement.

Regression analysis showed a significant ($p<0.01$) difference between expert versus trainee endoscopy scores in duodenal evaluation trial 1, although repeat duodenal lesion
evaluation aided by use of a visual template improved the overall scores of trainee endoscopists to near that of expert endoscopists ($p=0.06$). For the validation study, the expert endoscopists had substantial to almost perfect agreement for each lesion assessed in the stomach ($k>0.8$), had moderate to substantial agreement assessing the small intestine ($k>0.6-0.84$) and substantial agreement when assessing the colon ($k>0.7-1$). The conclusion is that trained operators can identify and agree upon most endoscopic lesions of mucosal inflammation.
CHAPTER 1

INTRODUCTION:

Background

Inflammatory bowel disease (IBD) affects many species including humans, dogs, cats, and even rodent models have been used to study intestinal inflammation. In dogs, IBD refers to a group of chronic enteropathies with persistent or recurrent clinical signs such as vomiting, diarrhea, alterations in appetite, and weight loss.\textsuperscript{1-3} Chronic enteropathies in dogs is a general term describing IBD, food responsive diarrhea (FRD), and antibiotic responsive diarrhea (ARD).\textsuperscript{3-4} FRD can be divided into two main causes; immunologic and nonimmunologic.\textsuperscript{5} Non-immunologic causes of FRD are instances of food intolerance and dietary indiscretion, versus immunologic in which a dietary hypersensitivity occurs.\textsuperscript{5} These syndromes can have the same clinical signs as IBD and can sometimes be indistinguishable as some people theorize that IBD can cause FRD, or vice versa.\textsuperscript{5} ARD occurs when there are microbial imbalances in the gastrointestinal (GI) mucosa by a pathogenic species.\textsuperscript{3} A potential cause for certain dogs to have ARD is aberrant host-bacterial interactions.\textsuperscript{6} Certain dog breeds have been associated with ARD most notably the German Shepherd Dog.\textsuperscript{6}

Genetics and its role in gastrointestinal health have been an increasing area of research in veterinary medicine. There are distinct breed predispositions for canine chronic enteropathies that have been investigated. Some of these include immunoproliferative enteropathy in Basenjis, protein losing enteropathy and nephropathy in Soft-Coated Wheaten
Terriers, and Boxer granulomatous colitis.\textsuperscript{3,7} Additionally, Yorkshire terriers exhibit an increased prevalence of intestinal lymphangiectasia suggesting a possible genetic association.

As already stated, chronic enteropathies present with similar signs. When clinicians evaluate dogs for signs of chronic vomiting or diarrhea, it is important to rule out primary GI causes from non-GI causes. Clinicians perform a battery of tests including bloodwork, fecal analysis, and abdominal imaging, to try and differentiate these different causes since therapy varies in each instance. Once the non-GI causes for the clinical signs such as endocrine, kidney, and liver disorders have been ruled out, further GI diagnostic investigation is warranted.

When dogs have diarrhea, it is important to distinguish between a small and/or a large bowel origin.\textsuperscript{3,7} Large bowel diarrhea is often characterized by tenesmus, a small volume of feces with mucus or frank blood present, and an increased frequency in defecation. Dogs with small bowel diarrhea have a larger stool volume, weight loss, and may also present with signs of vomiting. Small bowel diarrhea is a common clinical sign of dogs diagnosed with IBD.\textsuperscript{7}

Questions that are often asked of the owner of these patients include volume of stool produced, frequency of defecation, straining, and the presence of blood or mucus in the stool.\textsuperscript{4} The answers to these questions allow the clinician to often localize the source of the intestinal problem. Once it is determined which area of the GI tract is affected (based on clinical signs and laboratory tests [ie, hypocobalaminemia]), GI endoscopy is often the next diagnostic step.\textsuperscript{2}

In order to obtain a diagnosis of IBD, one must exclude other causes of intestinal disease and obtain mucosal biopsies that demonstrate histopathologic evidence of mucosal inflammation.\textsuperscript{1-4} While the exact etiopathogenesis of IBD in dogs is unknown, disruption of
the GI mucosa, alterations in the GI microbiota (dysbiosis), and dysregulation of mucosal immunity have been implicated.\textsuperscript{1-5} The normal functioning GI tract has inherent protective mechanisms including the gut associated lymphoid tissue (GALT), development of oral tolerance, and the production of IgA.\textsuperscript{3,8} The current hypotheses in human, murine, and canine models of IBD suggest that genetics, the mucosal immune system, diet, and disruption of the GI microbiota can be contributing factors to disease.\textsuperscript{3,8}

In the normal animal, antigen presenting cells (APCs) continually sample antigens within the lumen of the GI tract.\textsuperscript{8} Pending the type of antigen, signals created by the APCs stimulate the adaptive immune response to eradicate the pathogen.\textsuperscript{8} Presentation of the antigen to the lymphocyte induces lymphocyte activation.\textsuperscript{3} Further activation signals from the APC and helper T cells occur, which induce sensitized effector T lymphocytes to travel through the mesenteric lymph nodes via lymphatic drainage and return to the lamina propria beneath the GI mucosal epithelium.\textsuperscript{3} Here, they are primed to respond to a given antigen they have been exposed.

In IBD, the mechanism of inflammation arises from a defect in recognition of normal commensal bacteria by the intestinal immune system.\textsuperscript{8} Pattern recognition receptors (PRR) on the mucosal epithelium such as toll-like receptors or the APCs that have mutated, lead to the “misguided” recognition of commensal bacteria as potential pathogens.\textsuperscript{3,8} This misrepresentation induces T cell differentiation and the production of robust quantities of pro-inflammatory cytokines and a subsequent influx of mucosal inflammatory cells.\textsuperscript{7} The most common pathologic type of IBD in both dogs and cats is lymphoplasmacytic enteritis.\textsuperscript{2,3}

Obtaining biopsies in dogs with IBD is paramount to making a correct diagnosis. Intestinal biopsies may be obtained using surgical or endoscopic techniques. There are
guidelines established by the World Small Animal Veterinary Association Gastrointestinal Standardization Group to help veterinarians diagnose and characterize the histopathological GI mucosal changes in dogs and cats with GI disease. The histopathologic evaluation of the biopsies obtained is paramount for an accurate diagnosis of IBD. Cell type infiltration of the lamina propria aids in clinical decision making. For instance, the presence of macrophages and neutrophils may suggest a potentially infectious pathogen, versus infiltration of eosinophils, suggesting a parasitic cause or response to dietary constituents.

Therapies for IBD are influenced by the histopathology of the GI biopsies and the severity of the clinical signs. The goal of therapy largely involves ameliorating the clinical (GI) signs. Empirical treatment with a highly digestible novel protein diet should be performed first. Typically, a novel protein diet is suggested for dogs with chronic enteropathies; some examples of novel proteins available as canine diets include bison, kangaroo, duck, salmon, and venison. Many commercially available dog foods have a mix of proteins which provides a potential for exposure to numerous dietary antigens. Some of the proposed dietary antigens present in commercial dog foods include soy, beef, gluten, lactose, and wheat. Currently, several major dog food companies have developed hydrolyzed protein diets that are suggested to be “hypoallergenic”. The hydrolytic process reduces the native protein to a size that is theorized to be less likely to stimulate the immune system. Individual response to diet trials varies between dogs. A diet trial should be done for at least 2-4 weeks to see if there is a favorable response. There have been many studies that have shown improvement of clinical signs in up to 50% of dogs with chronic enteropathies.

Empiric therapy for protozoa such as *Giardia* and other endoparasites is commonly prescribed by clinicians for dogs with chronic enteropathies. Typically, the anti-parasitics
fenbendazole and pyrantel pamoate are considered standard choices for dogs presenting with gastrointestinal disease. A 28 day trial of antibiotics such as metronidazole and tylosin can also be used empirically to rule out ARD. If the clinical signs recur after discontinuing the antibiotic, then long-term antibiotic therapy can be continued. If the response to these empiric/trial therapies are incomplete or poor, or if there is moderate to severe evidence of inflammation on histopathology, further treatment options should be considered. Since IBD is the result of mucosal immune dysregulation, oral immunosuppressive therapy to alleviate the offending clinical sign is warranted. Some common immunosuppressive agents include prednisone, budesonide, and cyclosporine, among others. Dosing of these medications are slowly tapered over weeks to months to ameliorate the offending clinical signs using the lowest effective dose.

IBD in humans describes a group of inflammatory GI disorders including Crohn’s disease (CD) and ulcerative colitis (UC) that induce an immune mediated inflammation within the mucosa of the gastrointestinal tract (but predominantly the colon and rectum). This inflammation induced by noxious substances in the lumen stimulates local production of TNF-a, IL-6 and IL-1b. Histopathology of the mucosa is paramount in diagnosing disorders of the gastrointestinal tract, as IBD is arbitrarily characterized by the type of inflammatory infiltrate present. Intestinal mucosal samples can be obtained via surgical intervention, or less invasively, by endoscopy.

Gastrointestinal endoscopy is commonly utilized in human and veterinary medicine as a relatively non-invasive diagnostic tool for the identification, assessment, and localization of disease in the gastrointestinal tract. Endoscopy allows human and veterinary physicians a means to obtain histopathologic samples to further evaluate and characterize disease in their
patients as well as response to treatment interventions.\textsuperscript{15,16} However, in veterinary medicine, few studies have examined and correlated histopathology and clinical signs.\textsuperscript{10} In addition, it is well recognized that variation of lesion interpretation exists among endoscopists and pathologists in both the human and veterinary literature.\textsuperscript{10,13,17,18}

In human medicine, development of multiple standardized disease indices and activity scores to define inflammatory activity for CD and UC have been proposed and validated.\textsuperscript{15,19} The purpose of these activity scores and indices has been to unify gastroenterologists’ endoscopic lesion description and to characterize severity.\textsuperscript{15,19,20} Human physicians have tried to provide an objective and reliable way to assess endoscopic lesions for patient disease monitoring and continuity of care between attending clinicians.\textsuperscript{15}

There is very little veterinary literature available that examines endoscopic mucosal lesions and the consistency of mucosal interpretation among endoscopists. There have not been any published studies that have assessed inter-observer agreement between veterinary endoscopists when evaluating mucosal lesions. The goal of the following thesis is to further describe the importance of gross endoscopic disease description, evaluate groups of endoscopists’ interpretations of endoscopic lesions, and suggest a simplified mucosal endoscopic index to implement when evaluating dogs with IBD.

**Thesis Formatting**

The following Master of Science thesis is arranged in a journal paper format. The references utilized for each chapter are listed at the end of each chapter. Chapter 1 begins with a general introduction to the thesis including a background of canine IBD and a description of human mucosal inflammation/lesion indices. Immediately following chapter 2,
the literature review, I have included two papers that either have been submitted to peer reviewed journals as novel and individual manuscripts, or are in preparation for submission. Chapter 5 of the thesis summarizes the overall concept of the research hypothesis, results, and areas for further research.

References


CHAPTER 2

LITERATURE REVIEW:

**Human endoscopic indices**

Most gastroenterologists agree that Dr. Sidney Truelove is the founding father of IBD therapy. He recognized in the early 1950s that just evaluating symptoms of the patient was an inadequate means of measuring treatment efficacy, and a more objective measure of disease activity was necessary. In the decade following Dr. Truelove’s studies and observations regarding IBD, Dr. Baron critically evaluated endoscopic scoring of lesions. Dr. Baron emphasized the importance of visual mucosal interpretation in IBD patients; however, the appearances of the diverse lesions can be difficult to describe without clearly defined terms. Dr. Baron emphasized the necessity of evaluating those mucosal characteristics that could be consistently reported by various observers. He placed more emphasis on the presence or absence of clearly defined mucosal features and thought that criteria with descriptive terms regarding shades of color and texture were less reliable. Initially, human mucosal assessments could only be made to the locations limited by rigid sigmoidoscopy. Through the years improvements have been made in endoscope technology and comfort. Today, we currently have the advantage of flexible and capsule endoscopes with high definition visualization capabilities to fully evaluate mucosal surfaces, but there is still no unified human measurement of mucosal endoscopic activity.

In humans, the two most commonly recognized and well-described inflammatory bowel disorders are Crohn’s disease (CD) and Ulcerative colitis (UC). It is well known among physicians that a definitive cause for the underlying inflammatory condition has not
been identified, and that therapy has largely been focused on ameliorating clinical signs.\textsuperscript{4-6} These diseases are notorious for relapse and characterized by abdominal pain, diarrhea and even bleeding.\textsuperscript{5} Monitoring therapy and identifying whether patients are relapsing or are in remission are integral components of patient management.

Although various non-invasive methods of intestinal disease measurement have been developed, there is a lack of specificity.\textsuperscript{5,6,11} Serologic and fecal markers such as C-reactive protein and calprotectin have been utilized, but further studies need to be performed to solidify their use in disease diagnosis and monitoring.\textsuperscript{5,6,11} The most common way to monitor disease progression/healing of lesions is via direct mucosal evaluation. This is done largely with gross endoscopic mucosal evaluation in combination with histological examination.\textsuperscript{3,5} High quality images of the mucosal surface are a means to critically evaluate lesions and allow review by experts without a bias of current treatment for assessment.\textsuperscript{1} Limitations that can affect mucosal interpretation include the preparation of the bowel, the equipment used, and the experience and training of the endoscopist on lesion identification.\textsuperscript{3,5,6}

The reliability of the individual endoscopist’s mucosal interpretation and assessment has been questioned.\textsuperscript{7-11} In addition, there have been numerous debates regarding lesion identification between endoscopists with varying training levels.\textsuperscript{7-11} As a result, multiple endoscopic grades and indices have been developed and validated. For instance, there are several activity indices in use when describing patients with UC, such as Truelove, Baron, Blackstone, and the modified 6-point activity index.\textsuperscript{2,4,7,11} Most of these indices prioritize several key features in the lesions assessed including; friability, granularity, ulcers, erythema, and mucosal vascular patterns.\textsuperscript{2,4,7,11} There are independent variables and scoring dependent upon which individual index is used. As already mentioned, the Baron index minimizes the
importance of lesion identification for features that cannot be clearly defined and rely on individual subjective interpretation such as color changes (erythema) or texture of the mucosal surface (granularity). This is in contrast to some of the other indices such as the Blackstone and modified 6-point scale that readily describe and even quantitate these types of lesions.

In regards to CD, there are also many disease activity indices currently in use. The two most common are the Crohn’s Disease Endoscopic Index of Severity (CDEIS) and the Simplified Endoscopic Activity Score for Crohn’s disease (SES-CD). The CDEIS was developed to report the global appraisal of lesion severity among endoscopists in a standardized way. The format of the CDEIS involves evaluating 5 specific colonic locations and describing if 4 specific pre-determined/specific lesions are present. This method was deemed reliable, albeit time consuming, and the scores obtained were reproducible by multiple endoscopists. In response, the SES-CD was developed to provide a simplified and more rapid clinical mucosal assessment by defining 4 variables and using a quantitative system of measurement.

Countless numbers of human endoscopic indices have been developed and utilized with various amounts of success since the advent of human endoscopy. There is no universally accepted and clearly comprehensive index for either UC or CD. The benefits of gastrointestinal endoscopy including disease diagnosis, monitoring treatment response, and disease surveillance have been widely accepted and proven in various studies. Limiting factors for the use of endoscopic indices are reliability of results, reproducibility of lesion assessment, simplicity of the score, experience of the endoscopist, and availability of mucosa for assessment. Based on all the scoring systems that have been evaluated, most agree
that clearly defined terminology of specific lesions such as friability, ulcers, and granularity should be noted, as well as the extent/size of the mucosal lesion. Most evaluation and validation of endoscopic indices agree that operator experience matters in regards to lesion assessment.

**Veterinary endoscopic indices**

The degree and severity of inflammatory bowel disease in veterinary patients is similar to that of humans. In comparison to the human literature, there are only a handful of studies that highlight or even quantify a mucosal endoscopic index. In fact, scoring indices for inflammatory bowel disease in veterinary medicine were not in use until the development of the Canine Inflammatory Bowel Disease Activity Index (CIBDAI). This is a clinical index developed to summate 6 variables with quantitative scores from 0-3 to describe mild, moderate, or severe IBD. There are currently no universally accepted endoscopic scoring indices although several methods for disease descriptions have been proposed in individual studies recently published.

In 2007, Allenspach and colleagues described risk factors for dogs diagnosed with chronic enteropathies. Seventy dogs were evaluated using activity indices CIBDAI and the Canine Chronic Enteropathy Activity Index (CCECAI), and these were compared to their endoscopic and histologic scores. An independent duodenal and colonic endoscopy score was developed for this prospective study that included a grading system from 0-3. The variables assessed were friability, erythema, white speckling, ulcers, cobblestone appearance, and difficulty insufflating the bowel during endoscopy. None of the variables were individual assessments, but rather a combination, as stated in the paper. There was no correlation of the endoscopy score to the disease activity indices, although a grade 3 duodenal
endoscopy score was associated with a negative outcome.\textsuperscript{14} This study did not develop a detailed and specific mucosal endoscopic index, nor did it find major correlations with other disease indices.

Also in 2007, Garcia-Sancho and colleagues published a prospective study evaluating clinical, macroscopic and histopathologic assessments of dogs with nonproteinemic lymphocytic-plasmacytic enteritis (LPE). Macroscopic variables were evaluated in the stomach and duodenum of the study dogs consisting of erosions, erythema, granularity, presence or absence of bile, lack of elasticity, and friability.\textsuperscript{15} The variables were rated in severity from 0-3 and for presence or absence of gross lesions.\textsuperscript{15} In all study dogs with LPE, there were macroscopic lesions, and 75\% of the animals post-treatment displayed macroscopic improvement.\textsuperscript{15} In contrast to the Allenspach et al study\textsuperscript{14}, this paper showed improvements in the gross endoscopic mucosal appearance in treated dogs and highlighted the most common pretreatment mucosal findings in both the stomach and duodenum of dogs with LPE.\textsuperscript{15}

Finally in 2012, Larson et al reported on the duodenal endoscopic findings and histopathologic confirmation of intestinal lymphangiectasia (IL) in dogs. An endoscopic grading scale of severity was developed based on mucosal granularity, active lymphatic discharge and white foci.\textsuperscript{16} The severity of the IL was categorized from 0-3.\textsuperscript{16} Based on the results of the study, there was poor sensitivity and specificity of the duodenal mucosal endoscopic appearance in predicting IL in dogs.\textsuperscript{16}

These 3 veterinary studies are by no means a comprehensive list of the available endoscopic indices available. Since there is no universal veterinary standard for evaluating the gastrointestinal mucosal surface, each study involving gastrointestinal endoscopy often
developed an individual scoring index. This makes the interpretation of lesion assessment difficult to analyze, as there are a multitude of lesion variables assessed, discrepancies in the importance of severity versus presence or absence of lesions, and the ability and experience of the endoscopist is often not considered. This highlights the inherent problems with endoscopic mucosal assessments made in dogs with various chronic enteropathies and emphasizes the necessity for a universal system of endoscopic mucosal assessment.

References


INTER-OBSERVER AGREEMENT IN THE DUODENAL ENDOSCOPIC ASSESSMENT
OF CANINE INFLAMMATORY BOWEL DISEASE


*Submitted and under review at the Journal of Veterinary Internal Medicine

J. Slovak chose images/cases utilized; organized the display/construction of images, templates and evaluation forms; collected, graded and interpreted data.

Abstract

**Background:** Measures of inflammatory activity are essential in defining disease burden at diagnosis and for determining effects of treatment in dogs with inflammatory bowel disease (IBD). GI endoscopy is performed for direct inspection of the mucosa and acquisition of biopsies for histopathologic evaluation. Endoscopic observations might also be used to determine the extent and severity of the disease.

**Aim:** To evaluate the inter-observer agreement in the endoscopic duodenal mucosal assessment in canines diagnosed with IBD.

**Methods:** Thirty-five archived endoscopic images of grossly normal (n=6) and inflamed (n=29) duodenal mucosa were displayed to 3 expert and 5 trainee endoscopists. Each image was assessed independently by endoscopists for inflammatory changes using established
indices (i.e., hyperemia, granularity, friability, lymphatic dilatation, erosions) or interpreted as normal mucosa (trial 1). A repeated trial (trial 2) was administered with images re-randomized one month later accompanied by a visual template.

**Results:** There was slight inter-observer agreement in initial lesion identification for expert and trainee endoscopists in trial 1 ($k \leq 0.02$, $p > 0.05$). Inter-observer agreement improved in trial 2 for both expert and trainee endoscopists, ($k=0.2$, $p > 0.05$) for experts and ($p < 0.05$) for trainees. There was a significant ($p < 0.01$) improvement in trainee endoscopy scores of lesions from trial 1 to trial 2. Regression analysis showed a significant ($p < 0.01$) difference between expert versus trainee endoscopy scores in trial 1. Repeat lesion assessment aided by use of a visual template (trial 2) improved the overall scores of trainee endoscopists to near that of expert endoscopists ($p = 0.06$).

**Conclusions:** Accurate assessment of IBD disease activity from endoscopic findings benefitted from operator experience.

Different indices have been proposed to measure the activity (and/or severity) of canine inflammatory bowel disease (IBD) to evaluate efficacy of treatment in clinical trials.$^{1-3}$ All of these indices are based on clinical signs and/or biological data. Gastrointestinal (GI) endoscopy is a well-established technique to directly visualize the mucosa and acquire targeted biopsy specimens for histopathologic confirmation of intestinal inflammation. Because endoscopy is routinely performed for diagnosis of canine IBD, endoscopic findings (i.e., abnormal mucosal appearances) could be used to measure disease activity.

Several different endoscopic indices for evaluation of inflammatory activity in human IBD (i.e., Crohn’s disease [CD]$^{4,5}$ and ulcerative colitis [UC]$^{6-9}$) have been designed. All of
these scoring systems were based on the severity/extent of mucosal granularity, vascular pattern, vulnerability of mucosa, and/or mucosal damage (mucus, fibrin, exudates, erosions, and ulcers) observed during colonoscopy. However, no standardized model has been established. Separate studies in dogs with small intestinal IBD have yielded conflicting results on the utility of endoscopic scoring as a measure of disease activity.\textsuperscript{1,10} One potential reason for this discrepancy could be inter-observer variation in identifying endoscopic abnormalities based on operator experience and the lack of systematic endoscopic assessment. The aim of the present study was to evaluate the inter-observer agreement in the assessment of endoscopic activity in canine duodenal IBD.

**MATERIALS AND METHODS**

**Selection of Images**

Two hundred archived endoscopic images from consecutive duodenoscopy procedures performed in dogs with IBD between 2004 and 2012 at Iowa State University were retrieved from a computerized database and reviewed. A total of 35 endoscopic images of normal (some images obtained post-biopsy) and inflamed duodenal mucosa from 25 IBD dogs were selected for study enrollment. Image selection was determined by joint agreement of authors JES and AEJ. A diagnosis of canine IBD was based on previously established clinicopathologic and histopathologic criteria.\textsuperscript{1,2,10-12} Endoscopic interpretation of intestinal lymphangiectasia included observation of multifocal to diffuse white foci within the mucosa suggestive of lymphatic distension.\textsuperscript{15} Duodenoscopy procedures were performed using a commercial video endoscope (Olympus GIF-160, Olympus Optical, Tokyo, Japan) with still
images of normal and abnormal mucosa captured by the endoscopist. The file size of the downloaded images was approximately 100 kb, with a pixel array of 640 x 480 and 24-bit color. These still images were then arranged in a PowerPoint presentation for testing purposes.

**Assessment of Images**

Endoscopic still images in PowerPoint format were assessed by 3 expert and 5 trainee endoscopists for inflammatory activity. Expert endoscopists were defined as individuals with advanced clinical training and active operator participation in GI endoscopy over the preceding 24 months. The experts were experienced and familiar with mucosal lesions of disease activity as identified with GI endoscopy. Trainee endoscopists had minimal endoscopic training and lacked consistent endoscopic operator experience over the same 24 month period.

Images were randomized by means of a Research report randomizer program and assessed independently by each endoscopist for inflammatory changes as originally determined by JES and AEJ. Neither the clinical data nor the date on which the image was taken was made available to the endoscopists. The endoscopic variables evaluated included hyperemia, erosions, granularity, friability, lymphatic dilation, or the mucosal appearance was interpreted as normal (Table 1). Written definitions of each variable were made available to all endoscopists. If an individual image contained more than one mucosal abnormality, the endoscopist was asked to identify the salient lesion (trial 1). The assessment of mucosal inflammatory changes was repeated 1 month after the first assessment (trial 2), although the endoscopists were not informed that they were going to assess the same images (order re-
randomized) a second time. Additionally, each endoscopist was instructed to review a template of representative mucosal lesions (Figure 1) before image re-assessment (trial 2) to see whether this exercise improved endoscopy scores.

**Analysis of Data**

Data was collected from each operator using pre-designed Excel® spread sheets for statistical analysis. Fleiss Kappa coefficients were calculated to assess agreement among multiple raters within expert and trainee groups and tested against null value 0 using the “irr” package in R. A mixed effects logistic regression model was used to analyze endoscopy score agreement with gold standard for comparison of assessment accuracy using the Glimmix procedure in SAS. Group (trainee vs expert), trial and their interaction were the fixed effects in model, whereas endoscopist was the random effect. A p value < 0.05 was considered statistically significant.

**RESULTS**

The base-line clinicopathologic characteristics in IBD dogs were similar to previous reports. The affected dogs were predominantly middle-aged (age range 1-11 years), exhibited chronic gastrointestinal signs, and had variable disease activity as evidenced by clinical scores (Table 2). There were 11 spayed females and 14 neutered males included in the study. Dogs with IBD included the following pedigrees: 4 West Highland White Terriers, 3 Golden Retrievers, 2 mixed breed dogs, 2 Boxers, 2 Labrador Retrievers, 2 Shih Tzus, 2 Yorkshire Terriers and 1 each of Wheaton Terrier, German Shepherd Dog, Viszla, English Bulldog, Cocker Spaniel, Gordon Setter, Beagle, and Miniature Poodle. None of the dogs had
evidence of extra-alimentary tract inflammation (based on results obtained from diagnostic testing), and each dog had failed to respond fully to previous dietary and antibiotic interventions.

Among the 35 images of the test set obtained during duodenoscopic examination in IBD dogs, 6 were of normal mucosa (some images obtained post-biopsy), 6 showed friability, 5 showed hyperemia, 6 showed increased granularity, 7 showed erosions, and 5 showed lymphatic dilation.

Based on Fleiss Kappa statistics evaluation, the inter-observer agreement within expert and trainee groups improved among experts from trial 1 $k<0.01$, $p>0.05$, to trial 2 $k=0.2$, $p>0.05$ and among trainees from trial 1 $k=0.02$, $p>0.05$ to trial 2 $k=0.2$, $p<0.05$. Using the Glimmix procedure, showing comparison within the groups, there was a significant ($p<0.01$) improvement (17.1%) in the trainee endoscopists regarding lesion assessment from trial 1 to trial 2. The expert endoscopists showed no statistical significant improvement between trial 1 and trial 2 ($p=0.19$), although there was a 7.6% improvement. Regression analysis showed a significant ($p<0.01$) difference between operator groups regarding trial 1 lesion assessment. Repeat duodenal image evaluation aided by use of a visual template, (trial 2) improved the overall scores of trainee endoscopists to near that of expert endoscopists, ($p=0.06$).

**DISCUSSION**

The results of this study indicate that operator experience matters when making endoscopic mucosal assessments. Operator experience may be gained from performing numerous endoscopic procedures and/or the provision of a written/pictorial template to aide identification of mucosal lesions.
Gastrointestinal endoscopy is an important tool in the diagnosis of canine IBD. Gastroscopy, enteroscopy, and colonoscopy are of value in the assessment of specific organ involvement in IBD and to differentiate IBD from other causes of chronic enteropathy. Recent advances in patient preparation and instrumentation, mucosal examination techniques, and the development of forceps biopsy standards have made GI endoscopy the preferred method for diagnosis of small and large intestinal inflammation.

Canine IBD is often characterized by a relapsing and remitting clinical course. Determination of inflammatory activity is important for assessing disease severity and for tailoring patient therapy. Different indices for assessment of disease activity have been proposed. Clinical indices utilize scoring systems derived from GI signs alone (CIBDAI) or in combination with laboratory testing (CCECAI) to quantify intestinal inflammation. Noninvasive serologic markers including perinuclear anti-neutrophil cytoplasmic antibodies (pANCA) and C-reactive protein (CRP) may provide only indirect assessment of disease activity. Histopathological examination, while required for diagnosis of IBD, is hindered by poorly standardized grading criteria and disagreement among pathologists in defining mucosal inflammation.

Since endoscopy provides immediate and direct assessment of intestinal mucosal damage, endoscopic findings might be used to measure inflammatory activity. Several endoscopic activity indices for CD and UC are in use. Salient lesions of human IBD (ie, CD and UC) range from erythema, loss of vascularity, friability, and granularity of the mucosa to erosions/ulceration. Similar endoscopic indices have been used in canine IBD including erythema, friability, erosions/ulceration, cobble-stone appearance (granularity), white speckling on the surface, and difficulty in insufflating (stenosis). Clinical trials
utilizing endoscopic scoring for canine IBD are limited and have provided conflicting results on the utility of endoscopic scoring as a measure of disease activity.\textsuperscript{1,10} A reason for this discrepancy could be inter-observer variation in identifying endoscopic abnormalities based on operator experience and the lack of systematic endoscopic assessment.

This study investigated inter-observer agreement in the assessment of endoscopic activity in IBD dogs using defined descriptors of mucosal inflammation. Both written descriptions and a visual template of mucosal lesions were used to assess the role of operator experience in defining duodenal endoscopic activity in dogs. Our results indicated that there was slight to fair inter-observer group agreement in lesion identification in expert and trainee endoscopists for either trial 1 or trial 2. However, there was significant inter-observer difference in lesion assessment when still images of IBD were evaluated by the 3 experienced versus 5 trainee endoscopists in trial 1. Analysis of inter-observer agreement showed a significant difference between operator groups regarding lesion assessment with expert endoscopists having less chance of disagreement regarding the identification of endoscopic abnormalities. This observation of improved inter-observer agreement of experienced versus trainee endoscopists emphasizes the value of operator experience and is similar to results in humans with IBD.\textsuperscript{8,27}

Allenspach et al. previously evaluated the association between endoscopic scores of the duodenum and colon with other inflammatory indices (ie, clinical activity [CIBDAI] and histopathology) as partial assessment of long-term outcome in dogs with chronic enteropathies.\textsuperscript{1} In this study, numerical scores (range 0-3; normal to severe mucosal inflammation) were assigned by one of 2 operators using mucosal assessment criteria of erythema, friability, white speckling, granularity, and luminal stenosis. No correlation was
found between endoscopy scores and histology scores pre- versus post-treatment; however, an endoscopy score of 3 in the duodenum, indicative of severe inflammation, was significantly associated with negative outcome. In a separate study, Garcia-Sancho et al. performed endoscopic examination in 16 dogs diagnosed with lymphocytic-plasmacytic enteritis and evaluated gastric/duodenal lesions of mucosal erythema, granularity, friability, erosions, and luminal distension before and after IBD therapy. While the number and relative experience of endoscopists were not noted in this report, these investigators showed significant differences between pre- and post-treatment macroscopic endoscopic lesions in the stomach and duodenum.

Our choice of endoscopic mucosal characteristics (ie, hyperemia, friability, granularity, erosions, and lymphatic dilatation) to evaluate was based on the personal experiences of AEJ in performance of duodenoscopic procedures over many years. We have observed that observer variation for graded characteristics (i.e., mucosal hyperemia – is it pale, pink, or red?) is quite high, while that for discontinuous variables (i.e., presence or absence of erosions) is generally low. More importantly, we have also observed that operator experience plays an important role in endoscopic assessment with trainee endoscopists more likely to miss mucosal lesions or misinterpret normal versus abnormal mucosa. The results of this study would confirm these previous anecdotal observations. While hyperemia and luminal distensibility have been used in previous endoscopic indices for dogs, we have not found them useful in the past or in the current study. In fact, of the endoscopic variables evaluated across a spectrum of GI organs (i.e., stomach, small intestine, and colon) we found mucosal hyperemia to have the greatest variability amongst all observers, including experienced endoscopists (JES and AEJ, unpublished observations).
With regards to statistical analysis, the use of the Fleiss Kappa coefficients were calculated to assess agreement among multiple observers versus Cohen’s kappa which compares inter-observer agreement between 2 observers. Although the kappa scores within the expert and trainee groups for trial 1 showed only slight agreement, they both improved to fair agreement for both groups in trial 2. The $p$ value for the experts in both trials were greater than 0.05 and could have been the result of a Type II error. However, the trainee’s $p$ value in trial 2 was less than 0.05, which is unlikely to have occurred by chance alone.

There were several potential limitations of our study. First, we utilized a single center for our study and focused only on duodenal endoscopic assessment of dogs with IBD. Whether the same results for inter-observer variability across different study centers, endoscopic interpretation by non-specialist clinicians, or evaluation of other alimentary tract organs (i.e., stomach, ileum, and colon) in dogs having different enteropathies might yield similar results was not assessed. Second, we used still images of endoscopic lesions versus video streams to evaluate variation between operator cohorts. Our rationale was that still images in texts or continuing education events are routinely utilized for endoscopic training purposes. Additionally, a manageable number of still images could be more easily evaluated twice by the same clinicians in this study, which assured good compliance in spite of their other professional duties. Accurate assessment of UC endoscopic activity from archived still images has been previously reported.8

In summary, a simple classification of the variable mucosal appearances in the duodenum of dogs with IBD is described. According to the results from this study, accurate assessment of IBD activity from duodenal endoscopic findings benefitted from operator
experience. Acceptable agreement rates can be obtained by endoscopists under training using well-defined endoscopic appearances.

Footnotes:

a Research report randomizer www.randomizer.org

b Powerpoint™

c Excel™

d “irr” package in R R Core Team, Vienna, Austria

REFERENCES


Table 1: Descriptions of endoscopic variables used in the study

<table>
<thead>
<tr>
<th>Endoscopic criteria for duodenal mucosal assessment</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperemia</td>
<td>Gradations of mucosal redness (pale $\rightarrow$ red)</td>
</tr>
<tr>
<td>Friability</td>
<td>Mucosal bleeding on contact with endoscope or biopsy forceps</td>
</tr>
<tr>
<td>Granularity</td>
<td>Alteration in the texture of the mucosal surface</td>
</tr>
<tr>
<td>Erosions</td>
<td>Superficial linear mucosal defect(s) with hemorrhage</td>
</tr>
<tr>
<td>Lymphatic dilatation</td>
<td>Multifocal to diffuse white foci within the mucosa</td>
</tr>
</tbody>
</table>
Figure 1: Representative lesion images

Normal (A,B); Hyperemia (C,D); Friability (E,F); Lymphatic dilation (G,H);

Granularity (I,J); Erosions (K,L)
Table 2: Clinical characteristics of IBD dogs

**Baseline characteristics of IBD dogs**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age (yr.)</td>
<td>6.8</td>
</tr>
<tr>
<td>Male sex, n (%)</td>
<td>14 (56)</td>
</tr>
<tr>
<td>Disease duration (mo.)</td>
<td>4.8</td>
</tr>
<tr>
<td>CIBDAI score (^a)</td>
<td>5.9</td>
</tr>
<tr>
<td>Endoscopic lesions (^b)</td>
<td>100%</td>
</tr>
<tr>
<td>Histopathologic grade (^c)</td>
<td></td>
</tr>
<tr>
<td>Mild IBD</td>
<td>24%</td>
</tr>
<tr>
<td>Moderate-severe IBD</td>
<td>76%</td>
</tr>
</tbody>
</table>

\(^a\) Mean disease activity at diagnosis; range 0 – 18
\(^b\) Mucosal lesions of friability, granularity, or erosions
\(^c\) Predominant lymphocytic-plasmacytic inflammation
CHAPTER 4

DEVELOPMENT AND VALIDATION OF A NOVEL ENDOSCOPIC SCORING INDEX FOR CANINE INFLAMMATORY BOWEL DISEASE


*In preparation for submission to the Journal of Veterinary Internal Medicine

J. Slovak co-chose images/cases utilized; organized the display/construction of images, templates and evaluation forms; collected, graded and interpreted data.

Abstract

Background: Endoscopy is the gold standard to evaluate dogs with IBD, but the reliability of the assessment of endoscopic lesions is unclear. Previous studies have provided conflicting results as to the benefit of endoscopic activity in the diagnosis and management of canine intestinal disease. Moreover, mucosal criteria have varied between trials, which may create confusion and confound accurate interpretation between endoscopists.

Aim: The aim of this study was to develop and validate a mucosal endoscopic index for assessing the severity of lesions in dogs with inflammatory bowel disease.

Methods: In total, 95 images of inflammatory and normal mucosa from dogs with chronic enteropathies were displayed to 3 expert and 5 trainee endoscopists (initial test). Each picture
was assessed independently by the endoscopist for inflammatory changes, using established indices or interpreted as normal mucosa from multiple areas of the GI tract (i.e., stomach, duodenum, and colon). Agreement was measured between the trainee and expert endoscopists for each organ. The results of the duodenal assessments have been reported in a previous study. The developed index was then applied to a prospective independent group of dogs (23 total) diagnosed with inflammatory bowel disease for a validation study. Comparisons were made between 2 expert endoscopists (JES & AEJ) to measure mucosal assessment agreement.

**Results:** There was only slight agreement among the experts when evaluating the stomach \( k=0.04 \). There was fair agreement among the trainees evaluating the stomach \( k=0.2 \) and fair agreement for both the trainees and experts when evaluating the colon \( k=0.2 \). For the validation study, the expert endoscopists had substantial to almost perfect agreement for each lesion assessed and for the total score in the stomach (\( k>0.8 \)), had moderate to substantial agreement assessing the small intestine (\( k>0.6-0.84 \)), and substantial agreement when assessing the colon (\( k>0.7-1 \)).

**Conclusions:** Trained operators can identify and agree upon most endoscopic lesions of mucosal inflammation.

**Introduction**

Inflammatory bowel disease is one of the most important differential diagnoses in dogs with chronic enteropathies.\(^1-4\) Additional causes include antibiotic responsive diarrhea (ARD) and diet responsive diarrhea (DRD).\(^1-4\) The diagnosis of IBD relies heavily on clinical
signs, response to therapy, serologic and fecal testing, abdominal imaging, and gastrointestinal endoscopy.\textsuperscript{4,8}

Gastrointestinal endoscopy provides an immediate method for assessing intestinal mucosal damage. Endoscopy is considered the gold standard to evaluate and obtain biopsies from dogs with IBD, but the reliability of mucosal lesion assessment is unclear.\textsuperscript{6,9-13} There is a paucity of published veterinary literature regarding mucosal assessment in dogs with chronic enteropathies.\textsuperscript{3,9-10} In human medicine, the most commonly recognized endoscopic mucosal indices are those established for ulcerative colitis (UC) and Crohn’s disease (CD).\textsuperscript{6,12-18} There is little agreement in the human literature as to which index is superior for evaluation, although most agree that an ideal endoscopic index should have high concordance between endoscopists and high reproducibility regardless of the endoscopist’s experience.\textsuperscript{14,16}

Previously published veterinary studies have failed to determine whether standardized endoscopic evaluation is useful in defining the severity and extent of inflammatory lesions in affected dogs. Former veterinary clinical trials have used multiple individual endoscopic indices in order to classify mucosal variables for each individual study.\textsuperscript{3,9-10} The aim of the present study is to develop and validate an endoscopic index for mucosal assessment in dogs with inflammatory bowel disease.

**MATERIALS AND METHODS**

*Selection of Images (initial test)*
Archived endoscopic images from gastrointestinal (GI) endoscopic procedures performed at Iowa State University from 2002-2012 were retrieved from a computerized database and reviewed. A total of 30 stomach endoscopic images from 27 dogs, 35 duodenal images from 25 dogs, and 30 colonic images from 23 dogs were selected for study enrollment. Image selection was determined by joint agreement of authors JES and AEJ. Endoscopic procedures were performed using a commercial video endoscope (Olympus GIF-160, Olympus Optical, Tokyo, Japan) with still images of normal and abnormal mucosa captured by the endoscopist. The file size of the downloaded image was approximately 100 kb with a pixel array of 640 x 480 and 24-bit color. These still images were then arranged in a Powerpoint presentation for testing purposes.

Assessment of Images (initial test)

Still endoscopic images in Powerpoint format were assessed by 3 expert and 5 trainee endoscopists for inflammatory activity (hyperemia, granularity, friability, lymphatic dilatation, erosion) or interpreted as normal mucosa. Expert endoscopists were defined as individuals with advanced clinical training and active operator participation in GI endoscopy over the preceding 24 months. The trainee endoscopists had either minimal endoscopic training and/or lacked consistent endoscopic operator experience over the same 24 month period.

Images of each organ were randomized by means of a Research report randomizer program and assessed independently by each endoscopist. No clinical data or the date on which the image was taken was made available to the endoscopists. The endoscopic variables evaluated for the stomach included; granularity (6), friability (6), erosions (8), hyperemia (3),
normal mucosa (3) and normal post-biopsy mucosa (4), for the small intestine (previously reported); granularity (6), friability (6), erosions (7), lymphatic dilation (5), hyperemia (5), normal mucosa (4), normal post-biopsy mucosa (2), and for the colon; granularity (7), friability (6), erosions (7), mass (2), normal mucosa (5), and normal post-biopsy mucosa (3). Written definitions of each variable were made available to all endoscopists. If an individual image contained more than one mucosal abnormality, the endoscopist was asked to identify the dominant lesion.

**Selection of Images (validation test)**

Selected endoscopic video clips of a prospective group of 23 dogs diagnosed with inflammatory bowel disease at Iowa State University from 2011-2013 were reviewed from a computerized disc of the previously performed GI procedures. All 23 dogs had an upper GI endoscopy performed, and 10 of those dogs had a concurrent colonoscopy made available for review. All endoscopic procedures were performed using the same equipment.

**Assessment of Endoscopic Video (validation test)**

Video was recorded for each endoscopic procedure on the endoscopy computer hard drive and a separate compact disc for the prospective group of 23 IBD dogs. Approximately 5 minute representative video clips of the endoscopic procedure including mucosal biopsies were viewed for each organ (stomach, duodenum, and colon) by 2 expert endoscopists (JES & AEJ). Answers were recorded using a 0-2 point system (0=absent, 1=mild-moderate, 2=moderate-severe) for the following variables for each organ; stomach (max total 6 pts); granularity, friability, erosions, duodenum (max total 8 pts); granularity, friability, erosions,
and lymphatic dilation, and the colon (max total 6 pts); granularity, friability, erosions.

(Figure 1)

**Analysis of Data (initial test)**

Data was collected from each operator (3 expert and 5 trainee) using pre-designed Excel spreadsheets for statistical calculation. Fleiss Kappa coefficients were used to measure agreement within the expert and trainee groups. Additional analysis was performed using the Glimmix procedure. A mixed logistics regression model used mean values to compare inter-observer agreement between trainee and expert endoscopists. A $p$ value $< 0.05$ was considered statistically significant.

**Analysis of Data (validation test)**

Data from both expert endoscopists were recorded using a 0-2 point assessment for statistical calculation. Agreement between the two endoscopists were assessed using Cohen’s kappa coefficient for each score and each organ. Additionally, agreement between the two endoscopists was also assessed using a lesion present (1) or absent score (0). The difference in the distribution of the scores between the endoscopists was assessed using the test of symmetry.

**Results (initial test)**

Thirty endoscopic stomach images were obtained from 27 dogs. The dogs’ ages ranged from 1-13 years with a mean of 8 years. There were 14 spayed females, 8 neutered
males, 4 intact males, and 1 intact female included with the following pedigrees: 2 Weimaraners, 2 American Eskimos, 2 mixed breed, 2 Vizslas, 2 Labrador Retrievers, 2 West Highland White Terriers, and 1 each of Pomeranian, Bernese Mountain Dog, Wheaten Terrier, Sheltie, Chow Chow, Shar Pei, Pembrooke Welsh Corgi, English Bulldog, Cavalier King Charles Spaniel, Basset Hound, Samoyed, Boston Terrier, Miniature Schnauzer, Gordon Setter, and an Akita. All dogs exhibited chronic gastrointestinal signs. Fleiss Kappa statistics performed testing for agreement among the experts was $k=0.04$ and $k=0.15$ for the trainees. There was no significant difference when comparing the expert to the trainee group ($p=0.10$).

As previously reported, 35 duodenal endoscopic images were obtained from 25 dogs ranging in age from 1-11 years (mean 6.8 years). There were 11 spayed females and 14 neutered males included in the study. The following breeds were represented 4 West Highland White Terriers, 3 Golden Retrievers, 2 mixed breed dogs, 2 Boxers, 2 Labrador Retrievers, 2 Shih Tzus, 2 Yorkshire Terriers and 1 each; Wheaton Terrier, German Shepherd Dog, Vizsla, English Bulldog, Cocker Spaniel, Gordon Setter, Beagle, and a Miniature Poodle. Fleiss Kappa statistics performed testing for agreement among the experts was $k=0.01< and k=0.02$ for the trainees. There was a significant difference when comparing the experts to the trainees for duodenal evaluation ($p=0.05$).

Thirty colonic endoscopic images were obtained from 23 dogs ranging in age from 1-11 years (mean 6.8 years). There were 11 spayed females, 11 neutered males, and 1 intact male included in the study. The represented breeds were as follows; 4 Boxers, 3 Shih Tzus, 2 German Short Haired Pointers, 2 Labrador Retrievers, and 1 each of the following; West Highland White Terrier, Miniature Dachshund, Brittany Spaniel, Basset Hound, mixed breed, Weimaraner, German Wire Haired Pointer, Australian Shepherd, Rottweiler, Siberian Husky,
Beagle, Pembroke Welsh Corgi, and a Bichon Frise. Fleiss kappa statistics performed testing for agreement among the experts was k=0.22, and k=0.15 for the trainees. There was no significant difference when comparing the expert to the trainee group (\(p=1\)).

**Results (validation test)**

Twenty-three dogs with histopathologic confirmed inflammatory bowel disease were prospectively enrolled in a study. Their ages ranged from 1-14 years (mean 6.7 years). There were 14 spayed females, 8 neutered males, and 1 intact male included for evaluation. Their pedigrees were as follows; 4 West Highland White Terriers, 3 mixed breeds, 2 Shih Tzus, 2 German Shepherd Dogs, 2 Labrador Retrievers, 2 Boxers, and 1 each of the following; 1 German Short Haired Pointer, 1 Gordon Setter, 1 Boston Terrier, 1 Scottish Terrier, 1 Vizsla, and 1 Collie. All 23 dogs had stomach and duodenal evaluations performed, and 10 of these dogs also had a colonoscopy performed.

Comparisons were made between 2 expert endoscopists when making mucosal evaluations by watching 5 minute representative video clips of the stomach, duodenum, and colon. For the quantitative scoring system (0-2), there was almost perfect agreement when comparing total scores of the stomach assessment for each patient between expert endoscopists (weighted \(k=0.87\)). There was moderate agreement in total duodenal mucosal assessment between expert endoscopists (weighted \(k=0.6\)), and there was substantial agreement between expert endoscopists when making colonic mucosal assessments (weighted kappa=0.74). (Table 1)
When assessing individual variables for each organ, there was substantial agreement between the expert endoscopists when grading the stomach (erosion: $k=0.84$, granularity: $k=1$, friability: $k=0.8$). The small intestine showed moderate agreement between the expert endoscopists when assessing mucosal lesions (erosion: $k=0.84$, granularity: $k=0.66$, friability: $k=0.67$, lymphatic dilation: $k=0.62$), and the colon assessments showed substantial agreement between expert endoscopists (erosion: $k=1$, granularity: $k=1$, friability: $k=1$).

Similarly, for the qualitative comparison of a lesion being present or absent, there was almost perfect agreement when comparing total scores of the stomach assessment for each patient between expert endoscopists (weighted $k=0.87$). There was moderate agreement in total duodenal mucosal assessment between expert endoscopists (weighted $k=0.58$), and there was substantial agreement between expert endoscopists when making colonic mucosal assessments (weighted $k=0.8$). (Table 2)

**Discussion**

Canine inflammatory bowel disease is a major component in dogs with chronic enteropathies. Evaluation of endoscopic mucosal surfaces of dogs with IBD is routinely done during the diagnostic process. As a result, mucosal disease is an important measurement in identifying disease and monitoring treatment in dogs with IBD. There is a need for a simple, reproducible, and validated veterinary scoring index. To date, there have been no veterinary studies proposing and validating a canine endoscopic mucosal scoring index. This paper proposes a simplified endoscopic mucosal index and applies the index to a prospectively enrolled group of dogs with IBD to compare agreement between experienced endoscopists. Our results indicate that there is little difference within and among groups of trainee and
expert endoscopists when assessing mucosal lesions, and there is moderate to substantial agreement between endoscopists when evaluating mucosal lesions in dogs with IBD.

In human medicine, multiple indices have been proposed to assess and quantify mucosal damage, especially in relation to UC and CD.\textsuperscript{12-18} In fact the first endoscopic mucosal assessments and scores were made by Sidney Truelove in 1955, known as the founding father of human IBD therapeutics.\textsuperscript{13,14,19} Since then, several human indices have been proposed to measure the disease activity and severity of both CD and UC.\textsuperscript{12-18} At present, many human gastroenterologists utilize the Ulcerative Colitis Colonoscopic Index of Severity (UCCIS) and the Simplified Endoscopic Activity Score for Crohn’s Disease (SES-DC).\textsuperscript{16,18}

In veterinary medicine, the diagnosis of canine IBD is the result of excluding metabolic, infectious, neoplastic, or obstructive disorders of the gastrointestinal system.\textsuperscript{3,4,20} Determining endoscopic inflammatory activity is crucial for the assessment of disease and for tailoring appropriate therapy.\textsuperscript{3,4} However, there is a wide variation among endoscopists in the mucosal assessment of disease.\textsuperscript{21-23} As such, there is a need for a simplified veterinary endoscopic scoring index.

In our study, the mucosal lesions chosen for assessment were based on former veterinary and human indices developed for various chronic enteropathies.\textsuperscript{3,9,10,12-18} There was no significant difference within the trainee group when assessing the stomach, duodenum or the colon. For the expert group, there was little agreement when assessing the stomach (k=0.04). This may be due to lesion bias based on prior experiences, such as “over or under-interpreting” lesions. When comparing the expert to the trainee endoscopy scores, there was
no significant difference in lesion assessment for the stomach or the colon. As previously reported in chapter 3, there was a significant difference in the duodenal assessment between experts and trainees \( (p<0.01) \). Again, this is likely a result of the inexperience of the trainee endoscopists. Overall, the agreement rate between the expert and the trainee endoscopist were very good.

When the developed simplified endoscopic scoring index was applied to the prospective group of 23 dogs with IBD, there was almost perfect agreement between the expert endoscopists evaluating the stomach, moderate agreement when evaluating the duodenum, and substantial agreement with colonic evaluation. The difference in the mucosal assessments could be a result of the subjective nature of scoring mild-moderate (1 pt) versus moderate-severe (2 pts). However, when kappa statistics were performed comparing the endoscopists’ assessment of the presence of absence of a lesion, similar scores were obtained for the stomach, duodenum and colon. Overall, there is very good agreement between the endoscopists in lesion assessment.

There were several limitations to our study. First, still images instead of video streams were used for the expert and trainee endoscopists for initial development of the scoring system. This was done to reduce the length of time utilized for individual mucosal assessment. A still image also allowed us to highlight one specific lesion. Video streams would have added unnecessary length to the study as well as caused further problems with mucosal lesion interpretation. Therefore, standardized archived still endoscopic images were included in the study.
The number of expert and trainee endoscopists included in our study was low and may have affected our results. We feel that having several different people with various training backgrounds provided variability and diversity to each endoscopy group. A multi-institutional study to increase the numbers of the expert and trainee groups could be pursued in the future to further corroborate our results. It is worthy to point out that although the overall number of endoscopist invited to participate in the study was low, there have been no other veterinary studies comparing endoscopic mucosal assessments between endoscopists. The three previously published canine studies regarding mucosal scoring had limited numbers of evaluators/endoscopists (n=2) performing the evaluation, and no reports of agreement between their scores.\textsuperscript{3,9,10}

In summary, our simplified endoscopic activity score for canine IBD was successfully applied to a prospective group of dogs with histopathologically confirmed IBD. There was very good agreement between the expert endoscopists evaluating their mucosal lesions in the stomach, duodenum, and colon using a quantitative 0-2 point scale or using a simplistic qualitative scale of lesion absence or presence. Based on these results, we have determined that accurate mucosal endoscopic assessments between trained operators depend on detailed descriptions of a limited number of descriptive variables.

**Footnotes**

\textsuperscript{a}Powerpoint\textsuperscript{TM}

\textsuperscript{b}Research report randomizer

**References**


Figure 1: Quantitative lesion grading scale; 0=absent, 1=mild, 2=moderate/severe

<table>
<thead>
<tr>
<th>Organ</th>
<th>Erosion</th>
<th>Granularity</th>
<th>Friability</th>
<th>Lymphatic dilation</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stomach</td>
<td>0-2</td>
<td>0-2</td>
<td>0-2</td>
<td>NA</td>
<td>6</td>
</tr>
<tr>
<td>Duodenum</td>
<td>0-2</td>
<td>0-2</td>
<td>0-2</td>
<td>0-2</td>
<td>8</td>
</tr>
<tr>
<td>Colon</td>
<td>0-2</td>
<td>0-2</td>
<td>0-2</td>
<td>NA</td>
<td>6</td>
</tr>
</tbody>
</table>
Table 1: Validation study quantitative (scoring 0-2) agreement values (weighted kappa)

<table>
<thead>
<tr>
<th>Lesion</th>
<th>Stomach</th>
<th>Duodenum</th>
<th>Colon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erosions</td>
<td>0.9</td>
<td>0.9</td>
<td>1.0</td>
</tr>
<tr>
<td>Granularity</td>
<td>1.0</td>
<td>0.66</td>
<td>1.0</td>
</tr>
<tr>
<td>Friability</td>
<td>0.8</td>
<td>0.67</td>
<td>0.78</td>
</tr>
<tr>
<td>Lymphatic dilation</td>
<td>NA</td>
<td>0.62</td>
<td>NA</td>
</tr>
<tr>
<td>Total</td>
<td>0.87</td>
<td>0.6</td>
<td>0.71</td>
</tr>
</tbody>
</table>
Table 2: Validation study qualitative (lesion present or absent) agreement values (weighted kappa)

<table>
<thead>
<tr>
<th>Lesion</th>
<th>Stomach</th>
<th>Duodenum</th>
<th>Colon</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erosions</td>
<td>0.83</td>
<td>0.83</td>
<td>1.0</td>
</tr>
<tr>
<td>Granularity</td>
<td>1.0</td>
<td>0.77</td>
<td>0.8</td>
</tr>
<tr>
<td>Friability</td>
<td>0.91</td>
<td>0.68</td>
<td>0.78</td>
</tr>
<tr>
<td>Lymphatic dilation</td>
<td>NA</td>
<td>0.5</td>
<td>NA</td>
</tr>
<tr>
<td>Total</td>
<td>0.87</td>
<td>0.58</td>
<td>0.8</td>
</tr>
</tbody>
</table>
CHAPTER 5

GENERAL CONCLUSIONS

The goal of this study was to evaluate inter-observer endoscopic agreement and make a simple, novel mucosal endoscopic scoring system for dogs with IBD to help standardize the endoscopist’s descriptions of mucosal surfaces. This is the first veterinary study concerning this specific topic despite a long history of human medical interest in endoscopic activity scores specifically for patients with CD and UC.

The results of the inter-observer endoscopist agreement using Fleiss kappa statistics showed slight to fair agreement within the expert and trainee groups when evaluating the stomach, duodenum, and colon mucosa of dogs diagnosed with IBD. In duodenal assessment, there was significant improvement in the trainee endoscopist group when a visual template of duodenal lesions were provided with written descriptions of lesions ($p<0.01$). As a result, the trainees scored similarly to the expert endoscopist group. Therefore, we concluded that endoscopists benefitted from clearly defined lesion terminology and visual examples when making endoscopic mucosal assessments.

A simplified endoscopic scoring system was developed based on the results of the previous pilot data and numerous human endoscopic indices. Lesions including erosions, granularity, and friability were used to describe the stomach, duodenal, and colon mucosal surface with the addition of lymphatic dilation when making duodenal assessment. A score of 0-2 for each lesion in each organ was given; 0=no lesion, 1=mild, and 2=moderate to severe.
Total maximum score for the stomach and colon was 6, while the total maximum score for the duodenum was 8. Mild mucosal disease was defined as a total score of 0-2, 3-4=moderate disease and 5-6/8=severe disease. Two expert endoscopists evaluated representative video clips of a prospective group of dogs with IBD for scoring purposes. Based on Cohen’s kappa statistics, the experts had moderate to almost perfect agreement for stomach, duodenal, and colonic mucosal assessments.

Our recently developed simplified endoscopic mucosal scoring system is easy to use and was well received by the endoscopists participating in our study. The lesions chosen for inclusion in the mucosal assessment index were those most often encountered in diseased canine GI tracts. A visual template accompanied by written lesion descriptions are useful resources for veterinary endoscopists, especially those with limited experience. We hope that the results of this study will further compel fellow clinician scientists to critically evaluate how they make endoscopic mucosal assessment in their patients. A uniform means of identifying, describing, and reporting mucosal lesions benefit the patient, fellow veterinarians, and will help create more consistent medical record/endoscopy reports.

Based on our results, we conclude there is moderate to near perfect agreement of expert endoscopists using our simplified mucosal scoring index in dogs with IBD. Future studies are needed to evaluate the significance and association of mucosal lesion assessments and other markers for IBD disease severity. A future study is planned for the comparison of our recently developed mucosal endoscopic disease activity score and serum albumin levels, histologic assessment, CRP, and pANCA levels.