

## Review

---

# A Critical Review of “A randomized trial of the effects of the no-carrageenan diet on ulcerative colitis disease activity (*Nutr. Healthy Aging*. 2017; 4(2): 181–192).”

James M. McKim<sup>a,\*</sup>, Jamin A. Willoughby<sup>a</sup>, William R. Blakemore<sup>b</sup> and Myra L. Weiner<sup>c</sup>

<sup>a</sup>*IONTOX, Kalamazoo, MI, USA*

<sup>b</sup>*Celtic Colloids Inc., Topsham, ME, USA*

<sup>c</sup>*TOXpertise, LLC, Princeton, NJ, USA*

**Abstract.** Carrageenan (CGN) is a widely used food additive commonly found in dairy products, dairy substitutes (such as almond milk and soy milk), deli meats, nutritional supplements, beverages, and infant formula. It is typically used as a gelling, thickening, emulsifying and stabilizing agent and to improve the mouthfeel of the product. Decades of research on CGN safety has been reviewed by numerous regulatory bodies across the world. These regulatory bodies have all agreed on the safety status of CGN used as a food additive. Despite this, a few groups have continually published reports using misinformation and misinterpretation of results to suggest that CGN is harmful. The study reviewed here, “A randomized trial of the effects of the no-carrageenan diet on ulcerative colitis disease activity (Bhattacharyya et al. *Nutr. Healthy Aging*. 2017; 4(2): 181–192)” from Joanne Tobacman’s group, is an example of this. In this study, Bhattacharyya et al. set out to assess the impact of CGN in the diet on the interval to relapse in patients with ulcerative colitis (UC) who were currently in remission. For the study, UC patients in remission were instructed to participate in a “no carrageenan” diet. Some of the participants also received placebo capsules while others received CGN capsules. CGN, being a food additive, is mostly ingested as part of the diet, and typically bound to food protein. It is never ingested as a bolus in capsule form. The study contains errors in the interpretation of CGN literature and has cited studies that used degraded CGN, not food grade CGN. There are also issues with the use of statistical tests and data interpretation. In one instance, the authors seemed to ignore their own study results, suggested that CGN in the diet had no effect. Here we will review these issues as well as discuss how this study could have been approached to eliminate obvious bias.

**Keywords:** Carrageenan, poligeenan, degraded carrageenan, colitis, inflammation

### Abbreviations

CGN carrageenan is a high molecular weight polydisperse polysaccharide synthesized by certain species of red seaweed and

made up of repeating galactose and 3,6-anhydrogalactose units with variable ester sulfate content by degree and location.

SCCAI Simple Clinical Colitis Activity Index  
SIBDQ Short Inflammatory Bowel Disease Questionnaire  
UC Ulcerative colitis

---

\*Corresponding author: James M. McKim, IONTOX, LLC, 4717 Campus Drive, Kalamazoo, MI 49008, USA. Tel.: +1 269 372 3755; E-mail: jmckim@iontox.com.

## 1. Introduction

Carrageenan (CGN) is a naturally occurring food additive isolated from various species of red seaweeds of the class *Rhodophyceae*. CGN is a high molecular weight linear polymer with a backbone comprised of repeating galactose sugars with varying degrees of sulfate groups attached (sulfated polygalactan). There are three major forms of carrageenan; lambda- ( $\lambda$ ), kappa- ( $\kappa$ ) and iota- ( $\iota$ ). These three forms differ in the conformation of the galactose linkages and the level and location of sulfate groups [1, 2], characteristics which impart the thickening and stabilizing properties of CGN via ionic interactions with food protein [1]. Though the three primary forms of CGN ( $\lambda$ ,  $\kappa$  and  $\iota$ ) show no major differences in terms of toxicological effects and human safety in foods [3], the minor differences in their chemistry impart different functional properties that are very useful to the food industry when mixed.

The high molecular weight ( $M_w$ ) food additive CGN (200,000 – 800,000 Da.), also called food grade CGN, has been shown to be safe [4, 5]. Regulatory agencies around the world have found CGN safe for use as a food additive. In the United States, CGN is an approved food additive [6, 7] and CGN from *Chondrus* extracts are considered “generally recognized as safe,” or GRAS [8], while the World Health Organization has placed carrageenan in the best, safest possible category for a food additive. The International Agency for Research on Cancer (IARC) has stated that carrageenan is non-carcinogenic [9]. Finally, CGN has even been deemed safe for use in infant formulas by the Joint FAO/WHO Expert Committee on Food Additives (JECFA), stating that “the use of carrageenan in infant formula or formula for special medical purposes at concentrations up to 1000 mg/L is not of concern” [10]. The report by JECFA also stated that, in making their decision, they “took account of the previous toxicological database on carrageenan, which did not indicate other toxicological concerns” [10].

There has been significant confusion in the nomenclature surrounding CGN. This is due primarily to confusion between food grade CGN and the acid hydrolysis products of CGN, called degraded carrageenan (d-CGN; 20,000–40,000 Da.) and poligeenan (PGN; 10,000 – 20,000 Da.). These chemicals, d-CGN and PGN, are made in the laboratory by subjecting CGN to very low pH (0.9–1.3) and

non-physiological temperatures (>80°C) for several hours. Neither d-CGN nor PGN have been definitively shown to be produced in the gut *in vivo*, under normal physiological conditions, upon ingestion of food grade CGN. Neither PGN nor d-CGN are used as food additives. PGN has been used commercially but not as a food additive. It is well known that PGN can cause intestinal inflammation and toxicity [4, 5, 11]. Confusion in the scientific community is often based on incorrect nomenclature. Many studies published over the last several years have repeatedly used the term “carrageenan” in their manuscripts when the actual chemical assessed in the studies was either d-CGN or PGN. Although these nomenclature issues have been clarified [12–14], these issues continue to persist, confusing both the scientific community and consumers. This, in turn, has resulted in misinterpretations of toxicological data by research groups [15, 16] and by consumer groups [17, 18] whereby the toxicological effects observed with ingestion of d-CGN and PGN are incorrectly thought to be the toxicological effects of food grade CGN. Despite this, these groups continue their assertions that CGN, when used as a food additive, is harmful, despite their claims having been rejected by regulatory authorities [10, 19].

While food grade CGN has been repeatedly demonstrated to be safe for human ingestion, d-CGN and PGN are known to cause lesions in some animal models that resemble the pathophysiology of ulcerative colitis (UC), including inflammation and lesions [20–25]. UC occurs in approximately 1 in 1.3 million people in the United States [26, 27]. The disease marked by inflammation and lesions in the intestines, which appear to be precipitated by a complex interaction of dietary, environmental, and genetic factors, as well as immune response dysregulation [28–31]. One of the primary ways a patient can reduce UC symptoms is to optimize their diet and to stop eating foods that may worsen the symptoms. Because diet likely plays a role in the relapse potential of UC and the intensity of UC symptoms, Bhattacharyya et al. (2017) assessed whether a CGN may contribute to UC relapse in patients who are in remission. Their hypothesis is based entirely on the false premise that CGN causes UC-like inflammation and lesions. In this review, we will discuss these issues with the study design, statistics and data interpretation, as well as discuss how this study could have been approached to eliminate obvious bias.

## 2. Discussion

### 2.1. Study hypothesis, references and actual effects of CGN

The first and obvious questionable aspect of the Bhattacharyya et al. (2017) publication is the hypothesis. The authors suggest that because carrageenan “predictably causes inflammation” in cell-based and animal models, it is possible that CGN exposure would shorten the relapse time in patients with UC that is in remission. The issue with this hypothesis is that CGN does not cause inflammation, predictable or otherwise, nor does it induce lesions or any other known symptoms of UC. The reason the authors believe this is likely due to misunderstanding and misinterpretation of previous studies. For example, the authors state the following in the start of their Introduction:

*“The common food additive carrageenan has been used in thousands of experiments in animal models and in cultured human cells to cause inflammation [1–8]. Carrageenan-induced inflammatory effects in the intestine in animal models resemble the histopathology of ulcerative colitis [6–8].”*

Reviewing the studies cited, it is possible to see where the confusion lies. Numerous studies done in the preceding decades used either d-CGN or PGN as a test material, and the authors of the Bhattacharyya et al. (2017) publication are applying the effects observed with PGN or d-CGN to CGN. This confusion is understandable as d-CGN and CGN were both often called “carrageenan” both in the title and in the body in numerous early works. The Materials and Methods are often the only place where it can be discerned whether CGN was subjected to acid hydrolysis prior to application, which would have made d-CGN and PGN. Referring to d-CGN or PGN as simply CGN is a problem that still exists in the literature to this day, thus in-depth analysis of each reference is key. The very first reference cited in the Bhattacharyya et al. (2017) publication is the same IARC Monograph cited in the introduction of this review. In this IARC Monograph, the authors state the following about Native (Undegraded) Carrageenan:

*“Groups of 30 male and 30 female MRC outbred rats, seven weeks old, were given 0.5, 2.5 or 5% carrageenan (Gelcarin HMR, largely composed of kappa components, high molecular weight) in*

*the diet for life. A group of 100 male and 100 female rats served as controls. Survival of the treated animals was not different from that of controls. No significant increase in tumour incidence was seen (Rustia et al., 1980).*

*Hamster: Groups of 30 male and 30 female Syrian golden hamsters, seven weeks old, were given 0.5, 2.5 or 5% carrageenan (Gelcarin HMR, largely composed of kappa components, high molecular weight) in the diet for life. A group of 100 males and 100 females served as controls. Survival was similar in treated and control groups. No significant increase in tumour incidence was seen (Rustia et al., 1980).”*

The cited IARC Monograph also states the following:

*“Several studies in which native carrageenan was administered to animals (guinea-pigs, rabbits, rats, mice, pigs, rhesus monkeys) in drinking-water (e.g., 0.5–2% for up to 14 weeks) or in the diet (e.g., 2–5% for 12 weeks) showed no ill effect (Maillet et al., 1970; Abraham et al., 1972; Benitz et al., 1972, 1973; Poulsen, 1973). However, in one study (Grasso et al., 1973) pin-point caecal and colonic ulcerations were found in guinea-pigs given 5% native carrageenan in the diet for 3–5 weeks.”*

This is the only noted mention in the entire IARC document that suggests adverse effects from CGN. These effects, observed at 5% CGN in the diet, are a consequence of a much higher concentration of CGN than used in real world applications. In fact, as a food additive, CGN is typically used at concentrations of 0.01 – 3.5% [32]. The study by Grasso et al. [23] additionally states that humans with colorectal cancer that were provided 5 grams of CGN daily for 10 days exhibited no sign of ulceration, nor was d-CGN detected in the colon. The authors of the Bhattacharyya et al. (2017) publication are clearly taking the known toxicological effects of d-CGN and PGN and simply applying them to CGN, which is a different polymer, both in terms of physiochemical effects and toxicological profiles.

The third reference in the Bhattacharyya et al. (2017) publication is a book chapter written by Blake-more and Harpell [1] and also referenced in this review. There is no discussion of the toxicological effects on CGN on human, animal or cell-based models in this reference, though the manufacturing, regulation, chemical structure and how that related

to the functional properties of CGN as well as its use as a food additive are discussed. The remaining 6 references cited in the first two sentences of the Introduction of the Bhattacharyya et al. (2017) publication either used d-CGN or PGN to induce lesions, or are reviews of CGN that also incorrectly cite studies using d-CGN or PGN to induce lesions.

When we begin to review the numerous studies that actually used food grade CGN in animal feeding studies, not the acid hydrolysis products PGN and d-CGN, we see a very different safety profile. For example, Abraham and Coulston [24] found that animals exposed to high concentrations CGN in the water and diet (1–5%, respectively) exhibited no adverse effects on the gastrointestinal tract (GI tract) compared to control animals. These results were also reported in other species including monkeys, rats, and Guinea pigs in the same study. Histopathological analysis was performed on PGN-induced lesions and the results suggested that the immune system plays a role in lesion formation. Rhesus monkeys, rats, and Guinea pigs that were fed PGN in their diets (0.5%, 1% or 2% PGN) for 7–14 weeks developed mucosal ulcerations marked with infiltration of macrophages and polymorphonuclear cells (PMNs) into the lamina propria. These effects were not observed with animals fed CGN [21]. Additionally, Song, Kim and Kim [33] provided Guinea pigs 1.5% CGN in drinking water for 40 days and found no GI lesions. Clearly the literature, when properly assessed, support global regulatory agencies' decisions that CGN is safe for human consumption. In fact, it is possible that there are some health benefits to CGN ingestion, including decreased cholesterol and low-density lipoproteins (LDL) and increased immune status parameters with a concomitant decrease in inflammation biomarkers in human volunteers [34–36]. It is apparent that the authors of the Bhattacharyya et al. (2017) publication either do not understand the toxicological differences between CGN and d-CGN/PGN.

## 2.2. Study design and participants

The criteria for patient acceptance into the Bhattacharyya et al. (2017) study was relatively strict. The patients had to be over 18 years of age, be diagnosed with UC as confirmed by biopsy, have had a previous need for corticosteroids to obtain remission, and they must have been off the corticosteroids and in remission for at least one month prior to the start of the study in December 2012. Finally, the sub-

jects must have had a Simple Clinical Colitis Activity Index (SCCAI) score of 2 or less, either be on no current medications, or a stable dose of medication, and be willing to try their best to follow a carrageenan-free diet. It is important to note that patients could be on stable doses of powerful anti-inflammatory or immune modulation medications and be considered viable participants, as long as they were no longer taking corticosteroids. Patients were also excluded if they were not currently in remission, unable to read the labels on food products to determine if CGN was included, or were unable to make choices on what they ate. Once selected to participate in the study, patients were randomized into two groups using a randomization website or by the participants year of birth, group 1 (capsules containing CGN) and group 2 (capsuled containing dextrose placebo). Upon starting the trial, patients completed two questionnaires via telephone with the study personnel every two weeks. These questionnaires were the SCCAI and the Short Inflammatory Bowel Disease Questionnaire (SIBDQ). Finally, every 3 months, blood and stool samples were collected from the patients at a clinic, however, during the course of the study, this was amended so that stool samples could be provided by mail and the blood draw was discontinued. Patients participated in the study for either one year, until they relapsed, or until they voluntarily dropped out of the study.

Sample size calculations performed by Bhattacharyya et al. (2017) estimated that a population size of 36 individuals would be required in order to achieve 80% power to detect a SCCAI score difference of at least 1.0 point between the two groups (placebo vs CGN). A total of 15 people agreed to participate in the study, however 3 dropped out prior to initiation of the study as they did not want to comply with the study procedures. This left 12 participants, one third of the total number of 36 that was originally determined by the authors to be necessary for statistical analysis. Despite this obvious issue in group size, the authors decided to continue to move forward with the study. The fact that the authors ignored their own stated participant numbers necessary to achieve 80% predictive power is the first sign of a potentially improperly designed study and immediately calls into question the robustness of the results. The authors of the Bhattacharyya et al. (2017) publication did not state what the percent power to detect a SCCAI score difference of at least 1.0 points was between the two groups when the total number of participants was dropped from 36 to 12. This information should have

been presented and discussed. In addition, the authors state that “. . . in view of the larger than anticipated difference in the mean SCCAI scores between the two groups . . . a smaller sample size than was originally proposed ( $n = 36$ ) could have been targeted.” However, this information couldn’t have been known or anticipated prior to initiating the study with a sample size that is one-third of the appropriate sample size. In addition, the data analysis and statistics presented for the SCCAI data, the same data used to make the decision that the smaller participant size was justifiable, are questionable. The statistics will be discussed later in this review. Regardless, this study should not have continued and the results should be viewed as highly suspect, and discussed as such, due to improper number of participants to achieve a minimum 80% power to detect a SCCAI score differences of at least 1.0 points between the two groups, per the authors initial calculations of the acceptable study size.

The 12 remaining participants were broken into two groups. The first group contained 5 individuals who received the CGN capsules while the second group contained 7 individuals who received the dextrose placebo capsules. The subjects were initially taking a single 100 mg capsule a day, but part way through the study, were instructed to double the dosage to two 100 mg capsules per day for a total of 200 mg. It is not known whether the participants were instructed to take both the capsules at the same time or were instructed to take a single capsule at two different times daily. All subjects received standard medical care from their own physicians and continued their maintenance medications, though they were not allowed to increase dosage of their medications or change medications they were on during the study. Although the two groups (CGN vs placebo) were randomized, it is important to point out that 10 of the 12 patients that took part in this study were on strong anti-inflammatory/immune modulation medications in order to maintain the remission of their colitis (Humira, Asacol, Sulfasalazine, Imuran). Two of the 12 participants were not on anti-inflammatory medications. This discrepancy was never called out or discussed by the authors of the Bhattacharyya et al. (2017) publication. If we assume the hypothesis posited by Bhattacharyya et al. (2017), that CGN “predictably” causes inflammation, is correct, the participants in this study were randomly grouped in a fashion that could only lead to proving their hypothesis correct, since the two participants that were not using powerful anti-inflammatories to maintain their remission were both in the group receiving CGN cap-

sules daily. This seems like a very important variable to the study hypothesis presented by Bhattacharyya et al. (2017). To properly control for this, these two individuals should have been left out of the study, as their lack of anti-inflammatory use would make them much more likely to relapse. At the very least, one of these patients should have been placed into each group. Three individuals in the CGN group relapsed and one of the individuals that relapsed was one of the two individuals not taking anti-inflammatory medication. If this individual was controlled for (or removed from the study) it would be interesting to see if there was still a statistical significance between the groups in terms of relapse given the small sample size of the study. Two of the three participants who experienced a relapse were on anti-inflammatory medications. One potential explanation for this is that carrageenan, with its high binding properties, could be binding their anti-inflammatory drugs and reducing efficacy hence leading to a relapse. This variable was not measured and was not controlled. In any event, the number of participants was far too small, per the authors own initial calculations, and the participants were randomized in a manner that favored a relapse in the CGN group, regardless of whether CGN was the reason for the relapse.

### 2.3. Route of exposure

Another questionable aspect of the study design was the route of CGN administration. CGN is used almost exclusively as a food additive, providing excellent gelling, thickening and emulsifying properties to food. The properties are imparted via water gel strength and protein reactivity. Gelation of CGN involves a heating and cooling cycle in the presence of specific cations, depending on the type of CGN used. For example, potassium ions are crucial for  $\kappa$ -CGN gelation and calcium ions are crucial for  $\iota$ -CGN gelation [1, 2]. CGNs also strongly react with proteins forming stable 3-D structures at concentrations as low as 50  $\mu\text{g}/\text{mL}$  [1, 2]. These stable 3-D structural combinations change slightly as the concentration of CGN is increased. For example, “stabilization” occurs at  $\sim 50 \mu\text{g}/\text{mL}$  (e.g. canned evaporated milk) while “stabilization + thickening” occurs at  $\sim 300 \mu\text{g}/\text{mL}$  (e.g. chocolate milk). As the concentration of CGN increases to  $\sim 750 \mu\text{g}/\text{mL}$  “stabilization + additional thickening” occurs (e.g. milk shakes) and finally, “stabilization + gelation” occurs at  $\sim 2,000 \mu\text{g}/\text{mL}$  (e.g. puddings). Note that

the strength and number of CGN-protein interactions are very much dependent on the pH of the application. For example, as pH of the application is decreased, the positive charge density on the protein increases, resulting in a stronger attraction to CGN via the negatively charged sulfate groups. When approaching the isoelectric point (IEP) of the protein at about pH 4.5 (and below), the CGN-protein structure, with increased ionic interactions, is strengthened to the point that the CGN-protein complex will precipitate. This means that the CGN remains bound to the protein until it reaches further down into the lower GI tract, and is only released from the protein after the protein has been at least partially broken down by gut enzymes. This explains why CGN is almost completely excreted intact in feces [37, 38] and not digested into smaller fragments.

Since the authors of the Bhattacharyya et al. (2017) publication wanted to assess the effect of a no-CGN diet on the timing of the relapse of UC patients, it seems that the optimal exposure scenario would be to provide either CGN or the placebo in the diet so that each participant in the study ingested 200 mg/day CGN, but in the context of its use as a food additive, bound to food protein or as a gel. Instead, however, CGN was provided as a 200 mg bolus every day (two 100 mg capsules) with a meal. This author's choice of dosing is similar to the dosing seen with supplements, however, carrageenan is not available as a supplement. It is occasionally used as an excipient in pharmaceuticals, however, the concentrations for use as an excipient (typically 0.1 – 1%) [39] are less than when carrageenan is used as a food additive so the overall carrageenan exposure is insignificant. Needless to say, this bolus dose is not how the human stomach nor the human GI tract is exposed to carrageenan in real world exposures, as carrageenan is not ingested as a pure material in a bolus fashion. It is ingested in a gel formation or bound to food protein. This is very important because upon binding food protein or forming a gel, the CGN conformation changes drastically and stays bound to the food protein through the GI tract. It would have been relatively easy to provide 200 mg carrageenan daily either in a gel formation or bound to food proteins to mimic real world exposures. For example, the authors could have provided food samples to the subjects where each participant in group two gets exactly 200 mg of dextrose (placebo) formulated with a small sample of food once daily, and the participants in group one get 200 mg food grade carrageenan formulated with a small sample of food once daily. This expo-

sure regimen could have even been done in capsule form and would have been a far more relevant exposure. Unfortunately, the route of administration chosen by the authors of the Bhattacharyya et al. (2017) publication leaves CGN free to interact with proteins and medications, potentially reducing the efficacy of the medication by altering absorption or metabolism. Any results obtained from this type of study design should be viewed alone, with no correlation to real world exposures. Despite this glaring difference in ingestion of CGN, as well as the conformational differences in CGN when bound, Bhattacharyya et al. (2017) attempted to correlate a bolus ingestion of unbound pure food grade CGN to real world exposures in the Western diet (e.g. bound to food protein). This correlation is false since in the Western diet, CGN is almost exclusively ingested as a food additive in a gel or bound to food protein [1, 2].

#### 2.4. SCCAI vs SIBDQ

Bhattacharyya et al. (2017) assessed for relapse in patients in part using two questionnaires that are commonly used to assess ulcerative colitis. The first questionnaire was the Simple Clinical Colitis Activity Index (SCCAI) and the second questionnaire was the Short Inflammatory Bowel Disease Questionnaire (SIBDQ). Interestingly, only the SCCAI was used as a valid endpoint to determine relapse. This is particularly troubling considering the SIBDQ results showed no difference between group 1 receiving CGN capsules and group 2 receiving the dextrose-containing placebo capsules. The results for the SIBDQ was shown in a figure alongside the SCCAI results, however the results were merely mentioned in a single sentence in the results section. The fact that only one of the two tests, the SCCAI, was discussed further and that the SCCAI happened to be the only one of the two questionnaires used to define relapse is also evidence of questionable study design, considering only the SCCAI data resulted in a statistically significant difference ( $p = 0.05$ ) between the placebo group and the CGN group. At the very least, Bhattacharyya et al. (2017) should have recognized this discrepancy and discussed the reasoning for leaving the SIBDQ results out of the final data used to determine relapse.

Both the SCCAI and the SIBDQ are short questionnaires (~5–10 questions) that can easily be completed by the patients that cover important quality of life factors involved with UC. The resultant scores

of the questionnaires can then be used to determine disease activity. The decision to only use and discuss the results from the questionnaire that led to the desired positive response was highly questionable, particularly because one test showed positive results while the other test did not. Therefore, the decision to use the SCCAI scores instead of the SIBDQ scores seems arbitrary. Both tests are widely used and reliable, and the SCCAI does take additional information into account such as such as urgency, incontinence and nocturnal bowel movements. However, the SCCAI is not validated, nor is it useful for assessment of Crohn's disease (CD) which is often confused with UC [40, 41]. The SIBDQ, on the other hand, is a questionnaire that has been validated in different populations and can be used reproducibly for both UC and CD [41–43]. In addition, the SIBDQ gives results that are robust and comparable to the full 32-item Inflammatory Bowel Disease Questionnaire (IBDQ) which is the gold standard for measuring health-related quality of life in adult patients with inflammatory bowel disease, UC, or CD [44]. The fact that the short questionnaire that most directly correlates to the gold standard was not selected by Bhattacharyya et al. (2017) as a measurement of relapse is further evidence of poor study design.

Another important part of the SCCAI data that should have been discussed further in Bhattacharyya et al. (2017) was that one of the patients in the placebo group actually had an increase in SCCAI of greater than 2 points. That is part of the criteria set forth by Bhattacharyya et al. (2017) to determine relapse. However, because the individual had not requested an increase in treatment from their physician, this was not defined by the authors as a relapse. This brings numerous questions to mind, such as why did the patient not receive treatment? Did the patient intend on receiving treatment before the study was over? Clearly, this patient did have a relapse, it just was not scored as such because of the author's defined criteria. Additionally, when did the increase in SCCAI occur during the course of the study? If the increase in SCCAI score occurred just as the study was concluding, perhaps the subject wasn't able to receive medical attention and was therefore considered to have not relapsed. Would the log-rank statistical test used to determine statistical significance between the placebo and CGN groups still show a statistical significance ( $p=0.05$ ) if this subject who received the placebo and clearly had an increase in UC symptoms and SCCAI score was included as a positive for relapse? These holes and discrepancies should,

at very least, have been discussed further, yet were largely ignored.

Finally, the scoring of the SCCAI and correlation to relapse utilized by Bhattacharyya et al. (2017) in this study is also questionable. The authors stated that the patients in their study had to start with an SCCAI score less than two, and remission was defined as an increase in 2 or more points in the SCCAI followed by an increase in treatment. Interestingly, in a validation study of the SCCAI by Jowett, Barton and Welfare [45] it was found that an SCCAI score of 5 or more (out of a maximum of 19) defined relapse with 92% sensitivity, 91% specificity, 85% positive predictive value (PPV) and 89% negative predictive value (NPV). The "relapse" group in the Bhattacharyya et al. (2017) publication had a mean SCCAI score of 4.20. This is 0.8 points away from the traditional score of 5 points that defines relapse. Why Bhattacharyya et al. (2017) chose to use a two point swing in SCCAI to define relapse is questionable and unknown, particularly when the article cited to support this decision [40] does not discuss a two point swing being related to remission, and a previous publication [44] showed that a total of 5 points is required for remission with optimal statistical accuracy.

## 2.5. Data analysis and statistics

Another of the questionable areas of the Bhattacharyya et al. (2017) is the data presented and the use of statistics. One of the main endpoints of the study, though interestingly not used to determine relapse, was the level of calprotectin present in fecal samples and the presence of six pro-inflammatory markers; interleukin 6 (IL-6), interleukin 8 (IL-8), tumor necrosis factor alpha (TNF $\alpha$ ), monocyte chemoattractant protein 1 (MCP-1) nuclear factor kappaB (NF- $\kappa$ B) and b-cell leukemia/lymphoma 10 (BCL10). The results show that there was no difference between the carrageenan group and the placebo group for 5 of the 6 pro-inflammatory markers (IL-8, TNF $\alpha$ , MCP-1, NF- $\kappa$ B and BCL10). There was a statistical difference between the carrageenan group and the placebo group when assessing IL-6 ( $p=0.02$ ). The authors also state that there is a statistically significant difference between the two groups when looking at fecal calprotectin levels ( $p=0.06$ ) despite stating in the Materials and Methods that for statistical analysis using the *t*-test, *p* values  $<0.05$  were the cutoff for statistical significance. For most scientific analysis, and in particular when using a simple two-

tailed *t*-test as the authors did, statistical significance is often cut off at  $p < 0.05$ . Any *p*-value larger than 0.05 is not considered significant as *t*-tests have an inherent error of 5% (type 1 error). This calls into question the validity or interpretation of the fecal calprotectin result. In general, a *p*-value of  $p = 0.06$  would not be considered a statistically significant difference. The results of the *t*-test claimed the fecal calprotectin levels were statistically significant from each other even though the *p*-value was 0.06, contradicting their own Materials and Methods where they state statistical significance determined when  $p < 0.05$ . In addition, it was noted that the same patient who had high values of calprotectin and IL-6, and was the primary cause for the statistical significance in fecal calprotectin and IL-6 between the two groups, was also one of the patients in the carrageenan group who was not on any anti-inflammatories, and therefore should have been removed from the study or, at a minimum, properly controlled for. This immediately makes all the inflammation data questionable at best, and further warrants the removal of the two subjects not taking anti-inflammatories from the inflammatory portion of the study, if not the entire study.

The results of inflammatory markers and fecal calprotectin analysis in the stool samples are not the only questionable use of statistics in this study. The authors use a *t*-test in a number of comparisons throughout the study. This was done using *p*-values greater than 0.05 as statistically significant, despite the fact that *t*-tests are not very stringent due to the expected high level or type 1 error (5% at minimum), and despite the fact that the authors admit this rate of error, and the small sample size of the study. Regardless, the authors still used *t*-tests with very lenient cutoffs for statistical significance throughout the study, including in the SCCAI results. In fact, it was used to analyze almost all of the data presented. One exception appears to be a re-analysis of the SCCAI scores (the initial analysis used a *t*-test, as previously mentioned). The final decision on whether group 1 (CCGN) and group 2 (placebo) were statistically significant from one another was made using the log-rank test (also called the Mantel-Cox test). In fact the authors immediately ended the trial as soon as a statistically significant difference was observed between group 1 (CCGN) and group 2 (placebo) when using the log-rank test. The log-rank test is typically used to compare the survival distribution of two samples. Technically, while log-rank test is a much better statistical method for this study compared to the unpaired *t*-test, it still may not be the best analysis for this study for two primary

reasons. First, a log-rank test is appropriate to use when the data are censored (a condition in which the value of a measurement is not fully known) and right skewed (when a participant drops out of the study). While there were dropouts in the study, making it a right skewed study, the study did not use censored data. The data used as the endpoint was the SCCAI scores, which provides a numerical value between 0 and 19. These are not censored data as obviously the values of the measurements are known. Secondly, the log-rank statistical analysis may not have been the best option as described by Berty, Shi and Lyons-Weiler [46]. Berty, Shi and Lyons-Weiler [46] found that log-rank tests are often far too permissive of a statistical test, meaning the test allows for significant results of survivorship distribution prediction models that have low accuracy. It seems that the most appropriate statistical test to be used in this study is the  $F^*$  test, which, under the conditions of this study, may have been a far more accurate test for statistical significance. It would be interesting to have the raw data so that statistical significance in relapse between groups 1 and 2 could be assessed using an  $F^*$  test.

The last questionable result to discuss is the statistical analysis (*t*-test) presented in regards to the SCCAI scores. Without being able to analyze the raw data ourselves, we can only use the results put forward by the authors and model what the individual SCCAI scores are, though this is not difficult as the mean and standard deviations (or standard error of the means [SEM]) are presented. Bhattacharyya et al. (2017) determined the mean  $\pm$  SD/SEM of the SCCAI scores for the CGN group to be  $4.20 \pm 3.70$  and  $0.86 \pm 1.46$  for the group who received the placebo. It is troubling that the authors claim there is statistical significance between the two groups when the standard deviation (or standard error of the mean; it was not specified which was used) is larger than the mean for the placebo group, and the SD/SEM for the CGN group is 88% of the mean. We have modeled a number of scenarios that could lead to these values (where the mean is correct and the  $\pm$  value is the standard deviation as well as scenarios where the mean is correct and the  $\pm$  value is the standard error of the mean). There was not a single possible combination of individual replicate SCCAI scores where the mean of group 1 (CGN; 4.20 with  $n=5$ ) and the mean of group 2 (placebo; 0.86 with  $n=4.20$ ) were statistically significant, regardless of whether the  $\pm$  values were SD or the SEM of the means. The test yielding the closest result to that presented in the study was a two-tailed *t*-test assuming equal



variance, resulting in a  $p$ -value of 0.0529. This is not a statistically significant difference ( $p < 0.05$ ). In addition, Bhattacharyya et al. (2017) stated in the Materials and Methods that for statistical analysis using the  $t$ -test,  $p$  values  $< 0.05$  were the cutoff for statistical significance. The authors then claim that a  $p$ -value = 0.05 is statistically significant, once again claiming significance when it contradicts their Materials and Methods statement that significance is when  $p < 0.05$ . Perhaps if the authors presented their raw results in a table and described the software used to assess statistical significance, it would help clarify these glaring errors.

### 3. Conclusion

Clearly, there are a number of questionable decisions made with regards to the study design, data interpretation and data analysis. Examples include the low number of participants, the method of CGN administration, which had little to do with how consumers would actually be exposed, and the inconsistent statistical analysis employed by the authors. In addition, there was no discussion from the authors regarding the inclusion of the only individuals in the study NOT taking powerful anti-inflammatories/immune modulating drugs being placed in the same CGN treatment group, especially when inflammation is a key component of the author's hypothesis. The authors also chose to completely ignore the SIBDQ scores and only focused on the SCCAI scores as they seemed to exhibit statistics that favored the authors desired study outcome. Another study result that should have been discussed was the lack of a statistically significant inflammatory response in patients receiving the CGN pills, especially when inflammation is hypothesized to be critical in CGN induction of colitis and IBD. Finally, a subject receiving a placebo had an increase in SCCAI score that should have counted as a positive for relapse, but wasn't counted since the patient did not receive care from their physician. Taken together, these issues result in a compromised study of questionable design.

### Acknowledgments

The authors would like to acknowledge Eunice Cuirle, Manager of Global Regulatory Affairs FMC

Health and Nutrition business unit of FMC Corporation, for her valuable comments and support.

### Funding

The authors report no funding.

### Conflict of interest

The authors have no conflict of interest to report.

### Disclosure

Financial support for the preparation of this review was provided by Marinalg International.

### References

- [1] Blakemore WR, Harpell AR. Carrageenan. In Imeson A, Ed. Food Stabilisers, Thickeners and Gelling Agents, Wiley-Blackwell, Blackwell Publishing Ltd., Hoboken, NJ. 2010; pp. 73-94.
- [2] Blakemore WR. Polysaccharide Ingredients: Carrageenan. Reference Module in Food Sciences. 2015; Elsevier, pp. 1-8. <http://dx.doi.org/10.1016/B978-0-08-100596-5.03251-0>.
- [3] Joint FAO/WHO Expert Committee on Food Additives (JECFA). Toxicological evaluation of certain food additives, including anticaking agents, antimicrobials, antioxidants, emulsifiers and thickening agents. WHO Food Addi Series. 1999;42:147-69.
- [4] McKim JM. Food additive carrageenan: Part I: A critical review of carrageenan *in vitro* studies, potential pitfalls, and implications for human health and safety. Crit Rev Toxicol. 2014;44:211-43.
- [5] Weiner ML. Food additive carrageenan: Part II: A critical review of carrageenan *in vivo* safety studies. Crit Rev Toxicol. 2014;44:244-69.
- [6] FDA. Food and Drugs: Food Additives. 2017; 21 CRF 172.620.
- [7] FDA. Food and Drugs: Food Additives. 2017; 21 CRF 172.626.
- [8] FDA. Food and Drugs: Food Additives. 2017; 21 CRF 182.7255.
- [9] IARC: International Agency for Research on Cancer. IARC Monographs on the evaluation of the carcinogenic risk of chemicals to humans. Some food additives, feed additives and naturally occurring substances. IARC Monographs. 1983;31:79-94.
- [10] Joint FAO/WHO Expert Committee on Food Additives (JECFA). Safety evaluation of certain food additives, WHO Food Additives Series 70. In: Prepared by the Seventy-ninth Meeting of the JECFA. 2015. [http://apps.who.int/iris/bitstream/10665/171781/3/9789240693982\\_eng.pdf?ua1/41](http://apps.who.int/iris/bitstream/10665/171781/3/9789240693982_eng.pdf?ua1/41)
- [11] McKim JM, Willoughby JA Sr., Blakemore WR, Weiner ML. Clarifying the Confusion Between Poligeenan,

- Degraded Carrageenan, and Carrageenan: A review of the chemistry, nomenclature, and *in vivo* toxicology by the oral route. *Crit Rev Food Sci Nutr.* 2018;14:1-70. <http://dx.doi.org/10.1080/10408398.2018.1481822>
- [12] Weiner ML, McKim JM, Blakemore W. Addendum to “Weiner ML. Parameters and Pitfalls to Consider in the Conduct of Food Additive Research, Carrageenan as a Case Study. *Food Chem Toxicol.* 2016;87:31-44.” *Food Chem Toxicol.* 2017;107(Pt A):208-14.
- [13] Weiner ML. Parameters and pitfalls to consider in the conduct of food additive research, carrageenan as a case study. *Food Chem Toxicol.* 2016;87:31-44.
- [14] Beattie IA, Blakemore WR, Dewar ET, Warwick MH. Corrigendum to “A study of orally-administered degraded carrageenan in the baboon” [*Food Cosmet. Toxicol.* 8 (1970) 257-266]. 2015; *Food Chem Toxicol.* 75:189.
- [15] Tobacman JK. Review of harmful gastrointestinal effects of carrageenan in animal experiments. *Environ Health Perspect.* 2001;109:983-94.
- [16] Tobacman JK. The common food additive carrageenan and the alpha-gal epitope. *J Allergy Clin Immunol.* 2015;136:1708-9.
- [17] Cornucopia Institute. Carrageenan: How a “natural” food additive is making US sick. A report by the Cornucopia Institute. March, 2013. [www.cornucopia.org](http://www.cornucopia.org).
- [18] Cornucopia Institute. Carrageenan: New Studies Reinforce Link to Inflammation, Cancer, and Diabetes. Updated report by the Cornucopia Institute. April, 2016. [www.cornucopia.org](http://www.cornucopia.org).
- [19] Department of Health and Human Services. Citizen Petition, Food and Drug Administration. 2012; Docket No. FDA-2008-P-0347.
- [20] Watt J, Marcus AJ. Ulcerative colitis in the guinea pig caused by seaweed extract. *J Pharmaceut Pharmacol.* 1968;1:187S-188S.
- [21] Benitz KF, Abraham R, Golberg L, Coulston F. Carrageenan: An ulcerative agent? *Tox App Pharm.* 1972;22:282.
- [22] Fabian RJ, Abraham R, Coulston F, Goldberg L. Carrageenan-induced squamous metaplasia of the rectal mucosa in the rat. *Gastroenterol.* 1973;65:264-76.
- [23] Grasso P, Sharratt M, Carpanini FMB, Gangolli SD. Studies on carrageenan and large-bowel ulceration in mammals. *Food Cosmet Toxicol.* 1973;11:555-64.
- [24] Abraham R, Coulston F. Ulcerative lesions due to carrageenan. *Z Gastroenterol.* 1979;17:S154-S158.
- [25] Dworkin BM, Deschner EE, Fath Jr. RB, Winawer SJ. Degraded carrageenan as a model for acute right-sided colitis in mice. *Gastroenterol.* 1982;82:1048.
- [26] Loftus EV Jr. Clinical epidemiology of inflammatory bowel disease: Incidence, prevalence, and environmental influences. *Gastroenterol.* 2004;126:1504-17.
- [27] Kappelman MD, Rifas-Shiman SL, Kleinman K, Ollendorf D, Bousvaros A, Grand RJ, Finkelstein JA. The prevalence and geographic distribution of Crohn’s disease and ulcerative colitis in the United States. *Clin Gastroenterol Hepatol.* 2007;5:1424-9.
- [28] McAlindon ME, Hawkey CJ, Mahida YR. Expression of interleukin 1 beta and interleukin 1 beta converting enzyme by intestinal macrophages in health and inflammatory bowel disease. *Gut.* 1998;42:214-9.
- [29] Pizarro TT, Michie MH, Bentz M. IL-18, a novel immunoregulatory cytokine, is up-regulated in Crohn’s disease: Expression and localization in intestinal mucosal cells. *J Immunol.* 1999;162:6829-35.
- [30] Hanauer S. Inflammatory Bowel Disease: Epidemiology, pathogenesis and therapeutic opportunities. *Inflamm Bowel Dis.* 2006;12:S3-9 (Suppl 1).
- [31] Baumgart DC, Sandborn WJ. Inflammatory bowel disease: Clinical aspects and established and evolving therapies. *Lancet.* 2007;369:1641-57.
- [32] Moirano A. Sulfated seaweed polysaccharides. In Graham HD, Ed. *Food Colloids*, AVI Publishing Co., Westport, Connecticut. 1977; pp. 347-81.
- [33] Song IS, Kim NY, Kim CY. Ulcerative colitis induced by carrageenan in Guinea pig. *Gastroenterol.* 1992;102:A699.
- [34] Shiau SY, Chang GW. Degraded carrageenan increases permeability of phenol red in the rat intestine. *Fed Proc.* 1981;43:1065.
- [35] Panlasigui LN, Baello OQ, Dimatangal JM, Dumelod BD. Blood cholesterol and lipid-lowering effects of carrageenan on human volunteers. *Asia Pac J Clin Nutr.* 2003;12:209-14.
- [36] Sokolova EV, Bogdanovich LN, Ivanova TB, Byankina AO, Kryzhanovsky SP, Yermak LM. Effect of carrageenan food supplement on patients with cardiovascular disease results in normalization of lipid profile and moderate modulation of immunity system markers. *Pharma Nutrition.* 2014; 2:33-7.
- [37] Dewar ET, Maddy ML. Faecal excretion of degraded and native carrageenan by the young rat. *J Pharm Pharmacol.* 1970;22:791-3.
- [38] Uno Y, Omoto T, Goto Y, Asai I, Nakamura M, Maitani T. Molecular weight and fecal excreted quantity of carrageenan administered to rats in blended feed. *Jpn. J Food Chem.* 2001;8:83-93.
- [39] Singh KK. In Rowe RC, Shesky PJ and Quinn ME, Eds. *Handbook of Pharmaceutical Excipients 6th Edition*, RPS Publishing, High Street, London, UK. 2009; pp. 122-6.
- [40] Walmsley RS, Ayres RC, Pounder RE, Allan RN. A simple clinical colitis activity index. *Gut.* 1998;43:29-32.
- [41] Walsh AJ, Bryant RV, Travis SPL. Current best practice for disease activity assessment in IBD. *Nat Rev Gastroenterol Hepatol.* 2016;13:567-79.
- [42] Guyatt G, Mitchell A, Irvine EJ, Singer J, Williams N, Goodacre R, Tompkins C. A new measure of health status for clinical trials in inflammatory bowel disease. *Gastroenterol.* 1989;96:804-10.
- [43] Irvine EJ, Zhou Q, Thompson AK. The Short Inflammatory Bowel Disease Questionnaire: A quality of life instrument for community physicians managing inflammatory bowel disease. CCRPT Investigators. Canadian Crohn’s Relapse Prevention Trial. *Am J Gastroenterol.* 1996;1:1571-8.
- [44] Pallis AG, Mouzas AI, Vlachonikolis IG. The inflammatory bowel disease questionnaire: A review of its national validation studies. *Inflamm Bowel Dis.* 2004;10(3):261-9.
- [45] Jowett SL, Barton JR, Welfare MR. Use of the simple clinical colitis activity index (SCCAI) to define relapse of ulcerative colitis (UC). *Gut.* 2001;48 (Suppl 1):A1-A124.
- [46] Berty HP, Shi H, Lyons-Weiler J. Determining the statistical significance of survivorship prediction models. *J Eval Clin Pract* 2010;16(1):155-65.