Gastritis Reflux CLIFF Peer Review

1. *Strength of evidence presented in CLIFF in modeling the characteristics and mission impact of this medical condition within the spaceflight environment?*

   Score: 3

   Grading Scale:

   4 - Confident

   3 - Somewhat confident

   2 – Neutral

   1 – Somewhat unconfident

   0 – Not confident

2. How confident are you that these data accurately represent the spaceflight environment?	*Tables 1, 2* *Composite Incidence and Level of Confidence*	*Table 3, 4 BC/WC* *Loss of Crew and probability*	*Table 4 - Clinical Phase 1, 2 durations* *Table 5 – CP 2 Loss of Crew*	*Table 4 – RTDC %* *Table 6 – RTDC Loss of Crew*	*Table 4 – Loss of Crew Life %* *Table 7 – Loss of Crew Life, Loss of Crew*	*Table 8 – TI*
Score	*2*	*4*	*4*	*4*	*4*	*3*

Comments and limitations:

1. **General comments.** The medical condition that is the subject of this CLIFF includes inflammation or discomfort in the esophagus, stomach, or duodenum. Symptoms of gastritis in mild cases often include epigastric burning, abdominal discomfort, nausea, indigestion, “heartburn,” postprandial fullness, but can progress to significant chest or abdominal pain. Given that the symptoms may be ill-defined or confused with other conditions such as acute coronary syndrome, pancreatitis, or cholecystitis, the categorization of disease may be imprecise in terrestrial studies as well as in historical spaceflight data. Therefore, it was difficult to find references that precisely matched the conditions defined in the CLIFF. Also, astronauts may experience symptoms that may be quite mild and may go unreported. Analysis of the usage of medications to treat gastritis such as antacids would be useful to determine the frequency with which astronauts treat themselves for these complaints. This might further refine the mean incidence.

2. **Incidence.** There is considerable heterogeneity in the references that were found using the MESH search terms. Thus, none of the terrestrial or analogue references adequately studied the condition as defined. There is fairly robust spaceflight data for this condition and although the

sample size is far smaller than terrestrial studies, this data is felt to be the best evidence available to represent the actual incidence value. Whenever actual spaceflight data is available, it should be strongly considered in the calculation of mean incidence. Unfortunately, the statistical methods being used strongly weight the mean incidence towards studies with large sample sizes. Since spaceflight datasets, by their very nature, have small sample sizes, they may have a negligible influence on the calculation of mean incidence when included with larger sample size terrestrial studies. Thus, given its small sample size, the only way for the spaceflight data to be used is to consider it as the only relevant reference. It should be noted that the biostatistician did not agree with this decision to use the spaceflight data, and it was noted:

"…initial Bayesian meta-analysis showed mean rate of 0.037 per py with SD 0.00250 leading to a recommended distribution of Gamma(shape=216.67, rate=5892.6). After seeing the distribution, expertise rankings changed with the intention to change the distribution. The decision was made to use only the in-flight data. This resulted in a mean rate of 0.3514 per py with SD 0.0883 which is approximated by a Gamma (shape=15.84, rate=45.07). This changing of expert opinion rankings after seeing the resulting distribution should be approached carefully; the biostatistician does not concur with this decision."

In response to these comments, the distribution was not changed in response to seeing the initial distribution. Rather, it was a decision made after careful deliberation between reviewers and involved making a choice between ignoring fairly robust spaceflight data or giving preference to a terrestrial study that did not completely describe the breadth of the medical condition as defined in the CLIFF. The results of the initial analysis were also questioned because they did not reflect the very high frequency with which this condition (which includes indigestion, heartburn, reflux, stomach upset) occurs in space. The initial distribution of 0.037 per py indicates that, on average, we would not expect a single case of such mild symptoms for nearly 30 person-years. This would imply that a four-person crew could be expected to fly a three year mission to Mars TWICE without having a single case of indigestion. Having flown in space four times, I experienced what could be described as heartburn on numerous occasions over a total of about a month in space. Therefore, it was felt that the analysis, that included both the terrestrial study and the spaceflight data and was overwhelmingly influenced by the terrestrial study, was not representative of the true incidence. In reality, if I had a preference for the study that most closely reflects the actual incidence, I would have chosen the Storms, et al paper, which had a mean incidence of 0.75 per py. However, I do not think that we should ignore spaceflight data when it is available, and the only way to do so is to choose the Antonsen paper as the top-rated reference.

There is probably significant under-reporting of upper gastrointestinal tract complaints by the crew as they are often self-limited, and the crew most likely has self-medicated in the past without reporting the condition to the ground. This could potentially impact the calculation of the amount of medication that needs to be carried on board. The best case/worst case probabilities, clinical phase durations, RTDC and Loss of crew life probabilities are supported by terrestrial evidence.

3. **Clinical phase data**. I am in agreement with the clinical phase durations entered in this CLIFF. This data was compiled by a team of expert clinicians, and I was involved in the discussions leading to the calculation of the clinical phase durations.

4. **Task impairment**. The analysis was performed and reviewed by a separate team of expert clinicians. I am aware of the methods that were used but did not directly participate in the discussions.

5. **Dependence on mission phase or duration of mission**. This medical condition should not be dependent on mission phase although except that it may occur more commonly early in the mission.

6. **Limitations**

- 'Gastritis' is a broad term not consistently used in medical practice and can be taken to describe a variety of conditions. Therefore, it was difficult to find terrestrial evidence that encompassed all the possible nuances of "gastritis."
- Although H. pylori resources were not included for this condition's CRT; however, studies used in this CLIFF to determine certain parameters may have incorporated this condition. It is not believed that this significantly impacts the veracity of the parameters although H. pylori in the worst case scenario is still a possibility.
- Under-reporting for this condition could have been significant during past space missions and may result in underestimation of the incidence. There is the possibility that sufficient medication to treat gastroenteritis will not be carried on board if the condition incidence used is too low.
- The statistical methods being used strongly weight the mean incidence towards studies with large sample sizes. Since spaceflight datasets, by their very nature, have small sample sizes, they may have a negligible influence on the calculation of mean incidence when included with larger sample size terrestrial studies.

Reviewed by:

David C. Hilmers