

## Peer Review File

Article information: <http://dx.doi.org/10.21037/amj-21-15>

### Reviewer Comments

The research carried out in southeastern North Carolina is very interesting and highlights the importance of fighting against Hepatitis C virus and the improved HCV screening.

#### SOME OBSERVATIONS

##### Major Issues

1. In methods, the authors need to explain better the model created using a forward-selection approach to provide the minimum Akaike information criterion.
2. The retrospective analysis probably resulted in the absence of essential data. The authors need to detail how many data were affected and the impact of this in the results.

##### Minor Issues

1. Discuss whether new studies using the model created using a forward-selection approach to provide the minimum Akaike information criterion data are worthwhile and improve data loss.

**Comment 1:** “In methods, the authors need to explain better the model created using a forward-selection approach to provide the minimum Akaike information criterion.”

**Reply 1:** We have added additional information regarding our approach as below. Feel free to let us know if you have any additional questions regarding our methods, and we would be glad to elaborate.

**Changes in text:** Pages 4-5, Lines 81-93 were added, “We started our forward-selection approach by categorizing patients within our pre-determined age ranges (18-24, 25-44, 45-64, and 65+) and used this as a dummy variable, serving as a way to include age as both a categorical and explanatory variable within our regression. We then sequentially tested the addition of each of the demographic or socioeconomic characteristics listed in Table 1. If the characteristic was statistically significant and strengthened the model, then it was kept, added to the regression, and the next characteristic was tested. If a characteristic was not statistically significant then it was not added to the model. The stepwise building of this model can be seen in both in Table 3 and Table 4 with each column listed 1-8 representing a step in the building of the model by characteristic, finally represented by the equations seen in the results section below. Each model in this process was compared to others developed in this stepwise development to ensure the final equation had the minimum Akaike information criterion, including as many relevant criteria as possible, creating the truest achievable model for our data.”

**Comment 2:** “The retrospective analysis probably resulted in the absence of essential data. The authors need to detail how many data were affected and the impact of this in the results.”

**Reply 2:** Acknowledgement of potential absence of essential data and biases has been added as below.

**Changes in text:** Pages 5-6, Lines 94-115 were added: “There is inherent risk of absence of

some data in all retrospective analyses, however via utilizing a matched case-control style approach including date-based sampling, the two groups (screen versus unscreened) were paired as best as possible to minimize selection bias and the variation of extrinsic factors differently affecting them. Information regarding the completion of a negative screening test was not taken from verbally reported patient history, but only obtained from negative results from laboratory HCV antibody testing performed within the healthcare system where this study was performed. Reported history of HCV infection was accepted as a positive diagnosis in addition to positive laboratory testing, so there is potential for some information bias to a portion of 295 patients per table 2, though the percentage of individuals with recall bias falsely reporting a history of HCV infection is thought to likely be low. Other information bias includes potential history not obtained from patients or not included in the electronic medical record, however this was minimized as best as possible by using the same strict criteria for gathering of all measured demographic and social factors in the study and requiring a “yes” or “no” response to be recorded in a field, with no assumptions made for null fields. Such fields may include including primary language, access to a primary care provider, history of IV drug use, insurance status, and race. A healthcare system-based sampling method was deemed appropriate, as opposed to a general population sampling, as the variable being primarily examined was whether or not a patient was screened for HCV when eligible when seen at a medical visit, though likelihood to test positive when screened was examined as a secondary point. Also, healthcare system-based sampling was important for this study as to provide a baseline to potentially later build upon for quality improvement measures, which may also later positively contribute to the field of knowledge of hepatitis screening disparities.”

**Comment 3:** “Discuss whether new studies using the model created using a forward-selection approach to provide the minimum Akaike information criterion data are worthwhile and improve data loss.”

**Reply 3:** This information has been added as below.

**Changes in text:** Page 13, Lines 284-290 have been added: “We believe using this forward-selection approach to provide the minimum Akaike information criterion is a worthwhile approach for future studies of similar focus as it would allow direct comparison of numerous models to determine which individual characteristics significantly increase the strength of the model while minimizing data loss, and establishing the best overall fit. This way, only key characteristics are included in final predictive models, but also the singular contribution of each characteristic to a model can be evaluated in a step-wise fashion so that no data contribution from any particular subset is overlooked or unaccounted for.”