Could Systematic Error Explain Observed Effects in Epidemiologic Research on H. pylori Transmission Pathways?: A Bias Analysis

Hastings, Emily V., (evhastin@ualberta.ca), Y. Yasui, K.J. Goodman, and the CANHelp Working Group.

Introduction

H. pylori bacteria colonize the lining of the stomach and/or duodenum, where they cause chronic inflammation. While evidence from across Canada indicates that the nation-wide prevalence is declining, Canadian Arctic Aboriginal populations continue to experience a disproportionately high frequency of this infection. The CANHelp Working Group conducts research to address concerns about health risks from H. pylori infection in Aboriginal communities where prevalence ranges from 55-70%.

Public health measures to control the infection have not been developed due to insufficient information on transmission pathways. Developing such measures requires accurate effect estimation in studies of transmission risk factors.

Standard epidemiologic analysis fails to assess uncertainty introduced by systematic error (bias), reducing the usefulness of such estimates for effective decision-making. We present simple quantitative methods for assessing the hypothesis that a specific bias explains an observed effect.

Methods

We used cross-sectional data from community projects to estimate odds ratios (OR) for associations between environmental exposures and H. pylori prevalence, using multivariable logistic regression models that account for random effects of household.

To evaluate the potential impact of selection bias resulting from factors that influence participation in a community research project, the distributions of relevant factors were compared between project participants and available census data for each community.

Sensitivty analysis was used to assess the hypothesis that a specific bias explains effects estimated for selected exposures. Bias hypotheses based on exposure misclassification were assessed for exposures with the greatest degree of uncertainty: untreated water consumption (ever/never); exposure to sewage in the home (ever/never); and age of household head (younger than 21, 21-49, 50+). Sensitivity analysis was used to assess the hypothesis that a specific bias explains effects estimated for selected exposures. Bias hypotheses based on exposure misclassification were assessed for exposures with the greatest degree of uncertainty: untreated water consumption (ever/never); exposure to sewage in the home (ever/never); and age of household head (younger than 21, 21-49, 50+).

The following equation was used to estimate bias-adjusted ORs:

\[
(\frac{A_u - Fn_u T_u}{A_u - Fn_u T_u}) / (\frac{B_u - Fn_n T_u}{B_u - Fn_n T_u})
\]

Where:

- \(A_u = \) exposed/exposed cases
- \(B_u = \) exposed/unexposed non-cases
- \(Fp_{1,0} = \) false-positive probability in cases/non-cases
- \(Fn_{1,0} = \) false-negative probability in cases/non-cases
- \(T_{1,0} = \) total cases/non-cases

Results

Exposure Misclassification:

A scenario that seems plausible is poor sensitivity of exposure classification due to recall error characterized by a higher probability of failing to report exposure that occurred (false negative exposure classification) than of falsely reporting exposure that did not occur (false positive exposure classification).

We hypothesized that differential exposure misclassification resulted from H.pylori-negative participants being more likely to report exposure that occurred than H.pylori-positive participants, given their higher average education levels. Such differential misclassification would result in a higher probability of false negative exposure classification among cases than non-cases. We calculated bias-adjusted ORs for a series of hypothesized scenarios with varying degrees of contrast between cases and noncases in false negative exposure classification probabilities.

The estimated OR (without bias adjustment) for the effect of consuming untreated water (ever/never) on prevalent H. pylori infection was: 0.36 (95%CI: 0.14,0.94).

The estimated OR (without bias adjustment) for the effect of exposure to sewage in the home (ever/never) on prevalent H. pylori infection was: 0.49 (95%CI: 0.2,1.1).

Discussion

We have shown how a simple quantitative method can be used to assess a hypothesis regarding the impact of a suspected systematic error on study results. Such methods can be used to assess bias hypotheses pertaining to selection errors and confounding in addition to misclassification errors.

This type of analysis provides a clear demonstration of how vulnerable study results are to hypothesized sources of systematic error; thus it is useful for informing appropriate interpretations of results.


<table>
<thead>
<tr>
<th>Exposure Misclassification Scenarios</th>
<th>Scenario 1</th>
<th>Scenario 2</th>
<th>Scenario 3</th>
<th>Scenario 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>False-negative probability in cases</td>
<td>40%</td>
<td>50%</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>False-positive probability in cases</td>
<td>20%</td>
<td>20%</td>
<td>20%</td>
<td>20%</td>
</tr>
<tr>
<td>Bias-adjusted OR</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
<td>0.37</td>
</tr>
</tbody>
</table>

Table 3: Exposure misclassification scenario and corresponding bias-adjusted ORs for the effect of untreated water consumption (ever/never) on H. pylori prevalence

Conclusion

The distributions of ethnicity and sex were similar between participants in all three projects and the respective census populations. Young individuals (0-19 years) were underrepresented in the study populations. Study participants from both Aklavik and Old Crow had a higher median income than reported by the census population for each community. These observations could be used to develop selection bias hypotheses for quantitative bias analysis.

Results from bias analyses aimed at assessing the hypothesis that estimated effects were due to differential exposure misclassification revealed that cases and non-cases would need to differ drastically with respect to recall accuracy in order to explain the large inverse association estimated for untreated water consumption and somewhat less so for the more moderate inverse association estimated for exposure to sewage.

In the absence of a plausible reason to expect that cases and non-cases would differ by the amount necessary to explain all of the estimated effects, it is unlikely that this source of bias is solely responsible for the observed effect sizes.