Problems with epidemiologic studies on *Helicobacter pylori* in India

K M MOHANDAS

Division of Digestive Diseases and Clinical Nutrition, Tata Memorial Hospital, Mumbai 400 012

The discovery of *Helicobacter pylori* (*H. pylori*) in the early 1980s resulted in a deluge of articles in the medical literature. Many studies focused on the association between the infection and various gastrointestinal and extra-gastrointestinal diseases. Indian workers have been no exception to this phenomenon.1 Unfortunately, numerous among these studies, including many from India, have major drawbacks which have resulted in misleading conclusions.

In general, most Indian studies are designed as either cross-sectional or case-control and therefore many limitations are to be anticipated (Table 1). The common mistakes are either in the study design (selection bias, lack of proper controls, small sample size and the use of variable methods to detect *H. pylori* infection) or in the interpretation (not appreciating the limitations in the study design or overlooking confounding factors). We therefore have to be cautious in interpreting these data from an epidemiological viewpoint. One of the reasons for these limitations is that knowledge about *H. pylori* increased rapidly within a short span of 8-10 years. Thus, many observations of the previous year became irrelevant the year after.

This article will focus on the limitations of epidemiologic studies with reference to general epidemiologic principles.

Problems with epidemiologic studies

Selection bias of patients and controls

*H. pylori* infection is primarily a disease of the socio-economically deprived populations. High rates of infection will be found in patients and their relatives (often used as controls) attending general hospitals in India.2,3 The prevalence of *H. pylori* infection in these populations may not be representative for Indians and will also vary in different parts of India.

A selection bias is also introduced by the younger age of individuals enrolled in various studies. This is inevitable because of the skewed age distribution of the Indian population, with over 80% of the population less than 40 years (Table 2).4 Therefore diseases in India appear to occur a decade earlier than in the West. For the same reason, the prevalence of diseases like gastric intestinal metaplasia or gastric ulcer which are age-related will be underestimated. It is possible to correct some of these flaws using statistical models. However, this will entail a large sample size and proper data collection.

Recall bias and misuse of potent drugs

Because of the cross-sectional and case-control study designs, other important causal factors that affect the development of the disease or the presence of *H. pylori* infection may be left out. For example, there is widespread over-the-counter abuse of omeprazole and H₂ blockers in India for dyspeptic symptoms; these drugs are known to reduce the detection rates of *H. pylori* by the rapid urease test (RUT) and other tests.5-8 History of recent use of these medications is often left out, and this may contribute to discrepancies in the prevalence data.

Let me cite the example of duodenal ulcer. The prevalence of *H. pylori* infection in patients with duodenal ulcer in Western countries is 95% (95% CI 94-96).9 while in Indian patients it ranges widely from 70% to 90%.10 Does this mean that the etiology of duodenal ulcers in India is different from that in the West? Rather, this is more likely to be due to the reduced sensitivity of RUT and histology, at least partly as a result of abuse of omeprazole and other drugs. As evidence is a study from New Delhi which reported the prevalence of *H. pylori* infection in patients with duodenal ulcer by RUT, brush cytology and serology to be 71%, 82% and 100%, respectively.11

Bias due to inappropriate measurements

Histology is a popular method of detecting *H. pylori* infection in India. The story of *H. pylori* is one of the best examples of the dictum “what the mind does not know the eyes do not see”. In spite of the widespread distribution of *H. pylori* infection, it was not discovered till two decades ago. This is also illustrated by the wide interobserver variations in sensitivity and specificity among pa-

---

Table 1: Summary of limitations in epidemiologic studies

<table>
<thead>
<tr>
<th>Study design</th>
<th>Selection bias</th>
<th>Recall bias</th>
<th>Measurement bias</th>
<th>Confounding factor</th>
<th>Lost to follow up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ecological</td>
<td>NA</td>
<td>NA</td>
<td>High</td>
<td>High</td>
<td>NA</td>
</tr>
<tr>
<td>Cross-sectional</td>
<td>High</td>
<td>High</td>
<td>Medium</td>
<td>High</td>
<td>NA</td>
</tr>
<tr>
<td>Case-control</td>
<td>Medium</td>
<td>Low</td>
<td>Low</td>
<td>Medium</td>
<td>Low</td>
</tr>
<tr>
<td>Cohort</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>NA</td>
</tr>
<tr>
<td>Randomized</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>Very low</td>
<td>Medium</td>
</tr>
</tbody>
</table>

Table 2: Age distribution in population-based cancer registry of Chennai and Mumbai (1993-87)

<table>
<thead>
<tr>
<th>Age group</th>
<th>Chennai (%)</th>
<th>Mumbai (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-19</td>
<td>42.4</td>
<td>38.8</td>
</tr>
<tr>
<td>20-39</td>
<td>34.6</td>
<td>39.3</td>
</tr>
<tr>
<td>40-59</td>
<td>17.4</td>
<td>17.8</td>
</tr>
<tr>
<td>60+</td>
<td>5.4</td>
<td>4.0</td>
</tr>
</tbody>
</table>

Source: Ref 6
Mohanadas

Table 3: Sample size required for unmatched case-control study

<table>
<thead>
<tr>
<th>Anticipated odds ratio</th>
<th>Hypothetical prevalence of H. pylori infection in controls 50%</th>
<th>60%</th>
<th>70%</th>
<th>80%</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5</td>
<td>178</td>
<td>204</td>
<td>254</td>
<td>362</td>
</tr>
<tr>
<td>3</td>
<td>130</td>
<td>152</td>
<td>192</td>
<td>276</td>
</tr>
<tr>
<td>3.5</td>
<td>106</td>
<td>124</td>
<td>158</td>
<td>230</td>
</tr>
</tbody>
</table>

Estimates are for 95% confidence, 80% power and 1:1 case-control sample

Source: Ref 13

The kits used for RUT in Indian studies are developed and standardized in-house and are read at different time intervals ranging from one hour to 24 hours. This alters the sensitivity and specificity of RUT and makes comparisons between institutions difficult.

While serology is one of the best methods for epidemiologic studies, few workers in India have used it. Because of differences in strains of H. pylori in different geographic regions antigen selection is very important. Imported serology kits were never standardized or tested in duplicate before their clinical use in India. Performance of serologic methods varies with the antigens chosen, the population from which the reference sera is drawn, age, ethnicity and homologous and heterologous infection rates in the population being studied.

Lack of statistical power

Most epidemiologic studies from India lack in statistical power (Table 3). Assuming that 70% of adult Indians who may serve as controls in epidemiologic studies are infected with H. pylori, studies with small samples may not be sufficient to detect significant disease associations. It is important to have a sample size of 40 to 50 subjects in each age group and gender group to eliminate sampling bias.

Future directions

In spite of well-meaning individual efforts, little new knowledge about H. pylori has been added from India. There are several important epidemiologic issues that need to be resolved on a priority basis.

First, we need to determine the true prevalence of H. pylori in the normal Indian population from different regions. Although it is widely assumed — based on a few Indian and some immigrant studies — that the prevalence rates in India reach 70% to 80%, there is not a single study that gives complete information about the normal population. In fact, the prevalence of H. pylori infection in the control groups of many Indian studies has ranged from 10% to 80%. A recent seroepidemiologic study from Chennai, a high-incidence area for peptic ulcer and gastric cancer, reported the prevalence of H. pylori infection in the control subjects to be 37% to 47% using different methods. It is obvious that a multicenter seroepidemiologic study is urgently required. This should be done using methods that are validated, with adequate sample size for each age group and sex group. While a house-to-house survey will provide a truly normal population sample, sampling of blood donors or asymptomatic industrial employees is easier to execute. Some useful information can also be obtained by using the serum that was collected from the hepatitis epidemic. For prevalence in children, serologic methods using salivary antibodies are probably the most acceptable. Once we have the true population prevalence of H. pylori infection, case-control studies can be designed to quantitate the extent of association.

Second, a unique epidemiologic feature of peptic ulcer in India has been a 1:30 ratio of gastric to duodenal ulcer, with excess male predominance and younger age of onset. My personal observation and the results of some other reports suggest that this may not be true. It is likely that the 1:30 ratio was recorded in the pre-endoscopy era when the detection of small gastric ulcers on barium studies was comparatively more difficult. The age and gender discrepancies are due to bias in the selection of cases. We need large multicenter studies to resolve this controversy.

It has also been hypothesized that there is a high peptic ulcer and gastric cancer incidence area comprising south India and the east coast, while the northern regions of Punjab, Haryana, Uttar Pradesh and Himachal Pradesh are relatively low-incidence areas. We need to study the population prevalence of ulcer disease and H. pylori infection in these regions to ascertain its role in the causality.

Thirdly, one group of investigators from Mumbai have documented the universal presence of H. pylori in dental plaques and have hypothesized that this may be a permanent reservoir for reinfection. It is very important to study the recurrence of H. pylori infection long after eradication to ascertain if this hypothesis is true.

Finally, several recent studies have reported the development of intestinal metaplasia during omeprazole therapy in the presence of H. pylori infection. Drug manufacturers abroad now include a warning in their advertisements. This phenomenon needs to be tested in India because there is widespread abuse of the drug in the presence of high prevalence of H. pylori infection. Whether young patients with reflux esophagitis, for example, can be put on long-term omeprazole therapy in the presence of H. pylori infection needs to be clarified.

Conclusions

The massive literature on H. pylori in the last 15 years has resulted in substantial information on this organism and its association with human disease. Unfortunately, a large amount of published literature with major drawbacks has resulted in a lot of confusion. To overcome this problem,
societies in Europe and Japan have appointed groups of experts to propose guidelines on how research on Helicobacter pylori should be conducted and how the results should be published even in Abstract form. I hope that the Position Paper in this issue of the Journal will serve as an eye-opener for Indian researchers and will improve the standard of H. pylori studies in India.

References