To: Rob Watters, Chairman of the Technical Subcommittee (TSC)

From: The Expert Advisory Committee (EAC)

Re: CHAP A report, prepared by the Ventana Clinical Research Corporation

Date: 24 January 2005

The CHAP A report dated 24 September 2004 encompasses a large effort on the parts of both Ventana and the Port Colborne respondents, and all should be congratulated on the scale and scope of the work undertaken. Still, there are a number of aspects of the report arising from the survey which need to be considered in the process of interpretation, and other aspects which may pose barriers to developing strong conclusions. These aspects are as follows:

Response rates

The staff conducting the survey clearly expended substantial efforts to obtain an acceptably high response rate from the community. Yet while the expectation was to achieve 75-80% response across all five of the GSA's, the actual response rate was disappointingly low (adults: 41.6%, p. 47; children: 35.1%, p. 47). In addition, there is some variability in response across the GSA's, ranging from 28 to 43% for children and adolescents, and from 31.4% to 50.1% for adults (p. 50). While a low response rate does not by definition pose serious interpretative problems, the opportunity for obtaining a non-representative sample increases as the response rate declines. In addition, in the circumstances represented in Port Colborne, there are few data which describe the characteristics of the non-respondents; these are limited to age, sex, length of residence, and household size (p. 55). The lack of information on key health status indicators makes it very difficult to place in context the characteristics of the responding sample vis a vis the health of the city as a whole. In these circumstances, the opportunity is great for bias to explain any deviations.

Survey questionnaire development

While it was the intention of the investigators to compare data derived from the Port Colborne survey with external data sources, there are a number of problems which were introduced during questionnaire development which make those comparisons flawed.

The SF-36, for example, is a standard and validated questionnaire for obtaining general and comparative data on health status. It was intended to be used as part of the Self-Reported Health Questionnaire (SRHQ), but when it was inserted, only 35 of the 36 questions appeared.

As another general example of deviation in question wording, the SRHQ generally asked "Have you ever been told by a physician that you have…." The main comparison data for the Port Colborne adults derive from the Canadian Community Health Survey (CCHS). The questionnaire used in that survey generally asked, "Do you have…?" The clarifying statement "Remember, we're interested in conditions diagnosed by a health professional…" was reiterated periodically throughout the CCHS questionnaire, as a reminder to respondents. Two differences are significant: The first, and as indicated by the investigators, is the fact that SRHQ identified lifetime prevalence ("Have you *ever* been told"), while the CCHS included period prevalence ("Do you *have*"). For many conditions, which last a lifetime, this distinction would be minimal; it is difficult, however, to predict what someone would respond to "do you have" if he had been treated successfully (i.e., "cured") in the distant past for a condition. In this instance, the wording distinction might be highly significant indeed, and the comparison between the two data sets invalid.

The second difference in the wording of the questions between SRHQ and CCHS is the focus on diagnosis by physicians or by health professionals. Many people seek health advice from, for instance, pharmacists, naturopaths, and physiotherapists. It is not clear how to compare data from questions that define the diagnosing individual so differently, although one would expect that, in general, prevalence estimates from the more inclusive question would be greater than those from the narrower question.

It's important to note that the investigators commented on the issue of external comparisons (see, e.g., p. 34: "The comparison of lifetime and current prevalence estimates has limitations that must be noted when interpreting the results…."); yet they ignore this qualifier when presenting their key findings (p. 18). Instead, they (and any readers of the report) should substantially de-emphasize what are potentially invalid comparisons.

Definitions

It is not always clear what questions from the SRHQ were used in various analyses, nor what questions from the external data sets were used for prevalence comparisons. A few examples follow:

The difficulty with the implementation of the SF-36 has been described. The key to comparisons with the SF-36 depends entirely on the numbering of the 36 questions, since all of the measures which are derived from the SF-36 (i.e., the physical and mental domains encompassed in the 36-item scale) rely on different combinations of the 36 questions. We do not know how the investigators accounted for the missing question, nor how they numbered the remaining 35. Thus, it is not clear which of the

scales derived from the 35 items is correct: depending on how the remaining questions were numbered, it is possible that much of the information presented is incorrect. Without clarification on the coding, the reader should take care in using the data on health status for comparisons both within Port Colborne and between Port Colborne and other surveys.

Current asthma is clearly defined in the report as being a combination of a series of questions that explore disease intensity. While one can review the CCHS questionnaire and identify similar questions to those in the SRHQ, no documentation provided by the investigators clarifies that they did, in fact, use those similar-looking questions.

Emphysema is problematic because the SRHQ included only emphysema, while the CCHS question stated "emphysema or chronic obstructive pulmonary disease". Thus, comparison between the two data sets appears to be invalid.

Sinusitis poses another sort of comparative problem, for the investigators' documentation of the sources of their questions indicates that this question arose from the "respiratory literature", yet the data are compared with the CCHS (p. 15), which has no question relating to sinusitis in its questionnaire. (In the main body of the report, the investigators refer to a comparison with a journal article - p. 126.)

Cancer represents a particular illustration of the problem of the difference between "ever been told" and "have", especially as it appears that the SRHQ included non-melanoma skin cancer, a highly treatable and fairly common form of the disease which has few long-term health implications if treated promptly. (One assumes that the CCHS data likewise included all of these cancers.) It is uncertain what a respondent would say to the question of having cancer (now) if he had had a non-melanoma skin cancer removed 25 years ago. A substantially more useful comparison would have been the exclusion of non-melanoma skin cancers, limiting the comparison to the more serious but less common cancers, including breast, lung, and colorectal cancers. These data should be available but are not presented.

Childhood disorders and symptoms as documented in this report are difficult to compare, not due to any fault of the investigators, but because there is a paucity of useful survey data from other jurisdictions. Many of the variables are compared to individual study reports, covering age groups that are not comparable to the children/adolescents comprising the Port Colborne sample (e.g., p. 137). The meaning of these comparisons is not altogether clear.

Analysis

The analysis of the SRHQ data is reasonably straightforward. Where difficulties arise is in comparison within the Port Colborne data, and comparison between these and other data. Unfootnoted in the table on p. 18, but identified in its various source tables throughout the report, is the fact that many of the comparisons are adjusted for different sets of potential confounders. The investigators stated (p. 118) that they explored confounding by entering the additional variables (e.g., sex, age, income) into the regression models and looking for a fixed percentage of change in the risk estimate. It appears that the exploration of confounding was somewhat haphazard and inconsistent. As an example, tobacco smoke would be expected to confound results for many of the disorders (e.g., circulatory diseases), but smoking is not indicated as having been adjusted in some of these analyses (e.g., p. 128).

Household income was categorized in the SRHQ as <$15,000, $15,000-$29,000, $30,000-$49,000, $50,000-$79,000, and >$80,000. In contrast, the CCHS used <$50,000, $50,000-$59,000, $60,000-$79,000, and $80,000+ for this same variable. These differences make it difficult to understand the income comparisons presented between the two sets of data (p. 85). Additionally, since approximately 40% of the Port Colborne sample reported income <$50,000 (the lowest CCHS category) and another 25% preferred not to answer the income question, it might be expected that the analytic adjustments for income - an important confounding variable - would not be complete.

The investigators mention the lack of information on other potentially important confounding variables (p. 41), such as physical activity and diet. These factors are rather potent risk factors for some of the conditions included in the SRHQ analyses (e.g., circulatory, digestive, and malignant conditions). The lack of adjustment for these risk factors again emphasizes the need to interpret any of the presented analyses with caution.

The risk estimates were weighted using1996, rather than 2001, census data (see, e.g., p. 42, 47), with the weighting relying on the assumption that responders and non-responders in Port Colborne had "the same underlying disease prevalence", an assumption that "could not be evaluated". Two effects of this weighting precedure are that: 1) the 1996 vs. 2001 census estimates may yield different conclusions regarding representativeness of the sample (p. 88); and, 2) the estimates may be flawed. Studies of response have shown that responders and non-responders may differ on many characteristics. Notably for this report, one would expect that residents of Port Colborne with health concerns are more likely to have responded to the SRHQ, while residents with no health concerns are less likely to do so (as the investigators duly noted on p. 47-48).

Conclusion

It is unfortunate that the main findings (p. 18) in the Executive Summary have been stripped of the many qualifications the investigators were careful to insert into the body of the report, qualifications which would caution a reader not to make conclusions based on the apparent statistical significance. It is also unfortunate that the two columns of Port Colborne risk estimates do not reflect independent analyses, relying as they do on much of the same data. (The data in the right-most column is nearly a subset of the data presented in the middle column.) In addition, a full dose-response analysis (examining exposure for long-term, short-term, and never residence in GSA 3) could have been usefully presented in this primary summary table.

The reader needs to be wary of making comparisons between the Port Colborne data and any external sources, due to problems described above. Indeed, it is difficult to predict in which directions some of these problems might move the results. For instance, the use of "ever had" vs. "now have" could be hypothesized to increase the Port Colborne prevalences compared to other surveys, but the reliance on diagnosis by a physician vs. any health professional might work in the opposite direction. Absent further explanation from the investigators, the health status data from the SF-36 are suspect, and the childhood comparisons are weak.

Internal Port Colborne comparisons are probably reasonably stable and reliable, with the exception of the data on health status and keeping in mind the caveat relating to the potential for response bias. Focusing, then, on the table (p. 18) in the Executive Summary (and substantially ignoring the column comparing Port Colborne with provincial data), one can examine the two columns of internal Port Colborne comparisons: the middle column showing GSA 3 vs. GSA 1, 2, 5; and the incomplete dose-response data in the last column. Odds ratio estimates relating to cancers, and to respiratory, skin, cardiovascular or other conditions are not statistically significant, with the single exception of high blood pressure among long-term GSA 3 residents. (In addition, there is generally no dose-response relationship apparent for these conditions. See respiratory p. 153, skin p. 158, cardiovascular p. 167, cancers and other conditions p. 173.)

There are a number of significantly increased (and one significantly decreased) odds ratio estimates among the other conditions - endocrine, musculoskeletal, digestive, and neurological - which are noteworthy. Four of these are elevated in both Port Colborne comparisons: multiple hypothyroid symptoms, multiple chronic fatigue symptoms, ulcers, and peripheral neuropathy, although none of these four exhibits a strong dose-response (p. 164, 173). For the two symptom lists, this lack of dose-response, coupled with the fact that these are self-reported

symptoms rather than self-reported physician diagnoses, would not support a firm conclusion of increased risk. Alternate explanations certainly would include reporting bias among the respondents.

Thus, the data suggest some associations which are not consistently seen across all of the multiple comparisons presented in the report, which generally do not exhibit a dose-response, and which may be subject to reporting bias by concerned respondents. These associations need to be considered in a larger context, with results from the other Port Colborne health studies and in light of their biological and toxicological plausibility.

Paul Corey — Jan 24/2005 (Date)

Nancy Kreiger — 24 Jan. 2005 (Date)

Marie Roy — 24 Jan 2005 (Date)

Doug Templeton — Jan 24/05 (Date)