President’s Column

It seems fitting that my last editorial as President comes as we are planning another CAPE congress in Vancouver. You see it was Vancouver where I first decided to run for President and it was also there that the first edition of the CAPE Chronicle was unveiled. As that was before Gus took the reins, I put the entire edition together myself. I had promised we would start up a newsletter again and I did not want to renege on that commitment. Since Gus took over the quality and diversity of the contributions has improved greatly. The Chronicle is an important connecting piece for our organization and I am extremely proud of its success.

In reflecting on other aspects of my tenure as President, I am pleased about the transformation we have made from a very informal organization to one with just enough formality to help us to continue to prosper and grow (without changing the culture too much). The ratification of our constitution/terms of reference in particular I am very proud to say occurred under my watch. The increased participation of students and trainees, the updating of our website too, are all signs of the growth and maturity of CAPE. So, all in all, I feel as though I am leaving my post as President with the organization as strong as it has ever been and well positioned to continue to be so into the future. I truly believe if Alec were here to see how we have grown, he would approve of what CAPE has become. I believe this to be so because while we have grown, we have managed to maintain the collegial and supportive environment that makes CAPE a unique organization. I think Alec would be proud and that makes me feel the same.

No one single person though can take credit for CAPE’s success. In the end, the organization is only as strong as its members. I encourage all of you to continue to support CAPE, not only through attendance at annual conferences. Volunteer your time to help continue to develop our website. Participate in governance. Respond positively to Gus when he asks you to write a piece for the Chronicle. Promote CAPE to students, fellows and junior colleagues who you think might be interested in joining. Take an active role in your organization.

In closing, I would like to thank the Executive and all those members who responded when tapped on the shoulder to make a contribution. It was a pleasure serving you and CAPE these past few years. While this most certainly is not goodbye, it is indeed my last communiqué in this role – thank you very, very much for the opportunity and for the memories. Merci pour votre soutien, votre confiance et d'encouragement. Nous sommes plus forts aujourd'hui et le restera dans l'avenir en raison de votre engagement.

John Cairney, Ph.D.
CAPE President
EDITORIAL
Can We Trade Rights for Security?

This began as an opinion piece about the benefits of increasing the involvement of CAPE members in the provision of health policy advice to decision-makers at all levels. There are good reasons for this, but further thoughts on the matter led first to a broader view of health and well-being, and then swerved to a seemingly unavoidable consideration of the role of national government policies. But first:

A Role for Psychiatric Epidemiology? A recent report indicated that the average academic journal article is read in its entirety by about 10 people. The requisite worries about methodology aside; this led me to wonder about the current activities of CAPE members. I conducted a content review of the most recent first-author papers of 12 “prominent” CAPE researchers. Ten of the 12 involved work that is directly relevant to mental health policy &/or program development, such as the adequacy & improvement of services, MI and mortality, child development, recruitment of therapists, and the nature of MI. Three dealt with research methods that support the kind of work exemplified by the first ten (one paper was in both camps). We often hear about ivory-tower research in a pejorative sense, but I found nothing frivolous in this group of 12.

It is worth noting that many of the findings, and nearly all of the skills displayed, have useful, but relatively untapped, implications for many of the current social issues that are reported daily. i.e. unemployment, crime, suicide, race relations, and the recently rejuvenated anti-trophy hunting sentiment and public shaming, and perhaps the most salient of all, terrorism. Most such matters have mental health components to them whether it be as cause or effect, major or minor. Some significant number of “terrorist” actions undoubtedly reflect mental health issues, not political events. In any case, mental health researchers tend not to get involved. On the rare occasion when one does, it occurs after public opinion has already been shaped by personal reactions to the events, media interpretations, and interpretations by those who feel strongly that something should be done. Well, it’s a sort of irony that those who spend the most time studying the nature and impact of mental illnesses, tend to comment later (or never), display the greatest reserve, criticize their own findings, and insist on being decidedly unbombastic. This may be preferable in the general case, but it certainly does not bode well for getting a public policy message out there in a timely and clear way.

A Particular Example. In the context of real terrorism concerns, one wonders whether an apparent Canada-wide emphasis on point-of-problem solutions might cause more evils than it resolves. This approach inevitably leads to greater surveillance and more punitive attitudes toward Canadian citizens, it has a questionable upside, and I worry that our Canadian way of life could be detrimentally transformed without appeal to whatever facts might be available. That is, there is no valid upside – it may seem like something has been done, but it has not been demonstrated that we can stop terrorism, and the fact that mortal attacks are ongoing Worldwide would suggest that the opposite is true. Who doesn’t believe that almost anyone can come up with several ways to wreak havoc?

But there is a definite downside. Ignoring such evidence to the contrary leads to the wrong intervention and cannot be expected to be fruitful. Second, many have argued that a hawkish approach could increase terrorism, not reduce it. Third, citizen scrutiny, reduction of democratic safeguards, and a punitive approach can be expected to increase national mental health sequelae like suicide, anxiety, and antisocial behaviour.

It is actually about our Federal Government. To make a long story short, the usual means of addressing national issues have been overshadowed by the actions of the “Harper Government” (we should have been concerned when it was first reported that the PMO had stated a strong preference for this term, rather than calling it the Canadian Government).
Notably, the Federal Government has impeded the activity of health researchers by (1) placing restrictions on government scientists that have directly prevented accurate reporting of scientific findings, (2) fostering large business influence on health research by consolidating most public funding for health under a new body, the Canadian Institutes for Health Research and placing the head of a major pharmaceutical company on its governing board, (3) shortening the National Census and ceasing to publish the Canada Year Book in 2012 (after 145 years in print!). It was a yearly compendium of information on the nation’s social and economic progress, and (4) with exquisite and subtle collaboration, continued the trend of removing the security once treasured by academic scientists who could speak against government policy with relative impunity. This has come about because of a move toward a business model applied to Canadian universities and the weakening &/or removal of tenure (an easy target).

Furthermore, the Harper Government has produced anti-democratic legislation that will create a whisper thin pallor of social anxiety that will negatively affect all Canadians, and will disproportionately disrupt those with a mental illness, thus producing a circular effect, exacerbating the impact on all of us. These acts include (1) retroactive change in legislation to protect earlier RCMP behaviour, (2) legislation that will allow Canadians to be stripped of their citizenship - an atrocity in civilized nations – the USA Supreme Court, for example, concluded that American citizenship could not be revoked by the Government, (3) loss of freedoms via Bill C-51 in spite of the lack of evidence (and low likelihood) that anyone will actually be protected, and (4) The use of omnibus bills to bring forth new legislation – such bills are highly undemocratic because they contain new legislation on many topics, thus allowing a government to bury controversial laws within one dense package, producing a deceptive “take it or lose it all” dilemma for all MPs.

As if this was not enough, more subtle dicta that weaken our protective institutions and gradually place Canadians under increased surveillance include (1) undermining of the CBC by reducing funding and other actions, (2) increased emphasis on the military component of Canadian life (glamourizing the War of 1812, blatant photo-op setups at taxpayers’ expense. Note that this is not to be confused with the way that we feel about our soldiers), and (3) an apparent “out of the blue” decree that changed the name of our “Canadian Museum of Civilization” to the “Canadian Museum of History” in 2012. On the face of it, this may not sound like such a bad idea, but its stated purpose being to shift focus to “Canadian history and people” does keep us looking inward, backward and angry, rather than looking to the future in a constructive way.

**All in All.** Others have noted that researchers tend not to voice their opinions on controversial matters. This may be a natural reticence, but now there are even more reasons to avoid speaking out, including an increased number of sanctions and restrictions that mean many academics will jeopardize their careers if they act within their duties – not to mention within their rights.

Perhaps after the next election we will have a government that will, among many other things, set up a Commission on all aspects of open communication between Canadians and their governments. Hoping, though, isn’t enough. There is a need to attend to many issues regarding the mental health and well-being of Canadians, and the best way to do at this juncture is to stop Harper.

Gus Thompson
Editor
Welcome to British Columbia

The 2015 CAPE Conference
The Harbour Centre, Simon Fraser University

Featured Speaker CAPE 2015
Katherine M. Keyes

Mailman School of Public Health
Columbia University

Dr. Keyes’ is active and productive in a broad number of areas of epidemiology including life-course studies, psychiatric disorders, long-term outcomes of adverse childhood environments, and cross-generational cohort effects on substance use, mental health, and chronic disease. Katherine has completed key work on methodological issues in age-period-cohort effect estimation, and has examined a range of outcomes including obesity, perinatal outcomes, substance use disorders, and psychological distress. In a young career, she has authored more than 140 peer-reviewed papers plus the 2014 textbook, Epidemiology matters: A new introduction to methodological foundations (Oxford University Press), which forms the basis for a Columbia University digital course (http://epidemiologymatters.org/).

September 30
The Harbour Centre
Simon Fraser University
515 West Hastings Street

For details on CAPE 2015, visit the CAPE Website
http://www.psychiatricepидemiology.ca/

Coordinating Committee:
Dr Elliot Goldner, Simon Fraser University
(Chair, egoldner@sfu.ca)
Dr Chris Richardson, Univ. of British Columbia
Dr Tracie Afifi, University of Manitoba
Dr Gus Thompson, Inst. Of Health Economics
Dr Denise Zabkiewicz, Simon Fraser University
Dr Michael Krausz, Inst. of Mental Health UBC
Should We Correct for Multiple Tests?
David L. Streiner, Ph.D., C.Psych

Suppose you have done a survey looking at the prevalence and sequelae of physical impairments among the elderly. There are a number of outcome measures (e.g., activities of daily living, depression, social isolation). You have run a series of multiple regressions looking for predictors for each of these, and a slew of analyses of variance (ANOVA) examining the effects of factors such as gender and site of the impairment. At the end of the day (or year), you end up with a large number of statistical tests, each with its associated \( p \) level. The question that arises is whether we should take those \( p \) levels at face value, or if we have to apply some correction factor to them.

The first issue is, why is this an issue? We all know from Statistics 101 that if the null hypothesis is true and we use the traditional \( \alpha \) level of .05, then there’s a 5% chance that a statistical test will come out significant. If we run five tests, the probability that at least one will be significant is almost 23%; and by the time we reach 10 tests, the probability is 40%. In general, if \( k \) tests have been performed, the probability of at least one being significant is \( [1 - (1 - \alpha)^k] \). (If you’re interested in the logic behind this equation, you couldn’t do better than read Chapter 5 in Norman & Streiner, 2014.) So, given enough tests, you’re almost doomed to find something significant, even if nothing is going on. This is referred to as the problem of “multiplicity”: the increasing probability of finding significance as more and more statistical tests are conducted.

Over the years, a number of procedures have been proposed to correct for multiplicity, of which the most widely known (and unfortunately, the most widely used), is the Bonferroni inequality. Its popularity is likely due to its utter simplicity – if you have run \( k \) statistical tests, use a nominal significance level of \( \alpha/k \) rather than \( \alpha \). So, if you’ve done 10 tests, you would declare significant anything with a \( p \) level of .05/10 = .005 or smaller. Why did I say “unfortunately”? For two reasons. First, Bonferroni is extremely conservative, resulting in too many Type 2 errors (declaring findings to be non-significant that are actually true). Second, it is based on the often unwarranted assumption that all of the tests are uncorrelated. Many alternatives have been developed, including those of Holm, Hochberg, and Šidák; as well as more sophisticated approaches, such as the false discovery rate and resampling techniques. (For a more complete discussion of these, see Streiner, in press.)

However, using one of these approaches presupposes that we should correct for multiplicity, and this issue is far from settled. There are many people who argue that we must, citing the argument of the inflated \( \alpha \) level if we don’t. There are far fewer people arguing against it, such as Rothman (1990) and Schultz and Grimes (2005), but in my (not so) humble opinion, they have a better case. Indeed, they have forced me to change my view of correction. When I wrote an editorial about it six years ago (Streiner, 2009), I argued strongly in favour of correcting for multiplicity. However, after thinking and writing about it more, I have come closer to the views of

---

David Streiner is Emeritus Professor, Department of Psychiatry & Behavioural Neurosciences, McMaster University and Professor, Department of Psychiatry, University of Toronto.
Rothman and Schultz and Grimes (taking refuge in Emerson’s statement that “A foolish consistency is the hobgoblin of little minds, adored by little statesmen and philosophers and divines ”). Let me lay out their positions and mine.

Rothman (1990) has two arguments. The first is that “Scientists presume that the universe is governed by natural laws, and that underlying the variability that we observe is a network of factors related to one another through causal connections. To entertain the universal null hypothesis is, in effect, to suspend belief in the real world and thereby to question the premise of empiricism” (p. 45). The second argument is that it is better to have Type 1 errors than Type 2; the former can be disproven by later studies, but the latter will almost surely prematurely cut off research into promising and unexpected relationships.

Schultz and Grimes (2005) come at it from a different tack. Let’s go back to our opening example, of the effects of disability. Most likely, our outcome variables (activities of daily living, depression, isolation) will be correlated with each other to some degree. Now imagine that they are all significant at the .04 level. On a logical level, the fact that all three came out in the same direction should reinforce our hypothesis that disability in the aged has adverse effects. However, if we apply a correction for multiplicity, none of the results would be significant. In fact, the more variables that perform as we expect them to, the less likely we are to find a significant result, which is counter-intuitive.

Another problem with correcting for multiple p values is that it’s done inconsistently. Following a significant omnibus F ratio in an ANOVA, we would run one of the many post-hoc tests, such as the Newman-Keuls, Tukey’s Wholly Significant Difference or Honestly Significant Difference (does that imply that the Wholly Significant Difference wasn’t honest?), Fisher’s Least Significant Difference, and so on. All of these are just variations on a t-test, using some form of correction to control the overall α level. But let’s say that, instead of using an ANOVA, we ran a multiple regression, dummy coding the grouping variable. Differences between the groups would be shown by the βs for the dummy variables, but the p levels for these are not corrected for multiplicity. Why in one case but not the other, when Cohen (1968) has shown that the two techniques are mathematically identical? As far as I can tell, there is no reason, other than tradition.

There are other situations where corrections for multiplicity don’t make sense. The primary one is looking at baseline differences in randomized controlled trials (RCTs). There are usually two rationales offered for this: to see if the groups are truly balanced in terms of possible confounders, and to determine if any of the variables should be used as covariates in the analyses. But neither of these arguments makes sense (at least to me). To begin with, the p levels of these tests are meaningless. When we run a test following an intervention, we are testing two hypotheses: either the difference between the groups was caused by intervention, or it was due to chance. At baseline, though, there is no alternative hypothesis; if there are differences (and assuming no one wilfully subverted the randomization process), the probability that they were due to chance is 100%, irrespective of whatever values are printed out. As far as choosing covariates is concerned, we should not be relying on the data. Rather, we should be guided by theory or knowledge, and pick variables that are likely to affect the outcome. Even if they don’t differ significantly between the groups, covariance adjustment usually reduces within-group variance and thus increases the chances of finding something (Norman & Streiner, 2014).

In case-control or cohort studies, the p levels at baseline are meaningful, because there are two competing hypotheses. Here, though, it doesn’t make sense to correct for multiplicity, again for two reasons. We’ve dealt with one of them, looking for potential covariates, in the discussion of RCTs. The second reason echoes Rothman’s (1990) argument; it’s better to be over-inclusive, rather than miss
potentially interesting findings.

Does this mean that we should never correct for multiplicity? No, not at all. There is one area where it’s probably necessary, and that is with unplanned, post-hoc secondary analyses (also called “data dredging” or “fishing expeditions”). Because there may be many of these, it’s best to control for multiple tests. In this case, I would recommend (1) using any correction other than the Bonferroni, and ideally the false discovery rate approach (Benjamini & Hochberg, 1995), and (2) report both the corrected and uncorrected p levels, so readers can form their own opinions.

Even here, though, we have to tread carefully; not all secondary analyses are data dredging expeditions. For example, the At Home/Chez Soi project (Goering et al., 2011) had specific objective and hypotheses. Following its completion, there were a number of hypothesis-driven studies, none of which had (or needed) corrections for multiplicity (e.g., Goering et al., in press).

So, in conclusion, correcting for multiplicity has a place, but it’s probably more limited than has been generally proposed. This position is likely to arouse some objections, and I would welcome any discussion about it.

References


Rothman, K. J. (1990). No adjustments are needed for multiple comparisons. Epidemiology, 1, 43-46.


CAPE/ACÉP was organized in 1984 by a multidisciplinary group to: (1) facilitate communication among those involved in psychiatric epidemiology in Canada, (2) provide information & advice to policy makers, clinicians & scientists in the mental health field, and (3) support quality improvement in psychiatric epidemiology training in Canadian centres.

Visit the CAPE/ACÉP website: [http://www.psychiatricepidemiology.ca/](http://www.psychiatricepidemiology.ca/)

Send submissions & announcements to Gus Thompson, Editor gthompson@ihe.ca

**Important Conference Dates**

CAPE 2015 - Annual Scientific Symposium
September 30th, 2015
Harbour Centre, Simon Fraser University
515 West Hastings Street, Vancouver, BC
Coordinator: Elliot Goldner
egoldner@sfu.ca

IFPE 2015 - 15th International Congress
Bergen, Norway – October 7 - 10, 2015
Contact: Prof. Arnstein Mykletun
arnstein.mykletun@uib.no

International Association for Suicide Prevention (IASP)
28th IASP World Congress
*New Discoveries and Technologies in Suicide Prevention*
June 16 -20, 2015
Montréal, Québec, Canada
Contact : Prof. Brian Mishara

WPA Section, Epidemiology & Public Health
Munich, Germany, Mar 29 to April 1, 2016
“Psychiatric Epidemiology Meets Genetics”
Info: Viviane Kovess: info@wpaepi2016.org