Exclusive Interview With Developers Of STROBE Guidelines For Reporting Of Epidemiologic Studies -

Checklist Called A “Life-Jacket Not a Straight-Jacket” By Developers

Last month, several journals published Guidelines for Reporting Observational Studies in a statement titled Strengthening the Reporting of Observational Studies in Epidemiology (STROBE). The guidelines, which are in fact a checklist, were developed by an ad hoc network of researchers and editors led by Erik von Elm from the University of Bern. The stated purpose for developing the guidelines is to improve the quality of the reporting about observational studies. Publication of the guidelines in Epidemiology was accompanied by a lengthier document which is to serve as an explanation and elaboration of the elements in the checklist. Given the potential influence which such guidelines can have if they become widely adopted as standards in the field, The Epidemiology Monitor sought to better understand the rationale for the checklist and to learn more about the developers’ hopes and plans for the guidelines. Following is the interview we conducted with the lead developer, which included input from the development team.

EpiData: Free Software for Public Health Data Management And Analysis

by the Epidata Association represented by Coordinator Jens M. Lauritsen and Pedro Arias Bohigas

“What proportion of children in our schools are smokers?”, “Which types of fireworks were the cause of injuries during New Year?”, “What types of food were eaten by the 75 guests at the restaurant?” or “How can we visualise the hospital wards compliance with hygiene standards over time?” These and similar public health oriented questions arise in local settings and organisations. The typical mode of working is in accordance with general quality circle principles: Do we have a problem, who is affected, which appropriate counter measures will remedy the situation, and what is the effect and efficiency of this? To work along these lines, knowledge of content matter is required, but also software tools and skills in using them. Staff resources include software skills, time for acquiring skills, and management of tools. Epi Info version 6 software was a main tool in this respect developed and released by the Centers for Disease Control and Prevention (CDC) in USA in the mid 1980’s and from the mid 1990’s in collaboration with the World Health Organization.

- Interview begins on page 2

- EpiData, continues on page 5
“If STROBE can help enable better discussions about research evidence, it might strengthen epidemiology.”

EpiMonitor: Can you say more about the real need or the real reason for this checklist? Check lists do tend to standardize and prescribe actions, and we will not welcome the arrival of STROBE unless we are persuaded that something is broken in epidemiology which such a checklist can fix. Or if you take a more positive view, what will be the real "value added" to epidemiologic work of having such a statement?

Von Elm: The “value added” is the transparency of reporting which can be increased by using the STROBE checklist and from which users of the scientific literature will benefit. Many discussions around epidemiologic studies are about bias or confounding. But to discuss such issues appropriately, it is important that the studies are reported with enough detail in the first place. If STROBE can help enable better discussions about research evidence, it might strengthen epidemiology.

In general, there is nothing more broken in epidemiology than there was with randomised trials when the need for CONSORT was felt more than 10 years ago. The same is true for diagnostic studies when the STARD guidelines were drafted. Of course, there are epidemiologists who will not feel that STROBE is the long awaited guidance and who already write clear and complete study reports. But many other researchers have less experience with observational studies and might find the checklist useful when they write up their study, in particular if they have not trained in epidemiology before.

EpiMonitor: What are the use scenarios which might unfold now that the guidelines are published? Which scenario do you see as most likely and why?

Von Elm: One important “use scenario” is the endorsement of STROBE by journals in their instructions for authors. There is now a wide range of opinions and of degrees of acceptance of STROBE by journal editors. Some specialised epidemiology journals such as EPIDEMIOLOGY are more reserved, whereas general medical journals are enthusiastic. For instance, the Lancet considers adherence to STROBE as a "must" in its instructions, and several journals of the Public Library of Science (PLoS) and the Annals of Internal Medicine recommend STROBE to authors who intend to submit observational studies. At present, more than 20 journals have already endorsed STROBE, among which are several with an epidemiologic or public health signature (see list www.strobe-statement.org). This range of reactions can be understood. Editors of epidemiology journals will themselves be in lesser need of a checklist - moreover, they may be able to judge for themselves when it is right that an author leaves the "well-trodden" path. It could well be that most editors will arrive at a view that is close to that of journals that now recommend use of STROBE without enforcing it.

Second, some of the first reactions after publication of the STROKE Statements were requests from academics asking for permission to use the STROBE articles in course materials. That is a very welcome use of STROBE. I think that authors who start out writing their first few papers are among those who benefit the most from the present checklist. Of course, this checklist of items does not mean necessarily that all important things are said about any given observational study only because...
Someone has written something on each item, from 1 to 22. Given the diversity in observational research, it is easy to think of more points that could be worth being reported depending on the setting, field of research etc.

**EpiMonitor:** Have there been any persons opposed to the creation of the checklist? If so, what arguments have they made against the checklist?

**Von Elm:** Several persons still are opposed: see the commentary by MacMahon and Weiss in EPIDEMIOLOGY. Their main argument is that "they do not judge an apple by how well it is polished". They see a risk that the checklist will be used by authors to embellish their study reports, and by editors to judge the quality of research. On the other side, there is also a "sigh of relief" by Rothman and Poole in the same journal about the checklist being "benign". See also the editorial by the editors of Preventive Medicine who think that STROBE is important in the convergence of clinical medicine and epidemiology.

Interestingly, another type of criticism is that the checklist is too general (see commentary in Int J Epidemiol). It is true to a certain extent that some of the content, in particular of the STROBE explanation and elaboration article, can be found in elementary epi courses or text books. Since our checklist was not primarily aimed at professional epidemiologists, that is not a problem. The group of authors reporting observational studies in whatever medical journal is a much wider group. For them, some help in writing up papers might be particularly fruitful.

**EpiMonitor:** What do you see as the potential for misuse of the checklist by editors and others and what safeguards are in place to prevent the misuse of the checklist?

**Von Elm:** An important concern that we have is that the checklist might be seen as an instrument to measure study quality. That is also a deep concern of the editors of EPIDEMIOLOGY, and has also been voiced by some who commented on STROBE drafts in the past. We have clearly stated that the checklist is not for use as a quality instrument. When authors say in their article that they adhered to STROBE, they do make an attempt at reporting quality. That does not say anything about the quality of their actual study, however. The latter may have had major shortcomings, and if readers are informed about this, that is an example of good reporting. The inverse is also true: when innovative high-profile research is done, authors may be so enthralled by the innovative aspects that they forget to report about the more mundane aspects of their work, e.g. when and where it was done. In that case, the STROBE checklist might help restore the balance, i.e. give sufficient basic information to readers.

Of course, we cannot prevent all possible types of misuse of the checklist, however clearly we have indicated what it is for and what it is not intended for.

**EpiMonitor:** Did all of the signatories of the publication or all the contributors support the statement and agree with its contents?

**Von Elm:** As with all articles, the authors agree with the overall content of the papers. There is no exception here. Of course, it is inevitable that a particular co-author will not have seen his or her own "pet" idea make it into the text. Again, that happens with all...
papers on whatever topic. The contributors listed at the end of the articles are not authors, but they all made great efforts in giving us feedback and improving both papers - and we are very grateful to them. Given the spectrum of different backgrounds and views represented in the group of authors and of contributors, I believe it would be surprising if all of them agreed on every single aspect.

**EpiMonitor:** We did not see mention of any professional societies of epidemiologists contributing to the checklist? Were professional societies consulted and did they respond as a group?

**Von Elm:** We did not consult professional societies, but this was not because we thought that they or their members could not contribute to the STROBE Initiative. We had a wide-ranging consultation of individuals who are well-known in the field, and we had substantial input from several journal editors (from large general medical journals to specialised epidemiologic journals) who contributed during the various stages of making the checklist and writing both articles. We did not aim at formal approval of the checklist by the professional societies before publication. A formal endorsement by learned societies before publication probably would have added to the current concern of some commentators that STROBE could become an official or even mandatory document.

A formal endorsement by learned societies before publication probably would have added to the current concern of some commentators that STROBE could become an official or even mandatory document.

"You can only go beyond speculation and really argue about a study, if you know exactly what happened in that study."

"A formal endorsement by learned societies before publication probably would have added to the current concern of some commentators that STROBE could become an official or even mandatory document."

"You can only go beyond speculation and really argue about a study, if you know exactly what happened in that study."

**EpiMonitor:** The statement says that the checklist is not to be used to determine the quality of the research, yet in other places you say that it will highlight strengths and weaknesses of the studies. By encouraging the reporting of strengths and weaknesses, will not this de facto become a way of commenting on the quality of the research carried out—regardless of the validity of the findings?

**Von Elm:** When authors report in detail what was done and why it was done, others can see a particular methodological aspect or design feature as a strength or as a weakness (and sometimes both opinions exist!). You can only go beyond speculation and really argue about a study, if you know exactly what happened in that study. Of course, authors comment on the quality of their study when they report the details. But still, there is a difference: we do not expect authors to judge themselves on all the weaknesses of their work. This can be done later by others. But they could tell us what problems they encountered along the way and how they tried to solve them. On the other side, if authors write that their study is so good that it surpasses all others, as an editor or reviewer you would start to wonder, as well.

**EpiMonitor:** What was the reason for stating in regards to interpretation that investigators should give a “cautious interpretation” as opposed to a strong or firm interpretation of the results since each may be justified in different circumstances?

**Von Elm:** Right. I think we preferred to err on the side of caution for two reasons: First, scientific progress rarely comes from one single study, but needs
the accumulation and simultaneous interpretation of several studies. Therefore, caution is warranted if it comes to interpreting the results of a single study. Second, one of the reasons that epidemiology sometimes gets criticized is over-enthusiastic press coverage. Most of the time, this includes over-interpretation or generalisation of a study’s findings, even if limitations are mentioned in the original publication. Individual epidemiologists can often not help this, even if they clearly state such limitations in their paper, but they should at least not encourage it.

EpiMonitor: On a related note, why does the checklist not call for discussing the strengths of a study but only the limitations?

Von Elm: The former generally come natural to the author. The latter are crucial for the reader.

EpiMonitor: The examples of study reports in the discussion section and in the abstract section of the statement both give recommendations. However, the topic of making recommendations based on the findings is not addressed in the statement itself. Many in epidemiology would argue that epidemiologists as public health scientists have a responsibility in reporting public health relevant findings to point to desirable actions as your examples do in the statement.

Von Elm: Yes, I also feel that we have this responsibility, in particular if our research is publicly funded. But the topic of recommendations for action is not the same as the topic of quality of reporting. Authors may choose to take it a step further themselves and call for action, or just provide the scientific basis for further recommendations and policy making in an article. In both cases, it is important that such an article is clear enough to be used by people who have or have not content expertise.

EpiMonitor: What does the statement expect researchers to report when they describe the role of funders who have exerted undue influence over the study? How is the fact of undue influence to be described in the role of the funders?

Von Elm: It is, of course, not the authors themselves who will describe the funding influence as "undue". However, a factual statement like "data were collected, held and analysed by the sponsor" may greatly help the reader to assess how independent a study or the publication of its results were from the influence of a sponsor.

Organization (WHO), Geneva Switzerland.

Stability of tools and skills in using them is a major issue, since new or changing tools requires time off from immediate tasks. Over time computers have increased in speed, memory and data capacity, and the operating systems have changed accordingly. With new versions of operating systems it is becoming increasingly difficult if not impossible to use software that worked well on earlier versions. There are different approaches to the development of software and operating systems. One approach is the Open-Source Linux-oriented road of mutual exchange of principles and open standards, and another approach is the Microsoft way, which implies dependency, controlled and undisclosed principles. This is reflected on the political scene in the debates on software patents and the very large financial penalties imposed by the

"One of the reasons that epidemiology sometimes gets criticized is over-enthusiastic press coverage."

"With new versions of operating systems it is becoming increasingly difficult if not impossible to use software that worked well on earlier versions."