60th Anniversary Symposium

Epidemiology, Biostatistics and Occupational Health

2025.11.18

Lyrics to accompany the 'slide'-presentation.

Good Day. I was asked to say some words about the first 60 years. Rather than give my opinions on where we have been, and where we might go, I will mainly just present

data. And I look forward to what the other panelists have to say.

Recently, in an email to a still-active professor who used to be in our department in the 80s and 90s I used the abbreviation EBOH, only to get this reply "What is EBOH?". I will be referring to it simply as THE department. When I joined it, the department was called "Epidemiology and Health." My next department chair said that this was like having a "Department of Nephrology and Kidneys", and

changed it "Epidemiology and Biostatistics" in the mid 80s.
The name was later lengthened to its current 51 characters
(counting spaces!). 154/154

In 1979 I presented at a conference on clinical trials. After the session, a Duncan Thomas told me about a vacancy for a statistician in his department at McGill. He was an Assistant Professor there, and had gotten his PhD a few years earlier. Besides the small bit he told me, I knew nothing else about him or about the department and its work. 1979 was about 14 'BC' (if you use the looser term "Before Computers,' or 'BWWW' if you use the more accurate term Before the World Wide Web). So before I came up for my job interviews, I went to the Countway Library

in Boston, and found this booklet. 112/266

Apart from listing two courses for undergraduate medical students, here is all it contained about the department. During my visit, one of the first professors I interviewed with asked me **who** I would like to **work with** at McGill. Stuck for an answer, and not wanting to show my ignorance, I said "smart people." I joined the department in September 1980. From then on, I would tell prospective McGill students that our department had lots of smart people, so many in fact that one could choose to work with people who were not only **smart**, but also **nice**. 98/364

Whenever we has visitors to the department, we would take them into our small library/reading room that housed the Theses and Dissertations. Students had to deposit their thesis in McLennan Library, but our department insisted they pay for and give an extra copy to the department. For us, it was a great way to show off the department's 'scholarship'; and prospective students could see **what topics** they might want to **work on**, and also, if they looked at the acknowledgements, **who they might like to work with** at McGill. When the department moved from the Lyman Duff

Medical Building to Purvis Hall, these bound volumes with us. By the time we moved to 2001 McGill College, they had outlived their usefulness, other than maybe as a way to elevate laptops. 130/494

Today, the library's eScholarship platform allows world-wide/365/24/7 access to all McGill theses. 12/506

I was able to use `R` to scrape the data from the several pages worth of listings, and make a dataframe of the meta-information. 24/530

These are the earliest titles 5/535

and these are the latest. 5/540

Here is one measure of the growth in the amount of scholarship over the six decades. The overall MSc:PhD ratio is just under 2 to 1, but the fluctuations from decade to decade are a bit more than random. I will leave it to others to speculate on why we haven't had theses in French in recent decades. 58/598

As a supervisor, I fussed a lot about **titles**. I used to ask students to imagine they were the consumer rather than the producer of the information i.e. suppose all they had was the **title** of a Table, or Figure, or an Article, and they had to pay out of their own pocket to see/download the full contents. From the title alone, would they know ahead of time what type of information they were getting? Of course, the **number of words** doesn't translate into **how informative** a title is, nor does it fully capture how 'pithy' or how 'wordy' it is. Nevertheless, since it was easy to count

them, here is the distribution. The mean has gone up by about 5% per decade (The second inset uses the number of characters, familiar to you texters). 135/733

You might be wondering what the outliers were. Could one really have a 1-word title? or a 40 word title? The shortest was for a thesis written by a surgeon; I expect there are several with what I call "X and Y," or "exposure" and "outcome" structure; here is the shortest of those. Here is the genuinely longest title, 34 words; the 42 word title is bilingual (the thesis is in English). 72/805

The library people tell me that 'keywords' is a better umbrella term for what is currently called 'Subject' in the eScholarship platform. The headings represent a diversity of practices and conventions, some historical, some overlapping. Metadata is often a legacy of differing practices. The headings come from the Graduate and Postdoctoral Studies and are shared 'as is' with the Library. After some cleaning in R, I found that the1020 theses yielded 1700 'keywords." Here is the distribution. The 317 mentions of 'Epidemiology and Biostatistics' are all single-phrase entries from 2009 on-

wards, but there are also over 200 others from this era that are a bit more specific. If the department thinks that better subject-categories are worth having, then it might want to contact and work with the Graduate Studies and Library people, especially as the Library people will be doing some cleanup before they soon migrate to a new platform soon.

Given this legacy, I asked chatGPT if it would (retrospectively) categorize the titles from scratch. On my webpage, you can find the exact directions I gave it. Here are the categories it made, and how many it placed in each one. I haven't had time to carefully check its work, but I have put the file it produced on my website for anyone who wants to look at it. I wonder how much inter-categorizer variation there would be if we could also have it done by some human categorizers [it could be a nice teaching exercise in 'measurement']. I expect that chatGPT 'borrowed' these

categories from some relevant journals. I also wonder if you are surprised by the breakdown. 119/1075

Google tells me that the concept of a **word cloud** dates back to 1976, but of course it became a lot easier to make one once we got D.I.Y. graphics. All I did was make a frequency table of the words and invoke this R package. I will be interested to know whether the panelists or the attendees find it informative. [In other contexts I have often said that the more colours and glitz in a graph, the less I believe it.] 82/1157

At the very end, I will play a 'movie' that shows each of the 1020 authors and their thesis titles scroll up the screen, like film credits. I have also put that 3-minute movie on my website, where, if your video-player allows it, you can change the playback speed, and use fast-forward. 52/1209

But before I roll the credits, I want to peer into two of the earliest theses. I chose the first one, from 1969, as I was curious whether pharmaco-epidemiology at McGill had started earlier than I thought. 37/1246

The first student in our department, and indeed the first person in our department, to use the Saskatchewan Drug Database, used it to study "First trimester anticonvulsant therapy and the risk of congenital malformation in the offspring of women with epilepsy." By the way, you might be curious about the earliest known use of the term "Pharmacoepidemiology" **itself**. The usual citation is a 1984 article called "Pharmacoepidemiology: a new discipline," just as our department was changing chairs. It prompted this response from William Inman of the Drug Surveillance Re-

search Unit at the University of Southampton in the UK: "Precisely who first coined the term pharmacopidemiology is for the historians to decide. I have used it for nearly 20 years to describe my own work and that of Hershel Jick and Sydney Shapiro in the United States." 136/1382

Here is an image of the cover page, probably taken from a microfilm copy. 14/1396

And here is a rearrangement of the image that is bit easier to read. 14/1410

His focus was on better ways to detect adverse reactions to drugs. This is the late 1960s, soon after the wake up call from the thalidomide disaster earlier that decade.

So, what became of this promising 'PE' person? Well, not one of the people I contacted was able to help me trace him, so I again had to consult on chatGPT. Then, once I found out about the award named after him. I emailed some people. One answered by saying "I was nominated for the AFMC Charles Boelen International Social Accountability Award last year, but have no idea who he is..." I was a bit skeptical of the information from chatGPT, and it too was not 100% sure. But then Google was able to locate this interview with him, and lead me to his University of

Sherbrooke long-time collaborator, who tells me Dr Boelen now has Masters and Doctoral degrees from **three** Québéc universities. 124/1564

In it, he tells how his career evolved, taking him from 'a little village of Flemish Limburg,' to Montreal; thence to Gondar Ethiopia; and finally to WHO in Geneva. I recommend the full interview to all 3 SPGH departments. I 40/1604

Just this morning, as I was about to record this presentation, I got an email from Charles, with three items from his time in our department, and i have put all three on the webpage, But I can't resist showing the one he calls "really historical: a photo with three graduates and some staff members in 1969." It made me wonder if in 56 years time, those who were graduate students in 2024 will be able to readily find their historical photographs. 82/1686

Now to this thesis from 1974. 6/1692

I discovered it a year ago when preparing a presentation on early 'case-controlling' and conditional logistic regression for the Biostatistics Seminar Series. I had often heard Doug Liddell talk about the study it is based on, and the analysis, but I hadn't realized (or had forgotten) until I re-read Doug's 1988 review paper that it had been written up in an MSc thesis in this department in 1974. It dealt with the sequelae of infections in pregnancy. 77/1769

The abstract has a good summary, 6/1775

But its worth seeing some of the background and details. The data collection was carried out under the direction of Corbett McDonald. He had joined the Public Health Service in the London suburb of Colindale in 1951 and was head of the epidemiological research laboratory from 1960 to 1964, working on the epidemiology of viral infections, particularly influenza.

For 21 months in 1959-60, some 31 antenatal clinics in England and Scotland enrolled almost 12,000 pregnant women at their visit at the end of or just after their first trimester.

At that visit, they gave their reproductive history and a history of any infection they had had in this pregnancy, and they gave a sample of blood. Some of it was used for Rhesus testing and blood typing, and the rest was frozen and stored. 45/1910

The pregnancy outcomes were documented 7-28 days after the end of the pregnancy 13/1923

and the 1-year information recorded on the back of the card. 11/1934

The last of the children would have had their 1st birthday in the Summer of 1962. So, what next? There were over **11,000 pregnancies**, so over 11,000 pairs of cards. The number with adverse outcomes was just over 1,100. That Fall, the bloods of **all of these 1,100 mothers** were thawed out 52/1986
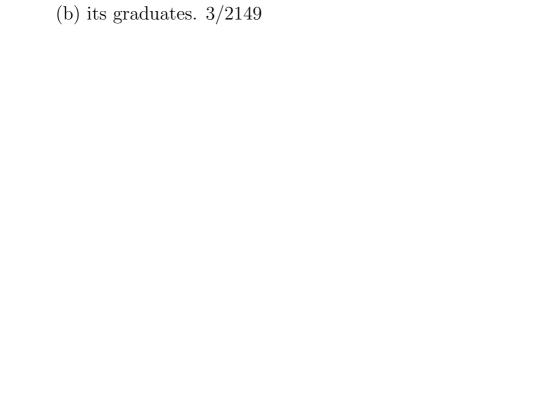
, along with a **1 in 10 random sample** of all the others. Thus, just under **2,200** sera were analyzed for 20 different types of antibodies. In my seminar a year ago I traced this strategy forward until it got the name 'nested case control study. It is now the workhorse for many areas of epidemiology.

I am still trying to find out where John Stewart was from and where he went afterwards. 75/2061

A FEW ENDING REMARKS: (1) the link between one's degree and later career is not always direct (2) it's a good foundation and point of departure (3) McGill should be proud of our department and 35/2096

(a) its 'no-name' designs and being part of what I call "the coming of age" of **the etiologic study**. Norm Breslow, a good friend of McGill, often quoted Ken Rothman, who said that "The sophisticated use and understanding of case-control studies is the most outstanding methodologic development of modern epidemiology." 50/2146

(b) its graduates. 3/2149

Here is the website that has the Presentation and Other Materials 11/2160

Rather than SAY all their names, I have put the thanks you's in the slide. 15/2175

Now let see our EBOH graduates parade again, but this time across the screen, virtually. This movie is also on the website. Depending in what video player your device uses, you may be able to control the playback speed, fast forward etc.

And very lastly, thanks to the symposium organizers for the opportunity to reminisce, and catch up with former colleagues, students and staff. 65/2240