

## Workshop "Thiết kế nghiên cứu và phân tích dữ liệu" tháng 12/2017

### Chương trình dự kiến (cập nhật ngày 08/11/2017)

Chương trình sẽ chia làm hai phần, mỗi phần 6 ngày. Chủ đề phần thứ nhất là thiết kế nghiên cứu và phân tích dữ liệu, với 18 bài giảng. Chủ đề phần thứ hai là về cách viết bài báo khoa học và báo cáo, với 13 bài giảng. Mỗi bài giảng sẽ kèm theo phần thực hành để học viên có kinh nghiệm thực tế.

### Phần I: Thiết kế nghiên cứu và phân tích dữ liệu

#### Ngày 1 (25/12/2017):

- **Bài 1 Qui trình nghiên cứu khoa học:** ý tưởng nghiên cứu đến từ đâu; phân biệt thế nào là nghiên cứu khoa học và những gì không phải là nghiên cứu khoa học; qui trình nghiên cứu; ý nghĩa của nghiên cứu khoa học.
- **Bài 2 Chọn mô hình nghiên cứu:** giải thích các điểm yếu và điểm mạnh của các mô hình nghiên cứu cắt ngang, bệnh chứng, đoàn hệ, nghiên cứu RCT, phân tích tổng hợp.
- **Bài 3 Giới thiệu R:** R sẽ được sử dụng cho phân tích dữ liệu trong lớp học. Bài giảng này sẽ giới thiệu ngôn ngữ R, vận hành, cách đọc dữ liệu, cách biên tập dữ liệu, và một số phân tích đơn giản.

#### Ngày 2 (26/12/2017)

- **Bài 4 Phương pháp thu thập dữ liệu:** giới thiệu cách thu thập dữ liệu và nguyên tắc thiết kế bộ câu hỏi.
- **Bài 5 Quản lý dữ liệu:** quản lý dữ liệu là khâu rất quan trọng cho nghiên cứu khoa học nhưng thường hay bị xem thường và dẫn đến nhiều hệ quả nghiêm trọng. Bài này sẽ giới thiệu cách sắp xếp dữ liệu sao cho thích hợp cho phân tích và những qui tắc đạo đức trong việc quản lý dữ liệu.
- **Bài 6 Phân tích mô tả** (biến liên tục, biến phân nhóm): giới thiệu các phương pháp phân tích mô tả như trung bình, độ lệch chuẩn, khoảng tin cậy 95%. Bài giảng nhấn mạnh đến ý nghĩa và cách diễn giải hơn là kỹ thuật.

#### Ngày 3 (27/12/2017)

- **Bài 7 Phân tích bằng biểu đồ:** trong khoa học biểu đồ đóng vai trò rất quan trọng, nhưng rất tiếc nhiều nghiên cứu có biểu đồ quá kém. Bài này sẽ giới thiệu các nguyên tắc và phương pháp soạn biểu đồ phẩm chất cao dùng các chương trình như *ggplot2* và *sjPlot*. Đây là một trong những khâu rất quan trọng để hiểu dữ liệu và ý nghĩa của dữ liệu.

- **Bài 8 Giới thiệu ý tưởng kiểm định giả thuyết:** bài giảng bàn về hai trường pháp ước tính (estimation) và kiểm định giả thuyết (hypothesis testing), và ý nghĩa của trị số P. Học viên sẽ nắm được ý nghĩa của trị số P và mối liên hệ với kiểm định giả thuyết.
- **Bài 9 Phương pháp so sánh 2 nhóm (biến liên tục):** một trong những mục tiêu nghiên cứu cơ bản là so sánh hai nhóm. Bài giảng sẽ giới thiệu phương pháp t-test cho dữ liệu tuân theo luật phân bố chuẩn, và phương pháp bootstrap cho các dữ liệu không tuân theo luật phân bố chuẩn.

*Ngày 4 (28/12/2017)*

- **Bài 10 Phương pháp so sánh 2 nhóm biến định tính:** Bài giảng sẽ giới thiệu phương pháp z-test, Ki bình phương, và các chỉ số như odds ratio và relative risk để sử dụng cho việc so sánh hai nhóm với biến outcome là định tính hay biến nhị phân.
- **Bài 11 Phân tích tương quan:** giới thiệu khái niệm "tương quan" (correlation) và phương pháp ước tính hệ số tương quan.
- **Bài 12 Mô hình hồi qui tuyến tính:** giới thiệu mô hình hồi qui tuyến tính đơn giản  $y = a + b \cdot x + e$ , giả định và phương pháp kiểm định giả định, ý nghĩa của tham số và cách diễn giải các chỉ số từ mô hình hồi qui tuyến tính.

*Ngày 5 (29/12/2017)*

- **Bài 13 Mô hình hồi qui tuyến tính đa biến:** tiếp tục mô hình hồi qui tuyến tính, bài này sẽ giới thiệu mô hình hồi qui đa biến tiên lượng (multiple linear regression) và những vấn đề liên quan đến đa cộng tuyến, cách chọn biến số có giá trị tiên lượng tốt nhất.
- **Bài 14 Mô hình logistic đơn biến:** bài giảng sẽ giới thiệu khái niệm odds và tỉ số odds (odds ratio) và ứng dụng trong mô hình hồi qui logistic. Bài giảng sẽ tập trung vào cách diễn giải ý nghĩa của các tham số trong mô hình và ý nghĩa thực tế.
- **Bài 15 Mô hình logistic đa biến:** tiếp tục mô hình hồi qui logistic đơn biến, bài này sẽ trình bày phương pháp phân tích với nhiều biến tiên lượng và những vấn đề kèm theo như đa cộng tuyến.

*Ngày 6 (30/12/2017)*

- **Bài 16 Phân tích sống còn (survival analysis):** nhiều nghiên cứu khoa học quan tâm đến yếu tố thời gian dẫn đến biến cố. Bài này giới thiệu phương pháp life-table và Kaplan-Meier trong việc mô tả các biến thời gian và xác suất sống còn. Phương pháp log-rank test cũng sẽ được giới thiệu qua chương trình "survival" trong R.

- **Bài 17 Mô hình Cox:** tiếp tục phương pháp Kaplan-Meier là [mô hình] Cox's proportional hazards model. Bài giảng sẽ giới thiệu khái niệm "hazard" (rủi ro) và cách xây dựng mô hình Cox để tiên lượng một biến phụ thuộc.
- **Bài 18 Sai sót phổ biến trong thiết kế nghiên cứu và phân tích dữ liệu:** trong thực tế có rất nhiều sai sót và thiếu sót trong phân tích dữ liệu, và những sai sót này dẫn đến bài báo bị từ chối công bố trên các tạp san quốc tế.

## **Phần II: Cách viết bài báo khoa học và trình bày**

### **Ngày 1 (2/1/2018): chủ đề kĩ năng trình bày Powerpoint**

- **Bài 1 Kĩ năng soạn Powerpoint:** bài này sẽ trình bày các nguyên tắc trong việc soạn Powerpoint, từ màu sắc, phông chữ, đến qui ước n\*n, sẽ được giải thích và ứng dụng.
- **Bài 2 Kĩ năng trình bày trong hội nghị:** bài giảng sẽ hướng dẫn cách trình bày và "protocol" trong hội nghị khoa học quốc tế. Từ cách nói, trang phục, đến dáng điệu sẽ được hướng dẫn.

### **Ngày 2 (3/1/2018): chủ đề cách viết bài báo khoa học**

- **Bài 1 Cấu trúc bài báo khoa học:** bài giảng sẽ đề cập đến cấu trúc IMRaD trong bài báo khoa học và mục tiêu của từng phần.
- **Bài 2 Tiếng Anh trong bài báo khoa học:** một phần lớn bài giảng sẽ xoay quanh vấn đề tiếng Anh (như văn phong và cách chọn từ ngữ) trong bài báo khoa học. Nguyên tắc IDEA sẽ được giới thiệu và lí giải trong cách viết bài báo khoa học.
- **Bài 3 Chuẩn bị trước khi viết bài báo khoa học:** những bước cần chuẩn bị cho việc soạn một bài báo khoa học.

### **Ngày 3 (4/1/2018)**

- **Bài 4 Cách đặt tựa đề:** tựa đề bài báo khoa học là một yếu tố rất quan trọng, có khi quyết định sự thành bại của bài báo, nhưng rất ít được các tác giả chú ý. Bài giảng này sẽ giới thiệu những nguyên tắc trong việc đặt tựa đề, cùng những điều không nên làm khi đặt tựa đề. Một số nghiên cứu khoa học về tựa đề bài báo cũng sẽ được trình bày để minh họa cho các nguyên tắc chung.
- **Bài 5 Cách viết phần dẫn nhập (introduction):** phần Dẫn nhập của một bài báo khoa học là lí do tồn tại của bài báo, nên cần phải được quan tâm đúng mức. Bài giảng sẽ giới thiệu công thức viết dẫn nhập có tên là CaRS (creating a research space). Bài giảng cũng sẽ giới thiệu những câu văn quen thuộc để giúp các bạn viết bài báo lần đầu

### **Ngày 4 (5/1/2018)**

- **Bài 6 Cách viết phần phương pháp (methods):** bài giảng sẽ giới thiệu những phần thông tin liên quan đến thiết kế nghiên cứu, đối tượng nghiên cứu, phương pháp đo lường, phương pháp phân tích, v.v. cho một bài báo y học. Mỗi mục sẽ được minh họa bằng những câu văn quen thuộc hay những mô tả đã được công bố trên các tập san y học nổi tiếng trên thế giới.
- **Bài 7 Cách viết phần kết quả (results):** kết quả là trái tim của một công trình nghiên cứu. Nhưng viết phần Kết quả có khi là một thách thức đáng kể cho những người mới vào nghiên cứu, vì không biết viết cái gì trước và cái gì sau, hay viết sao cho thuyết phục. Bài giảng này sẽ trình bày một số phương pháp và nguyên tắc quan trọng trong việc mô tả kết quả nghiên cứu. Phần đầu sẽ bàn về cách viết.
- **Bài 8 Cách trình bày biểu đồ / bảng số liệu:** phần hai hướng dẫn cách thiết kế bảng số liệu và những biểu đồ có phẩm chất cao.

#### *Ngày 5 (6/1/2018)*

- **Bài 9 Cách viết phần bàn luận (discussion):** đây là phần khó viết nhất trong một bài báo khoa học. Các sách hướng dẫn viết bài báo khoa học bằng tiếng Anh ít khi bàn đến phần này một cách cụ thể. Ngay cả những người ở cấp giáo sư vẫn viết phần bàn luận chưa đạt. Trong bài giảng này, các bạn sẽ học một công thức đơn giản (gồm 6 đoạn văn) nhưng rất hiệu quả trong việc cấu trúc phần bàn luận.
- **Bài 10 Cách chọn tập san:** chọn tập san thích hợp để công bố kết quả nghiên cứu đang là một vấn đề thời sự, vì có quá nhiều tập san "dòm" trên thế giới. Điều nguy hiểm hơn nữa là có những tập san nằm ở biên giới dòm và thật. Rất nhiều nhà khoa học Việt Nam đã trở thành nạn nhân của những tập san dòm. Bài giảng này sẽ giới thiệu các tiêu chuẩn để giúp các bạn phân biệt tập san dòm và tập san thật, cùng những tiêu chí để chọn tập san chuyên ngành thích hợp cho nghiên cứu.

#### *Ngày 6 (7/1/2018)*

- **Bài 11 Cách trả lời bình duyệt:** sau khi bài báo được gửi đi, nếu may mắn sẽ được gửi ra ngoài bình duyệt (peer review) bởi các chuyên gia. Nhiệm vụ của tác giả là trả lời những bình luận, phê bình, câu hỏi của chuyên gia bình duyệt. Trả lời bình duyệt vừa là một khoa học vừa là một nghệ thuật. Bài giảng này sẽ hướng dẫn các bạn cách trả lời cụ thể mà các tập san y khoa ưa chuộng.
- **Bài 12 Cách viết thư đệ trình (cover letter):** hướng dẫn cách viết cover letter sao cho thuyết phục và gây chú ý của ban biên tập.
- **Bài 13 Chia sẻ về cách viết cảm tạ, tài trợ, tài liệu tham khảo, và cách nộp bài báo:** trong bài này một tác giả ở Việt Nam sẽ chia sẻ các nộp bài báo khoa học trực tuyến.