

Buổi 6 (Trình diễn trực tiếp): cỡ mẫu, so sánh các nhóm, mô hình và dữ liệu thiếu

Nền tảng Nghiên cứu Lâm sàng · Tuần 6, Thứ Bảy 19:30 đến 22:30 (trực tuyến) · Neudata · kịch bản cho giảng viên

	Phần A: cỡ mẫu	Phần B: so sánh các nhóm	Phần C: mô hình và dữ liệu thiếu
Thời điểm	6.1, sau slide <i>Cỡ mẫu: tác động lớn hơn, ít nhiều hơn</i> (khoảng 20:05)	6.2 (khoảng 20:21)	6.4, trong khi chiếu các slide của mục này (21:30 đến 21:48)
Thời lượng	5 phút (Bước 2, 3 và một thay đổi; bản đầy đủ 8 phút nếu còn thời gian)	6 phút	lồng vào các slide 6.4 (tổng cộng khoảng 8 phút chạy R, tùy chọn)
Công cụ	công cụ tính trực tuyến miễn phí (ví dụ OpenEpi) và <code>Course/Resources/tools/sample_size_calculator.xlsx</code>	<code>Course/Data/R/03_compare.R</code>	<code>Course/Data/R/04_models.R</code> , <code>05_missing.R</code>

Học viên **theo dõi màn hình chia sẻ của giảng viên và dự đoán trong khung chat**. Học viên không tự tính và không chạy code. FLUID-ICU là dữ liệu **mô phỏng**, và mọi con số về FLUID-ICU dưới đây lấy từ `Course/Data/key_results.md`.

Lưu ý: bảng tính và kết quả R hiển thị bằng tiếng Anh. Kịch bản giữ nguyên tên trang tính, tên ô và đầu ra như trên màn hình.

Trước buổi học

- [] Mở thẻ trình duyệt với công cụ tính trực tuyến, thu phóng 150%, và mở sẵn bảng tính ở trang `3_Means` để dự phòng.
- [] Mở `Course/Data/FLUID-ICU.Rproj`. Tối hôm trước, chạy trọn `03_compare.R`, `04_models.R` và `05_missing.R` một lần từ đầu đến cuối (`05_missing.R` chạy 20 lần quy gán: cần khoảng một phút).
- [] Mở sẵn các hình dự phòng: `Course/Resources/figures/vi/data/` (`crude_vs_adjusted.png`, `forest_adjusted.png`, `km_balance.png`, `missing_lactate.png`, `cc_vs_mi.png`; bản tiếng Anh ở `Course/Resources/figures/en/data/`). Bảng so sánh giờ là bảng trực tiếp trên slide, *FLUID-ICU: bốn kết cục, bốn phép so sánh*.
- [] Một slide hoặc tài liệu dùng chung liệt kê bốn đầu vào để tính cỡ mẫu (tác động dự kiến, độ nhiễu (SD), alpha, lực thống kê) cộng với tỷ lệ bỏ cuộc.

Phần A: Cỡ mẫu trong 5 phút (công cụ trực tuyến và bảng tính)

Ở bản 5 phút, chạy Bước 2 và 3 và một thay đổi trong Bước 4 (lực thống kê 90%). Bước 1 và 5 mỗi bước nói một câu. Chỉ dùng bản đầy đủ 8 phút nếu lớp đang đi trước giờ.

Bước 1: Đặt vấn đề (1 phút). *"Mọi phép tính cỡ mẫu cần cùng những thành phần: anh chị dự kiến kết quả ở mỗi nhóm là bao nhiêu, phép đo nhiễu đến mức nào, nguy cơ báo động nhầm mà anh chị*

chấp nhận (alpha) và anh chị muốn chắc chắn đến đâu rằng sẽ phát hiện được một tác động có thật (lực thống kê). Phần mềm làm phần số học. Các thành phần là việc chuyên môn lâm sàng của anh chị."

Bước 2: Tái lập Bài báo 1 (hai trung bình, 2 phút). Chênh lệch trung bình / hai trung bình độc lập: nhóm 1 = 70 ml/kg, nhóm 2 = 40 ml/kg, SD 35 ở cả hai nhóm, độ tin cậy 95% (alpha 5% hai phía), lực thống kê 80%, tỷ lệ 1.

Dự kiến: **22 mỗi nhóm (tổng 44)**. Cộng thêm 10% cho tử vong được 25 mỗi nhóm, **tổng 50**, đúng như phần phương pháp của Bài báo 1. Bảng tính 3_Means: B20 = 22, B21 = 25, B22 = 50.

Bước 3: Câu hỏi thật của chúng ta, tử vong (hai tỷ lệ, 2 phút). 40% (chăm sóc thường quy) so với 30% (hạn chế dịch), alpha 5%, lực thống kê 80%, tỷ lệ 1.

Dự kiến mỗi nhóm: Fleiss **356**, có hiệu chỉnh liên tục **376** (công cụ có thể lệch ±1–2; Kelsey khoảng 358). Bảng tính 2_Proportions: B20 = 356, B21 = 376, sau khi cộng 10% mất theo dõi B22 = 396 mỗi nhóm, B23 = 792 tổng cộng.

"Cùng câu hỏi lâm sàng, kết cục khác: cân bằng dịch cần 50 người bệnh, tử vong cần gần 800."

Bước 4: Mỗi lần thay đổi một đầu vào (2 phút). Trước tiên mời cả lớp đoán trong khung chat: tăng hay giảm.

Thay đổi (hai tỷ lệ, Fleiss, không hiệu chỉnh)	n mỗi nhóm	Thông điệp
Ban đầu 40% so với 30%, lực thống kê 80%	356	
Lực thống kê 90%	477	muốn chắc chắn hơn thì cần thêm người bệnh
40% so với 35%	1471	khác biệt giảm một nửa thì n tăng khoảng bốn lần
40% so với 25%	152	tác động lớn chỉ cần ít người bệnh, nhưng có thực tế không?
Alpha 1%	530	ngưỡng báo động chặt hơn cần thêm người bệnh

Bước 5: Câu văn mẫu (1 phút). Chỉ vào câu màu xanh lá ở cuối trang 2_Proportions (bằng tiếng Anh). Nghĩa của câu: "Giả định kết cục xảy ra ở 40% người bệnh nhóm 1 và 30% ở nhóm 2, 356 người bệnh mỗi nhóm cho lực thống kê 80% ... chúng tôi sẽ tuyển 396 người bệnh mỗi nhóm (tổng cộng 792)."

Liên hệ với FLUID-ICU: SAP của nghiên cứu (mục 11) nêu rằng khoảng 600 người bệnh cho lực thống kê khoảng 80% với OR 1.7 và khoảng 55% với OR 1.5: "chỉ đủ lực thống kê cho một mối liên quan ở mức vừa phải." Đó là lý do KTC sau hiệu chỉnh của nghiên cứu này rộng.

Phần B: So sánh hai nhóm trong FLUID-ICU (03_compare.R, 6 phút)

Trước mỗi khối, mời cả lớp gõ tên kiểm định vào khung chat, dựa vào bảng quyết định.

Khối	Kết cục (loại)	Kiểm định	Kết quả dự kiến (key_results.md)
1	Tuổi (đối xứng)	kiểm định t Welch	trung bình 64.6 so với 61.4 tuổi
2–3	SOFA, lactat (lệch / điểm số)	Mann–Whitney	SOFA 6 so với 8; lactat 2.2 so với 2.9 mmol/L (91 người không có lactat bị loại khỏi phép so sánh)
4	Tử vong 28 ngày (nhị phân)	bảng 2×2 và khi bình phương	134/373 (35.9%) so với 45/245 (18.4%); RD 17.6 điểm phần trăm (KTC 95% 10.7 đến 24.4); RR 1.96; p < 0.001
5	AKI (nhị phân)	khi bình phương	230/373 (61.7%) so với 71/245 (29.0%)
6	Số ngày không thở máy (0–28, hai cụm)	Mann–Whitney	trung vị 20 (0 đến 28) so với 28 (22 đến 28) ; p < 0.001
7	Thời gian nằm ICU ở người sống (lệch): lỗi thường gặp	kiểm định t so với Mann–Whitney	chênh lệch trung bình theo kiểm định t 2.7 ngày so với Hodges–Lehmann 2.5 ngày (KTC 95% 1.8 đến 3.2); trung vị 7.5 so với 4.8

Ý giảng dạy: kiểm định đi theo loại kết cục, tác động được báo cáo kèm KTC 95% chứ không chỉ giá trị p, và mọi phép so sánh ở đây đều **chưa hiệu chỉnh** (nhóm dương nặng hơn: xem Bảng 1).

Dự phòng: slide *FLUID-ICU: bốn kết cục, bốn phép so sánh* (mọi kết quả trong một bảng).

Phần C: Mô hình và dữ liệu thiếu (04_models.R và 05_missing.R, cùng các slide 6.4)

Chạy từng khối khi đến slide tương ứng. Nếu thiếu thời gian, chỉ chiếu các hình (chúng chính là các slide).

Slide	Khối trong script	Kết quả dự kiến (key_results.md)	Nói
SAP có trước	mở SAP_FLUID-ICU.md	yếu tố gây nhiễu: tuổi, SOFA, lactat, thuốc vận mạch, thở máy, bệnh thận mạn; một tương tác (suy tim); CC và MI	"Đã quyết định trước khi ai đó nhìn thấy kết quả."
OR thô và OR hiệu chỉnh	04: khối 1–2	OR thô 2.49 (1.70 đến 3.70) , n = 618; OR thô trên cùng 527 người: 2.22 (1.48 đến 3.39) ; OR hiệu chỉnh 1.62 (1.02 đến 2.59) , p = 0.044, n = 527	"Nhiều do chỉ định, giờ đã nhìn thấy được."
Biểu đồ rừng	04: khối 3–4	mô hình hiệu chỉnh, hàng phơi nhiễm 1.62	"Chỉ hàng phơi nhiễm là câu hỏi của ta."

Slide	Khối trong script	Kết quả dự kiến (key_results.md)	Nói
Kaplan–Meier	04: khối 5	tỷ lệ sống ngày 28 81.6% so với 64.1% ; log-rank $p < 0.001$ (khi bình phương 21.4)	"Bậc thang, vạch kiểm duyệt, khoảng cách, số người còn nguy cơ."
Cox	04: khối 6	HR thô 2.17 (1.55 đến 3.05) , HR hiệu chỉnh 1.46 (1.00 đến 2.13) , $p = 0.053$; kiểm định PH tổng thể $p = 0.10$	"0.044 so với 0.053: cùng một câu chuyện."
(bổ sung) Tương tác	04: khối 7	suy tim OR 4.39 (1.65 đến 12.38) , $n = 90$; không suy tim 1.29 (0.77 đến 2.19) , $n = 437$; p tương tác 0.031	"Đã định trước, hợp lý, cần được xác nhận."
Ai bị thiếu?	05: khối 1–2	thiếu lactat 93/620 (15.0%) ; SOFA 5 so với 8; tử vong 19.4% so với 30.6%	"Không phải MCAR; MAR là hợp lý."
CC so với MI	05: khối 3–5	Q4 OR hiệu chỉnh cho mỗi mmol/L: CC 1.39 (1.24 đến 1.57) , $n = 527$ so với MI 1.45 (1.29 đến 1.64) , $n = 620$; Q1 OR hiệu chỉnh với MI 1.70 (1.09 đến 2.64)	"Vững: câu trả lời gần như không đổi."

Chỉ cho lớp thấy trong `05_missing.R` khối 4: mô hình quy gán có cả **kết cục** và các yếu tố dự báo việc thiếu dữ liệu (SOFA, MAP, thuốc vận mạch, ICU), 20 bộ dữ liệu quy gán, gộp theo quy tắc Rubin. Biểu đồ mật độ cho thấy lactat được quy gán thấp hơn (trung vị 2.4 so với 2.6 mmol/L ở các trường hợp đầy đủ), đúng như dự kiến vì những người bệnh bị thiếu lactat nhẹ hơn.

Kế hoạch dự phòng

Sự cố	Cách xử lý
Không có internet (Phần A)	Chạy Bước 2–4 trong bảng tính (ô màu vàng là ô nhập).
R lỗi (Phần B–C)	Mọi kết quả đều có trên slide, nên giảng giải trên slide.
mice chạy chậm	Bỏ phần chạy trực tiếp và chiếu <code>cc_vs_mi.png</code> .
Bị trễ giờ	Phần A: chỉ Bước 2–3. Phần B: khối 4 và 7. Phần C: chỉ các hình.

Các câu hỏi có thể gặp

"Kết quả cỡ mẫu nào đúng, 356 hay 376?" Cả hai đều chuẩn. Chúng dùng các phép xấp xỉ hơi khác nhau. Ghi rõ phương pháp trong đề cương và dùng nhất quán.

"Sao OR giảm sau khi hiệu chỉnh?" Người bệnh nặng hơn được truyền nhiều dịch hơn và tử vong nhiều hơn. Hiệu chỉnh so sánh những người bệnh có mức độ nặng tương tự, nên phần mối liên quan do mức độ nặng biến mất.

"p của mô hình Cox là 0.053, vậy là không có tác động?" Không. HR (1.46) và OR (1.62) tương tự nhau, KTC chồng lấp, và bằng chứng là như nhau. Hãy đọc ước lượng và KTC, đừng chỉ nhìn xem p nằm bên nào của 0.05.

"Quy gán đa bội chẳng phải là bịa dữ liệu sao?" Phương pháp này điền mỗi chỗ trống nhiều lần bằng các giá trị hợp lý được dự đoán từ dữ liệu đã ghi nhận, và nói rộng KTC để phản ánh sự không chắc chắn. Phương pháp hợp lệ khi dữ liệu là MAR. Luôn so sánh với phân tích trường hợp đầy đủ.

"Sao không hiệu chỉnh theo AKI?" AKI xảy ra sau khi truyền dịch và có thể nằm trên đường nhân quả (một biến trung gian). Hiệu chỉnh theo AKI sẽ loại bỏ một phần tác động mà ta muốn ước lượng. SAP loại AKI vì lý do đó.