

# Từ điển dữ liệu FLUID-ICU

> **DỮ LIỆU MÔ PHỎNG: KHÔNG PHẢI NGƯỜI BỆNH THẬT.** FLUID-ICU được tạo bởi R/00\_generate\_fluid\_icu.R để giảng dạy.

**Nghiên cứu:** FLUID-ICU, một nghiên cứu thuần tập tiến cứu đa trung tâm ở người lớn bị nhiễm khuẩn huyết nhập 4 ICU (ICU A-D), từ tháng 1/2023 đến tháng 12/2024.

**Tệp:** fluid\_icu\_raw.csv / .xlsx (đúng như các đơn vị gửi về, 623 dòng, dùng trong phần trình diễn làm sạch) và fluid\_icu\_clean.csv / .rds (620 người bệnh, 32 biến, sẵn sàng để phân tích).

**Phơi nhiễm chính:** cân bằng dịch tích lũy vào ngày 3 (cum\_fluid\_balance\_d3\_ml\_kg, positive\_balance\_d3). **Kết cục chính:** died\_28d (kèm time\_to\_event\_28d).

**Các yếu tố gây nhiễu xác định trước:** age, sofa, lactate\_mmol\_l, vasopressor, mech\_vent, ckd. **Yếu tố thay đổi tác động:** heart\_failure.

Trong tệp sạch, giá trị thiếu là một ô trống (CSV) hoặc NA (R). Cột *Mã giá trị thiếu (tệp thô)* cho biết tệp thô ghi giá trị thiếu như thế nào. Cột *Khoảng cho phép* là phép kiểm tra tính hợp lý dùng trong R/01\_clean.R; giá trị nằm ngoài khoảng này được đặt là thiếu và gửi yêu cầu xác minh (query). Các biến **DẪN XUẤT** được tính trong bước làm sạch và không có trong tệp thô.

Tên biến và các giá trị mã hóa (No / Yes, Female / Male, ICU A...) giữ nguyên tiếng Anh như trong tệp dữ liệu.

| # | Biến     | Nhãn                 | Loại                 | Đơn vị / mã                   | Khoảng cho phép | Mã giá trị thiếu (tệp thô) | Ghi chú                                                                                                     |
|---|----------|----------------------|----------------------|-------------------------------|-----------------|----------------------------|-------------------------------------------------------------------------------------------------------------|
| 1 | study_id | Mã số nghiên cứu     | Ký tự (mã định danh) | FLU-001 ... FLU-620           | Duy nhất        | Không được thiếu           | Cột thô 'Study ID'. Tệp thô có 3 mã được nhập hai lần (các dòng trùng hoàn toàn); đã loại khi làm sạch.     |
| 2 | site     | ICU tuyển người bệnh | Phân loại (4 mức)    | ICU A / ICU B / ICU C / ICU D | 4 mức           | Không được thiếu           | Cột thô 'Site' được gõ theo nhiều kiểu (icu a, ICU-A, ICUA); đã chuẩn hóa. ICU C báo cáo lactat theo mg/dL. |

| # | Biến           | Nhãn                       | Loại              | Đơn vị / mã             | Khoảng cho phép           | Mã giá trị thiếu (tệp thô) | Ghi chú                                                                                                       |
|---|----------------|----------------------------|-------------------|-------------------------|---------------------------|----------------------------|---------------------------------------------------------------------------------------------------------------|
| 3 | admission_date | Ngày nhập ICU (ngày tuyển) | Ngày              | YYYY-MM-DD              | 2023-01-01 đến 2024-12-31 | Không được thiếu           | Cột thô 'Admission Date' lần lộn các định dạng YYYY-MM-DD, DD/MM/YYYY và DD-Mon-YYYY.                         |
| 4 | age            | Tuổi                       | Số (số nguyên)    | Năm                     | 18 đến 110                | Để trống                   | Cột thô 'Age'. Một giá trị không thể có (210) được đặt là thiếu; đã gửi query cho đơn vị.                     |
| 5 | sex            | Giới                       | Phân loại (2 mức) | Female / Male           | 2 mức                     | Không được thiếu           | Cột thô 'Sex' có 8-9 cách viết (M, m, male, F, f, female, 'Female').                                          |
| 6 | weight_kg      | Cân nặng lúc nhập ICU      | Số                | kg (1 chữ số thập phân) | 30 đến 250                | Để trống                   | Cột thô 'Weight (kg)'. Một giá trị không thể có (7) được đặt là thiếu. Dùng để tính cân bằng dịch theo ml/kg. |
| 7 | diabetes       | Đái tháo đường             | Nhị phân          | No / Yes                | No, Yes                   | Không có                   | Mã thô Yes/No/Y/N/yes/no/1/0 (ICU B gửi 1/0).                                                                 |
| 8 | ckd            | Bệnh thận mạn              | Nhị phân          | No / Yes                | No, Yes                   | Không có                   | Yếu tố gây nhiễu xác định trước. Tệp thô mã hóa Yes/No lần lộn.                                               |
| 9 | heart_failure  | Suy tim mạn                | Nhị phân          | No / Yes                | No, Yes                   | Không có                   | Yếu tố thay đổi tác động xác định trước (Q1). Tệp thô mã hóa lần lộn.                                         |

| #  | Biến             | Nhãn                                       | Loại              | Đơn vị / mã                                                                                                   | Khoảng cho phép | Mã giá trị thiếu (tập thô) | Ghi chú                                                                                                                                               |
|----|------------------|--------------------------------------------|-------------------|---------------------------------------------------------------------------------------------------------------|-----------------|----------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------|
| 10 | hypertension     | Tăng huyết áp                              | Nhị phân          | No / Yes                                                                                                      | No, Yes         | Không có                   | Tập thô mã hóa lẫn lộn.                                                                                                                               |
| 11 | infection_source | Nguồn nhiễm khuẩn                          | Phân loại (5 mức) | Respiratory (hô hấp) / Urinary (tiết niệu) / Abdominal (ổ bụng) / Skin/soft tissue (da/mô mềm) / Other (khác) | 5 mức           | Không có                   | Thứ tự như liệt kê, dùng cho các bảng.                                                                                                                |
| 12 | sofa             | Điểm SOFA lúc nhập ICU                     | Số (điểm nguyên)  | Điểm                                                                                                          | 0 đến 24        | Không có                   | Mức độ nặng; yếu tố gây nhiễu xác định trước. Trung vị 7 (IQR 5-9).                                                                                   |
| 13 | apache_ii        | Điểm APACHE II (24 giờ đầu)                | Số (điểm nguyên)  | Điểm                                                                                                          | 0 đến 71        | -99                        | Cột thô 'APACHE II'; 6 giá trị mã -99 = không ghi nhận.                                                                                               |
| 14 | lactate_mmol_l   | Lactat máu lúc nhập ICU                    | Số                | mmol/L (1 chữ số thập phân)                                                                                   | 0.1 đến 30      | Đề trống, 999, -99         | Thiếu khoảng 15%, thường gặp hơn ở người bệnh ít nặng (MAR). Giá trị của ICU C trong tập thô tính bằng mg/dL (x 9.01): đã chia cho 9.01 khi làm sạch. |
| 15 | map_mmhg         | Huyết áp động mạch trung bình lúc nhập ICU | Số (số nguyên)    | mmHg                                                                                                          | 20 đến 200      | 999                        | Cột thô 'MAP (mmHg)'; 4 giá trị mã 999.                                                                                                               |
| 16 | vasopressor      | Thuốc vận mạch lúc nhập ICU                | Nhị phân          | No / Yes                                                                                                      | No, Yes         | Không có                   | Yếu tố gây nhiễu xác định trước. Tập thô mã hóa lẫn lộn.                                                                                              |

| #  | Biến                       | Nhãn                                 | Loại           | Đơn vị / mã                | Khoảng cho phép | Mã giá trị thiếu (tệp thô) | Ghi chú                                                                                                                                                                                   |
|----|----------------------------|--------------------------------------|----------------|----------------------------|-----------------|----------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 17 | mech_vent                  | Thở máy xâm nhập lúc nhập ICU        | Nhị phân       | No / Yes                   | No, Yes         | Không có                   | Cột thô 'Mech vent'. Yếu tố gây nhiễu xác định trước.                                                                                                                                     |
| 18 | fluid_in_d1_ml             | Tổng dịch vào, ngày ICU 1            | Số (số nguyên) | ml                         | 0 đến 20000     | Để trống                   | Toàn bộ dịch đường tĩnh mạch và đường tiêu hóa, kể cả thuốc và dinh dưỡng. Cột thô 'Fluid in D1 (ml)'.                                                                                    |
| 19 | fluid_out_d1_ml            | Tổng dịch ra, ngày ICU 1             | Số (số nguyên) | ml                         | 0 đến 20000     | Để trống                   | Nước tiểu, dẫn lưu, siêu lọc; không gồm lượng mất không nhận biết.                                                                                                                        |
| 20 | fluid_in_d2_ml             | Tổng dịch vào, ngày ICU 2            | Số (số nguyên) | ml                         | 0 đến 20000     | Để trống                   |                                                                                                                                                                                           |
| 21 | fluid_out_d2_ml            | Tổng dịch ra, ngày ICU 2             | Số (số nguyên) | ml                         | 0 đến 20000     | Để trống                   | Một giá trị âm (không thể có) trong tệp thô được đặt là thiếu.                                                                                                                            |
| 22 | fluid_in_d3_ml             | Tổng dịch vào, ngày ICU 3            | Số (số nguyên) | ml                         | 0 đến 20000     | Để trống                   |                                                                                                                                                                                           |
| 23 | fluid_out_d3_ml            | Tổng dịch ra, ngày ICU 3             | Số (số nguyên) | ml                         | 0 đến 20000     | Để trống                   |                                                                                                                                                                                           |
| 24 | cum_fluid_balance_d3_ml_kg | Cân bằng dịch tích lũy, ngày ICU 1-3 | Số (DẪN XUẤT)  | ml/kg (1 chữ số thập phân) | -150 đến 250    | Dẫn xuất                   | = (tổng dịch vào ngày 1-3 - tổng dịch ra ngày 1-3) / weight_kg. Không có trong tệp thô. Thiếu ở 2 người bệnh (thiếu cân nặng hoặc dịch ra). Phơi nhiễm chính (Q2, Q3: theo mỗi 10 ml/kg). |

| #  | Biến                | Nhãn                                                        | Loại                | Đơn vị / mã               | Khoảng cho phép | Mã giá trị thiếu (tập thô) | Ghi chú                                                                                                                                              |
|----|---------------------|-------------------------------------------------------------|---------------------|---------------------------|-----------------|----------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------|
| 25 | positive_balance_d3 | Cân bằng dịch tích lũy dương vào ngày 3                     | Nhị phân (DẪN XUẤT) | No / Yes                  | No, Yes         | Dẫn xuất                   | Yes nếu <code>m_fluid_balance_d3_m_l_kg &gt; 0</code> . Phơi nhiễm chính cho Q1 và Q5.                                                               |
| 26 | aki_stage           | Tổn thương thận cấp, giai đoạn KDIGO cao nhất, ngày ICU 1-7 | Thứ bậc (số nguyên) | 0 = không, 1, 2, 3        | 0 đến 3         | Không có                   | Cột thô 'AKI stage (KDIGO)'.                                                                                                                         |
| 27 | aki_any             | Tổn thương thận cấp bất kỳ (KDIGO giai đoạn $\geq 1$ )      | Nhị phân (DẪN XUẤT) | No / Yes                  | No, Yes         | Dẫn xuất                   | Yes nếu <code>aki_stage <math>\geq</math> 1</code> . Kết cục cho Q2.                                                                                 |
| 28 | rrt                 | Điều trị thay thế thận đến ngày 28                          | Nhị phân            | No / Yes                  | No, Yes         | Không có                   | Tập thô mã hóa lẫn lộn.                                                                                                                              |
| 29 | vent_free_days_28   | Số ngày không thở máy đến ngày 28                           | Số (số nguyên)      | Ngày                      | 0 đến 28        | Không có                   | Số ngày còn sống và không thở máy xâm nhập đến ngày 28; bằng 0 nếu tử vong; bằng 28 nếu không thở máy và còn sống. Cột thô 'VFD 28'. Kết cục cho Q3. |
| 30 | icu_los_days        | Thời gian nằm ICU                                           | Số                  | Ngày (1 chữ số thập phân) | $> 0$           | Không có                   | Lệch phải. Với người bệnh tử vong, thời gian nằm kết thúc khi tử vong. Kết cục cho Q5 (người sống).                                                  |
| 31 | died_28d            | Tử vong do mọi nguyên nhân trong 28 ngày                    | Nhị phân            | No / Yes                  | No, Yes         | Không có                   | Kết cục chính (Q1, Q4). Tập thô mã hóa lẫn lộn.                                                                                                      |

| #  | Biến              | Nhãn                                                  | Loại           | Đơn vị / mã | Khoảng cho phép | Mã giá trị thiếu (tệp thô) | Ghi chú                                                                                                                                             |
|----|-------------------|-------------------------------------------------------|----------------|-------------|-----------------|----------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------|
| 32 | time_to_event_28d | Thời gian từ khi nhập ICU đến tử vong hoặc kiểm duyệt | Số (số nguyên) | Ngày        | 1 đến 28        | Không có                   | Ngày tử vong nếu tử vong; 28 (kiểm duyệt) nếu còn sống tại ngày 28. Dùng cùng died_28d cho Kaplan-Meier và Cox. Cột thô 'Time to event 28d (days)'. |

Tên cột trong tệp thô được đặt theo kiểu bảng tính (ví dụ `Weight (kg)`, `MAP (mmHg)`); `janitor::clean_names()` cùng bốn lần đổi tên cho ra các tên ở trên (xem `R/01_clean.R`, bước 1).