

Trường Đại học Y tế công cộng
Mạng lưới Nghiên cứu Khoa học Sức khỏe Việt Nam

PHƯƠNG PHÁP CHỌN MẪU VÀ TÍNH TOÁN CỖ MẪU
TRONG NGHIÊN CỨU KHOA HỌC SỨC KHỎE

Chủ biên:

GS.TS. Hoàng Văn Minh, Trường Đại học Y tế công cộng

GS.TS. Lưu Ngọc Hoạt, Trường Đại học Y Hà Nội

Hà Nội- Tháng 8 năm 2020

TÁC GIẢ CUỐN TÀI LIỆU (THEO THỨ TỰ ABC)

PGS.TS. Đào Thị Minh An, Trường Đại học Y Hà Nội

PGS.TS. Nguyễn Thùy Dương, Viện Vệ sinh Dịch tễ Trung Ương

PGS.TS. Kim Bảo Giang, Trường Đại học Y Hà Nội

GS.TS. Lưu Ngọc Hoạt, Trường Đại học Y Hà Nội

PGS.TS. Phạm Minh Khuê, Trường Đại học Y dược Hải Phòng

BS. Khương Quỳnh Long, Trường Đại học Y tế công cộng

ThS. Trần Hùng Minh, Trung tâm Sáng kiến Sức khỏe và Dân số

GS.TS. Hoàng Văn Minh, Trường Đại học Y tế công cộng

PGS.TS. Võ Văn Thắng, Trường Đại học Y dược Huế

BS Ong Phúc Thịnh, Mạng lưới Nghiên cứu Khoa học Sức khỏe Việt Nam

TÁC GIẢ PHẦN MỀM TÍNH CỠ MẪU HSS 1.0 (<http://comau.tk>)

BS. Khương Quỳnh Long, Trường Đại học Y tế công cộng

BS Ong Phúc Thịnh, Mạng lưới Nghiên cứu Khoa học Sức khỏe Việt Nam

GS.TS. Hoàng Văn Minh, Trường Đại học Y tế công cộng

MỤC LỤC

CHƯƠNG I: MỘT SỐ KHÁI NIỆM CƠ BẢN.....	1
1. Quần thể và mẫu nghiên cứu.....	1
2. Chọn mẫu và tính toán cỡ mẫu	1
3. Phương pháp thống kê.....	1
4. Ước lượng khoảng tin cậy.....	2
5. Kiểm định giả thuyết	3
6. Sai lầm trong kiểm định giả thuyết.....	5
7. Mức ý nghĩa thống kê	6
8. Lực thống kê.....	6
9. Mức khác biệt và mức khác biệt có ý nghĩa thực tế	7
10. Kiểm định 1 phía và 2 phía.....	9
CHƯƠNG II: CHỌN MẪU NGHIÊN CỨU	10
1. Nguyên tắc chung	10
2. Chọn mẫu xác suất.....	11
2.1. Chọn mẫu ngẫu nhiên đơn (<i>simple random sampling</i>).....	11
2.2. Chọn mẫu hệ thống (<i>systematic sampling</i>)	12
2.3. Chọn mẫu ngẫu nhiên phân tầng (<i>stratified random sampling</i>).....	14
2.4. Chọn mẫu chùm (<i>cluster sampling</i>)	16
2.5. Chọn mẫu nhiều giai đoạn (<i>multistage sampling</i>)	18
3. Chọn mẫu không xác suất	19
3.1. Chọn mẫu thuận tiện (<i>convenience or accidental sampling</i>).....	19
3.2. Chọn mẫu chỉ tiêu (<i>quota sampling</i>)	19
3.3. Chọn mẫu có mục đích (<i>purposive sampling</i>)	19
3.4. Các ứng dụng của các kỹ thuật chọn mẫu không xác suất.....	19
3.5. Sự kết hợp giữa chọn mẫu xác suất và không xác suất	20
4. Chọn mẫu trong các thử nghiệm lâm sàng hoặc thực nghiệm	20
5. Đạo đức trong việc chọn mẫu nghiên cứu.....	21

CHƯƠNG III: TÍNH TOÁN CỖ MẪU NGHIÊN CỨU	22
1. Tính toán cỡ mẫu cho nghiên cứu sử dụng thống kê suy luận dựa trên ước lượng khoảng	22
<i>1.1.Nghiên cứu gồm 1 mẫu, xác định một trung bình.....</i>	<i>23</i>
<i>1.2.Nghiên cứu gồm 1 mẫu, xác định một tỷ lệ</i>	<i>26</i>
<i>1.3.Nghiên cứu gồm 2 mẫu độc lập, xác định sự khác biệt giữa 2 số trung bình</i>	<i>29</i>
<i>1.4.Nghiên cứu gồm 2 mẫu ghép cặp, xác định sự khác biệt giữa 2 số trung bình</i>	<i>31</i>
<i>1.5.Nghiên cứu gồm 2 mẫu độc lập, xác định sự khác biệt giữa 2 tỷ lệ.....</i>	<i>32</i>
2. Tính toán cỡ mẫu cho nghiên cứu sử dụng thống kê suy luận dựa trên kiểm định giả thuyết	33
<i>2.1.Nghiên cứu gồm 1 mẫu, kiểm định một trung bình</i>	<i>35</i>
<i>2.2.Nghiên cứu gồm 1 mẫu, kiểm định một tỷ lệ</i>	<i>37</i>
<i>2.3.Nghiên cứu gồm 2 mẫu độc lập, kiểm định 2 số trung bình.....</i>	<i>38</i>
<i>2.4.Nghiên cứu gồm 2 mẫu ghép cặp, kiểm định 2 số trung bình</i>	<i>40</i>
<i>2.5.Nghiên cứu gồm 2 mẫu độc lập, kiểm định 2 tỷ lệ.....</i>	<i>42</i>
<i>2.6.Nghiên cứu gồm 2 mẫu ghép cặp, kiểm định 2 tỷ lệ (McNemar)</i>	<i>44</i>
<i>2.7.Nghiên cứu bệnh chứng, kiểm định OR</i>	<i>46</i>
<i>2.8.Nghiên cứu thuần tập, kiểm định RR.....</i>	<i>48</i>
<i>2.9.Nghiên cứu sống còn (survival analysis study).....</i>	<i>50</i>
3. Nghiên cứu nghiệm pháp chẩn đoán.....	52
4. Nghiên cứu tương đương (Equivalence trial)	55
5. Nghiên cứu không kém hơn (Non-inferiority trial).....	59
6. Nghiên cứu thử nghiệm lâm sàng theo cụm (Cluster randomized design)	
62	
7. Cỡ mẫu cho mô hình hồi quy	64
8. Một số hiệu chỉnh.....	64
<i>8.1.Hiệu chỉnh đối với quần thể hữu hạn.....</i>	<i>64</i>

<i>8.2.Hiệu chỉnh theo hệ số thiết kế (Design effect)</i>	65
<i>8.3.Hiệu chỉnh tỷ lệ không trả lời, bỏ cuộc</i>	66
<i>8.4.Hiệu chỉnh khi cỡ mẫu 2 nhóm không bằng nhau</i>	67
PHỤ LỤC: CHỌN MẪU CHO QUẢN THỂ KHÓ TIẾP CẬN	69
1. Giới thiệu	69
2. Chọn mẫu RDS	70
3. Chọn mẫu TLS	77
TÀI LIỆU THAM KHẢO	81

Chương I: MỘT SỐ KHÁI NIỆM CƠ BẢN

GS.TS. Hoàng Văn Minh, BS. Khương Quỳnh Long

1. Quần thể và mẫu nghiên cứu

Trong mỗi nghiên cứu, quần thể nghiên cứu (*Population*) bao gồm toàn bộ các cá thể chúng ta đang quan tâm. Mẫu nghiên cứu (*Sample*) bao gồm các cá thể được nghiên cứu.

Do điều kiện nguồn lực có hạn, đối với mỗi nghiên cứu, chúng ta thường không thể tiến hành nghiên cứu trên toàn bộ quần thể mà thường chỉ tiến hành trên một mẫu.

2. Chọn mẫu và tính toán cỡ mẫu

Chọn mẫu (*Sampling*) là quá trình chọn ra các cá thể đại diện cho quần thể để đưa vào nghiên cứu. Để đảm bảo tính đại diện, cần áp dụng các kỹ thuật chọn mẫu xác suất (*Probabilistic Sampling*), hay còn gọi là chọn mẫu ngẫu nhiên (*Random sampling*).

Tính toán cỡ mẫu (*Sample size determination*) là việc tính toán số lượng cá thể được đưa vào mẫu nghiên cứu sao cho chúng ta có thể ngoại suy từ các đặc điểm của mẫu ra các đặc điểm tương ứng của quần thể (*Statistical inference*).

3. Phương pháp thống kê

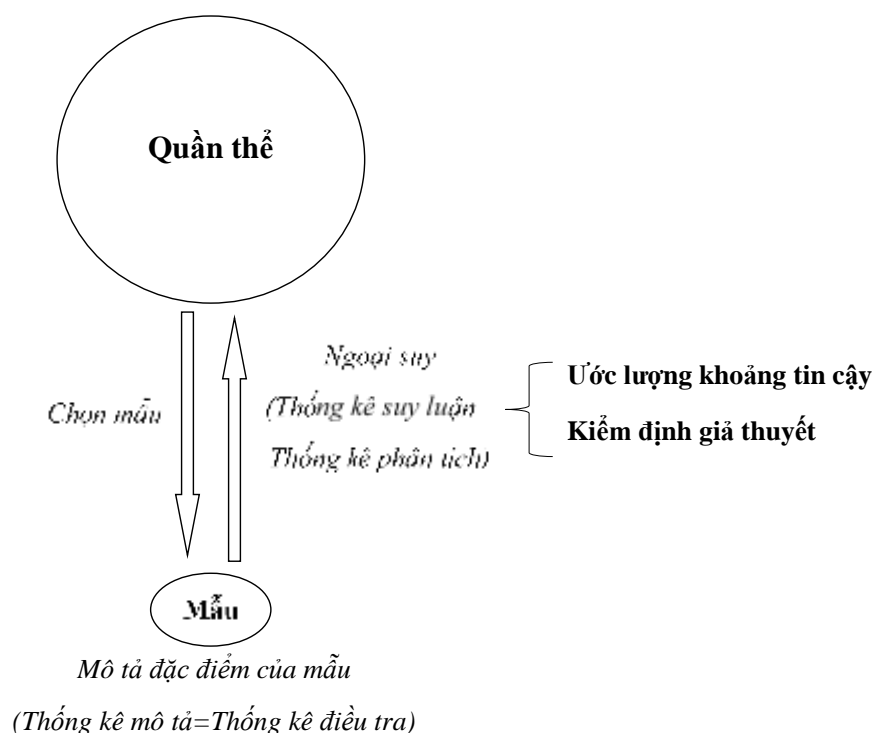
Sau khi tiến hành nghiên cứu trên mẫu, chúng ta thường thực hiện 2 loại phân tích thống kê, bao gồm:

- Thống kê mô tả (*Descriptive statistics*) hay còn gọi là thống kê điều tra (*Enumerative statistics*): Là việc mô tả các đặc điểm của mẫu nghiên cứu (các phân tích dựa trên các cá thể có trong mẫu);
- Thống kê suy luận (*Inference statistics*) hay còn gọi là thống kê phân tích (*Analytical statistics*): Là việc ngoại suy các đặc điểm của mẫu nghiên cứu ra thành các đặc điểm của quần thể nghiên cứu

Phần lớn các nghiên cứu đều có sử dụng thống kê suy luận/phân tích (thực chất là tiến hành ngoại suy).

Có 2 hai loại hình thống kê suy luận thường được áp dụng, bao gồm:

- Ước lượng khoảng tin cậy (*Confidence interval*): Dựa trên một giá trị tìm ra từ nghiên cứu để ước lượng một khoảng giá trị trong quần thể (với một mức tin cậy nhất định).
- Kiểm định giả thuyết (*Hypothesis testing*): Kiểm định liệu sự khác biệt, mối liên quan nào đó đã tìm ra ở mẫu nghiên cứu có xảy ra trong quần thể nghiên cứu hay không?



Hình 1: Một số khái niệm cơ bản

4. Ước lượng khoảng tin cậy

Ước lượng khoảng tin cậy (*Confidence interval*) là việc tính toán khoảng giá trị của tham số quần thể (trong khoảng từ X đến Y) với một mức tin cậy cho trước (các nghiên cứu khoa học sức khỏe thường dùng mức tin cậy 95%).

Khoảng tin cậy của biến định lượng được tính theo công thức:

- Cỡ mẫu lớn ($n \geq 30$): Tính theo phân bố z

$$95\% \text{ CI} = X \bar{\pm} Z \frac{\sigma}{\sqrt{n}}$$

Trong đó: X là giá trị trung bình, Z = 1,96, σ là độ lệch chuẩn, n là cỡ mẫu

- Cỡ mẫu nhỏ ($n < 30$): Tính theo phân bố t

$$95\%CI = X \bar{\mp} t \frac{\sigma}{\sqrt{n}}$$

t được tính tùy thuộc vào bậc tự do ($n-1$) (Xem bảng giá trị của t)

Khoảng tin cậy của biến định tính được tính theo công thức:

- Khi tỷ lệ ≥ 5 và ≤ 95 thì chúng ta sử dụng phương pháp xấp xỉ chuẩn (*Normal Approximation*):

$$95\%CI = p \bar{\mp} Z \sqrt{\frac{p(1-p)}{n}}$$

Trong đó: p là tỷ lệ mắc, $Z=1,96$, n là cỡ mẫu

- Khi tỷ lệ < 5 hoặc > 95 thì chúng ta áp dụng phương pháp Clopper–Pearson (hay còn gọi là *Exact methods*) dựa trên xác suất tích lũy của phân bố nhị phân (*Cumulative Probabilities of the Binomial Distribution*) để hiệu chỉnh. Hiện nay, phương pháp Clopper–Pearson (*Exact methods*) được áp dụng rất rộng rãi.

Giả sử, khoảng tin cậy 95% của tỷ lệ hiện mắc tăng huyết áp là nằm trong khoảng 10-20% có nghĩa là nếu chúng ta tiến hành nghiên cứu này 100 lần thì ít nhất có 95 lần giá trị khoảng tin cậy tính toán được (10-20%) sẽ chứa giá trị thực của quần thể.

Cỡ mẫu nghiên cứu càng lớn thì khoảng tin cậy càng hẹp, đồng nghĩa với tính chính xác (*precision*) của nghiên cứu càng cao, hay sai số ngẫu nhiên (random errors) càng nhỏ.

5. Kiểm định giả thuyết

Giả thuyết

Giả thuyết (*Hypothesis*) được hiểu đơn giản là những điều chúng ta kỳ vọng nhưng chưa được chứng minh là bất biến (khác với lý thuyết - *Theory*).

Trong nghiên cứu khoa học, có 2 loại giả thuyết được đưa ra:

- Giả thuyết H_0 (*Null hypothesis*): Thường đề cập đến “không có sự khác biệt” giữa các nhóm về giá trị biến số nghiên cứu nào đó

- Giả thuyết Ha (*Alternative hypothesis*): Do nhà nghiên cứu đưa ra, thường đề cập đến “có sự khác biệt” giữa các nhóm về giá trị biến số nghiên cứu nào đó

Kiểm định giả thuyết

Kiểm định giả thuyết (*Hypothesis testing*) là việc so sánh kết quả của nghiên cứu với giả thuyết H_0 . Kiểm định giả thuyết được thực hiện dựa trên các kiểm định thống kê (*Test statistics*).

Sau khi đặt ra các giả thuyết H_0 và H_a , chúng ta cần chứng minh H_0 hay H_a đúng. Theo lý thuyết của kiểm định giả thuyết thống kê, chúng ta rất khó hoặc không thể trực tiếp chứng minh một giả thuyết là đúng mà chỉ có thể chứng minh nó sai (bác bỏ), theo đó nếu có đủ bằng chứng cho thấy H_0 sai thì chúng ta bác bỏ giả thuyết H_0 và chấp nhận H_a , ngược lại nếu không đủ bằng chứng để bác bỏ H_0 thì chúng ta chấp nhận rằng H_0 đúng. Việc quyết định bác bỏ hay chấp nhận H_0 thường được thực hiện thông qua chỉ số P (*P-value*), chỉ số P là một xác suất có điều kiện mang ý nghĩa là xác suất quan sát được dữ liệu và những trường hợp hiếm hơn nếu giả thuyết H_0 đúng – $P(D|H_0)$. Theo đó nếu chỉ số P càng nhỏ càng cho thấy bằng chứng để bác bỏ giả thuyết H_0 , và khi chỉ số P nhỏ hơn mức ý nghĩa thống kê (α), thường được chọn là 0,05 (5%), thì chúng ta có thể kết luận có bằng chứng bác bỏ H_0 và chấp nhận H_a (sự khác biệt có ý nghĩa thống kê), ngược lại khi P lớn hơn α , chúng ta kết luận không đủ bằng chứng bác bỏ H_0 do đó chấp nhận H_0 .

****Kiểm định thống kê***

Kiểm định thống kê (*Test statistics*) là giá trị được tính toán và chuẩn hóa trong quá trình thực hiện kiểm định giả thuyết. Kiểm định thống kê bằng 0 có nghĩa kết quả nghiên cứu đúng như điều đã đề cập trong giả thuyết H_0 . Giá trị tuyệt đối của kiểm định thống kê càng lớn có nghĩa sự khác biệt giữa kết quả nghiên cứu và H_0 càng lớn.

Bản thân giá trị của kiểm định thống kê ít có ý nghĩa đối với kiểm định giả thuyết mà ta cần so sánh nó với phân bố của nó với giả định rằng nghiên cứu này

được lặp lại nhiều lần và Ho đúng để tính được xác suất xảy ra kiểm định thống kê nếu Ho đúng.

Nếu xác suất (*probability*) xảy ra kiểm định thống kê nếu Ho đúng nhỏ (thường chọn nhỏ hơn 5% hay 0,05), chúng ta kết luận là kết quả nghiên cứu của chúng ta khác với Ho. Nói cách khác là có thể bác bỏ Ho và chấp nhận Ha.

6. Sai lầm trong kiểm định giả thuyết

Thống kê là khoa học dựa vào nguyên lý xác suất, do đó không có một kết luận nào là hoàn toàn đúng hay hoàn toàn sai, các kết luận đều được diễn giải trong một số giả định hoặc sai số quy ước (chấp nhận được). Khi kết luận chấp nhận hay bác bỏ Ho, chúng ta đều có thể gặp phải những sai số tiềm tàng, và sai số này được chấp nhận trong một khoảng quy ước. Các tình huống xảy ra như bảng dưới đây.

Bảng 1: Các tình huống xảy ra khi kiểm định giả thuyết

Thực tế	Ho đúng	Ho sai
Kết quả nghiên cứu	(Không có sự khác biệt)	(Có sự khác biệt)
Chấp nhận Ho	Đúng ($1-\alpha$)	Sai lầm loại 2 (β)
Bác bỏ Ho	Sai lầm loại 1 (α)	Đúng ($1-\beta$)

Như vậy, có thể xảy ra 2 loại sai lầm:

- Sai lầm loại 1 (**hay α**) loại là bác bỏ giả thuyết Ho khi giả thuyết này đúng. Sai lầm loại 1 còn có thể hiểu là “dương tính giả”, tương tự như việc kết luận rằng bệnh nhân bị bệnh nhưng trên thực tế họ không bị bệnh. Trong kiểm định giả thuyết, sai lầm loại 1 thường được chọn ở mức 5%.
- Sai lầm loại 2 (**hay β**) là chấp nhận giả thuyết Ho khi giả thuyết này sai. Sai lầm loại 2 còn có thể hiểu là “âm tính giả”, tương tự như việc kết luận bệnh nhân không bị bệnh nhưng trên thực tế là họ bị bệnh. Trong kiểm định giả thuyết, sai lầm loại 2 thường được chọn ở mức <20%.

7. Mức ý nghĩa thống kê

Mức ý nghĩa thống kê (*Significance level*) đề cập đến ngưỡng quyết định bác bỏ hay chấp nhận giả thuyết H_0 dựa vào chỉ số P. Mức ý nghĩa thống kê thường được chọn là 0.05.

8. Lực thống kê

Lực thống kê (*Statistical power*) là xác suất bác bỏ giả thuyết H_0 (chấp nhận H_A) khi giả thuyết H_0 sai, hay có thể hiểu lực thống kê chính là xác suất dương tính thật. Lực thống kê đề cập đến khả năng loại bỏ sai lầm loại 2, hay lực thống kê bằng 1 trừ xác suất sai lầm loại 2, hay $\text{power} = 1 - \beta$. Lực thống kê thường được đòi hỏi là cao hơn 80% (sai lầm loại 2 <20%). Lực thống kê cần được tính toán và báo cáo trước khi nghiên cứu bắt đầu.

Như đã đề cập, lực thống kê là xác suất kết luận có sự khác biệt khi thực sự có sự khác biệt, hay lực thống kê biểu hiện cho khả năng phát hiện hiệu ứng khi nó thực sự tồn tại. Như vậy, nhà nghiên cứu đều mong muốn có được lực thống kê cao. Một nghiên cứu sẽ được đánh giá không cao nếu không đạt được lực thống kê cần thiết, thậm chí là không thể đưa ra kết luận nếu không đủ lực thống kê. Theo nguyên lý của kiểm định giả thuyết thống kê của Neyman và Pearson, có 3 trường hợp kết luận thống kê:

1. Nếu giá trị thống kê (z , t , χ^2 , F ...) nằm trong vùng bác bỏ H_0 , chúng ta bác bỏ H_0 và chấp nhận H_A
2. Nếu giá trị thống kê (z , t , χ^2 , F ...) nằm trong vùng chấp nhận H_0 , và có đủ lực thống kê, chúng ta chấp nhận H_0
3. Nếu giá trị thống kê (z , t , χ^2 , F ...) nằm trong vùng chấp nhận H_0 , và không có đủ lực thống kê, chúng ta không thể kết luận được gì. Vì khi đó chúng ta không thể quyết định là sự không khác biệt đó là do thực sự không có khác biệt hay do phép tính toán không đủ khả năng để phát hiện khác biệt

Như vậy lực thống kê đóng vai trò rất lớn cho tính giá trị của một kết luận thống kê. Trong thực hành, lực thống kê thường được quy ước là không nhỏ hơn

80%. Do đó, trong quá trình tính cỡ mẫu, nhà nghiên cứu cần phải tính toán cỡ mẫu để đáp ứng được điều này. Hơn nữa, sau khi thu thập số liệu, nhà nghiên cứu thường phải kiểm tra lại (*sensitive analysis*) xem với cỡ mẫu như vậy, lực thống kê đã đạt hay chưa

9. Mức khác biệt và mức khác biệt có ý nghĩa thực tế

Trong phương pháp kiểm định giả thuyết thống kê đã giới thiệu trong phần trước, giá trị P thường được sử dụng để kết luận sự khác biệt giữa các nhóm có ý nghĩa về mặt thống kê hay không. Việc kết luận dựa vào giá trị P mang ý nghĩa nhị phân – có/không có ý nghĩa thống kê, giá trị P không nói lên được *mức độ* của sự khác biệt. Trong thực hành, bên cạnh việc kết luận có hay không khác biệt, chúng ta cần biết mức độ khác biệt là bao nhiêu, và cỡ tác động (effect size - ES) là thể hiện mức độ khác biệt đó. Cỡ tác động là một thuật ngữ chung để chỉ các chỉ số khác nhau nhằm đo lường mức độ khác biệt giữa các nhóm, ví dụ khi so sánh huyết áp tâm thu giữa 2 nhóm dân số thì cỡ tác động chính là sự chênh lệch huyết áp tâm thu trung bình giữa 2 nhóm đo bằng đơn vị mmHg, hoặc khi thể hiện sự liên quan giữa cân nặng và chiều cao thì hệ số tương quan cũng là cỡ tác động. Đối với những nghiên cứu dịch tễ học có biến đo lường kết quả là biến nhị giá thì các chỉ số như tỉ số nguy cơ RR, tỉ số số chênh OR hay tỉ lệ hiện hành PR cũng đo lường mức độ sự khác biệt và được hiểu là cỡ tác động.

Cỡ tác động ảnh hưởng trực tiếp đến cỡ mẫu của một nghiên cứu, và thường là một thành phần trong công thức tính cỡ mẫu, cỡ tác động càng lớn (sự khác biệt giữa các nhóm càng lớn) thì cần ít mẫu hơn để phát hiện sự khác biệt này. Ngược lại nếu cỡ tác động nhỏ, chúng ta cần một cỡ mẫu lớn hơn. Một nhược điểm của cỡ tác động là khó so sánh kết quả nếu đơn vị đo lường khác nhau, ví dụ chúng ta sẽ khó so sánh sự khác biệt giữa chiều cao so với sự khác biệt về cân nặng. Một nhược điểm khác của cỡ tác động là không nói lên được mức độ giao động. Ví dụ sự khác biệt về chiều cao giữa 2 nhóm dân số của nghiên cứu thứ nhất (trung bình \pm độ lệch chuẩn) là 3 ± 1 cm sẽ rất khác so với nghiên cứu thứ 2 là 3 ± 5 cm, mặc dù sự khác biệt trung bình giữa 2 nhóm dân số của 2 nghiên cứu

đều là 3 cm. Do đó để có thể so sánh được các cỡ tác động với đơn vị đo lường khác nhau và tính toán đến sự giao động về kết quả, cỡ tác động chuẩn hóa (standardized effect size) được tính đến. Công thức tính cỡ tác động chuẩn hóa như sau:

- Đối với biến định lượng

$$ES = \frac{\mu_1 - \mu_0}{\sigma}$$

Trong đó

- μ_1 là kết quả nghiên cứu theo giả thuyết Ha
 - μ_0 là kết quả nghiên cứu theo giả thuyết Ho
 - σ là độ lệch chuẩn
- Đối với biến định tính

$$ES = \frac{p_1 - p_0}{\sqrt{p_0(1 - p_0)}}$$

Trong đó:

- n là cỡ mẫu
- p_1 là kết quả theo giả thuyết Ha
- p_0 là theo giả thuyết Ho

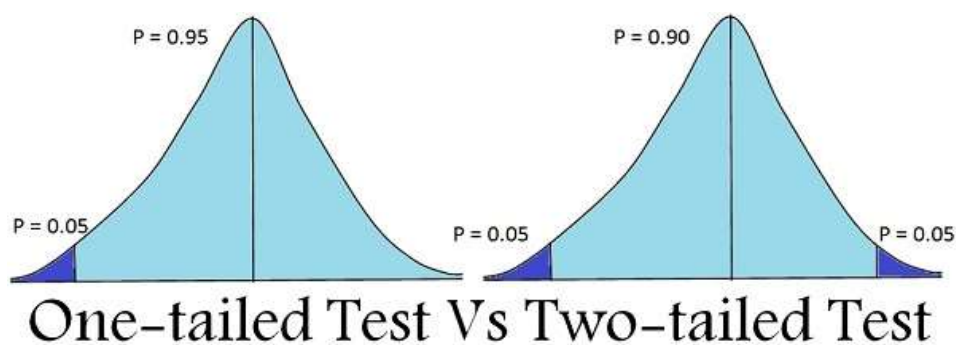
Đối với các nghiên cứu có kiểm định sự khác biệt, nếu cỡ mẫu lớn, dù sự khác biệt là nhỏ, ít ý nghĩa thực tế, vẫn có thể ý nghĩa thống kê (Ví dụ, huyết áp tâm thu được giảm đi 1 mmHg sau khi 1 tháng điều trị). Do vậy, mức khác biệt mong muốn hay còn gọi là mức khác biệt có ý nghĩa cần được xác định. Mức khác biệt mong muốn hay mức khác biệt có ý nghĩa thường do các nhà chuyên môn quyết định.

Cỡ tác động thường được tham khảo từ những nghiên cứu tương tự đã được làm trước đó. Tuy nhiên, khi thực hiện một nghiên cứu mới và không có thông tin tham khảo được từ nghiên cứu trước thì nhà nghiên cứu cũng có thể dựa vào kinh nghiệm lâm sàng, để lựa chọn ngưỡng thấp, trung bình hay cao vì cỡ tác động phản ánh sự khác biệt quan sát được trên thực tế.

10. Kiểm định 1 phía và 2 phía

Kiểm định 1 phía (*One-sided test*) được thực hiện khi chúng ta biết chắc chắn về xu hướng của sự khác biệt (chắc chắn lớn hơn hoặc chắc chắn nhỏ hơn). Kiểm định 2 phía (*Two-sided test*) được thực hiện khi chúng ta chỉ quan tâm có sự khác biệt hay không.

Với cùng 1 cỡ mẫu, kiểm định 1 phía có hiệu lực thống kê cao hơn. Tuy nhiên, kiểm định 2 phía được sử dụng nhiều hơn.



Hình 2: Kiểm định 1 phía và 2 phía

Tài liệu tham khảo

1. Julious SA (2009), Sample sizes for clinical trials, CRC Press, Boca Raton.
2. Rosner B (2011), Fundamentals of biostatistics, Brooks/Cole, Cengage Learning, Boston.
3. Perezgonzalez JD. Fisher, Neyman-Pearson or NHST? A tutorial for teaching data testing. *Frontiers in Psychology*. 2015;6:223.

Chương II: CHỌN MẪU NGHIÊN CỨU

GS.TS. Lưu Ngọc Hoạt, PGS.TS. Đào Thị Minh An, PGS.TS. Phạm Minh Khuê

1. Nguyên tắc chung

Một nghiên cứu sẽ có giá trị hơn nếu tất cả các cá thể trong một quần thể được bao hàm trong nghiên cứu. Tuy nhiên điều này thường rất khó thực hiện vì đa số các quần thể nghiên cứu thường quá lớn. Trên thực tế, một đặc trưng của quần thể có thể được ngoại suy với kết quả thu được từ một mẫu được rút ra từ quần thể này. Tuy nhiên sự ngoại suy chỉ được thực hiện nếu mẫu nghiên cứu được chọn đại diện cho quần thể và là đủ lớn.

Ba câu hỏi thường được đặt ra khi chọn mẫu là:

- Quần thể nào mà từ đó mẫu sẽ được lấy ra cho nghiên cứu?
- Làm thế nào để mẫu có thể đại diện cho quần thể nghiên cứu?
- Mẫu bao nhiêu là đủ cho một nghiên cứu?

Trả lời cho 3 câu hỏi này chính là giải quyết vấn đề xác định quần thể nghiên cứu (*study population*), chọn mẫu (*sampling*) và tính toán cỡ mẫu (*sample size*) cho một nghiên cứu.

Việc xác định quần thể nghiên cứu tùy thuộc vào nhiều vấn đề như ý tưởng của người nghiên cứu, vấn đề cần được nghiên cứu, các thông tin sẵn có cho việc chọn mẫu, kỹ thuật chọn mẫu, sự hiện diện của các nguồn lực phục vụ cho nghiên cứu, thời gian có thể dành cho nghiên cứu v.v... Quần thể nghiên cứu cần xác định rõ, bao hàm cả khái niệm thời gian và không gian để phục vụ cho việc chọn mẫu.

Một thiết kế mẫu được coi là tốt nếu như nó đáp ứng một số tiêu chuẩn sau:

- Đại diện cho quần thể nghiên cứu: Khi nó có tất cả các tính chất cơ bản của quần thể mà từ đó nó được rút ra.
- Mẫu là đủ lớn: Để có thể cho phép khái quát hoá một cách tin cậy cho quần thể nghiên cứu.

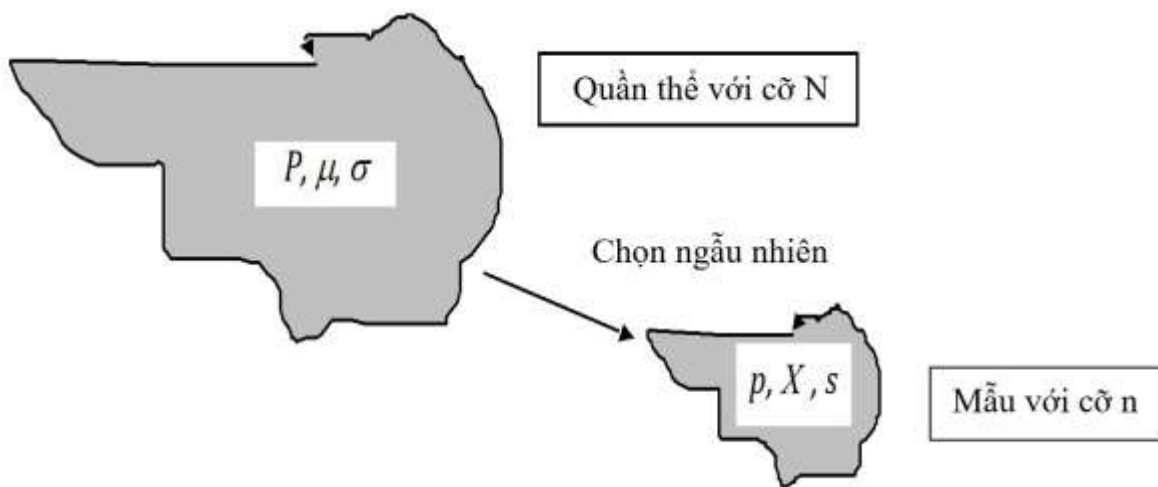
- Tính thực tế và tiện lợi: Để việc thu thập số liệu là dễ dàng và thuận tiện.
- Tính kinh tế và hiệu quả: Mẫu được chọn sao cho thông tin thu được là nhiều nhất trong khi chi phí là thấp nhất.

Trên thực tế có rất nhiều kỹ thuật chọn mẫu, tuy nhiên chúng có thể được xếp vào hai nhóm đó là chọn mẫu xác suất (*probability sampling*) và không xác suất (*non-probability sampling*). Mẫu được chọn theo phương pháp xác suất có tính đại diện cho quần thể hơn, nhưng thường tốn kém hơn khi thu thập do mẫu thường phân tán và cỡ mẫu lớn. Phần dưới đây trình bày các kỹ thuật chọn mẫu theo hai nhóm này.

2. Chọn mẫu xác suất

2.1. Chọn mẫu ngẫu nhiên đơn (*simple random sampling*)

Là kỹ thuật chọn mẫu trong đó tất cả các cá thể trong quần thể có cùng cơ hội (cùng xác suất) để được chọn vào mẫu.



Hình 3: Cách chọn mẫu ngẫu nhiên đơn

Ví dụ: Chọn 500 hồ sơ trong số 5.000 sản phụ đã đẻ tại bệnh viện A trong năm 2005 để nghiên cứu. Nếu theo cách chọn ngẫu nhiên đơn thì mỗi sản phụ có xác suất là 10% được chọn vào mẫu. Ta có thể tiến hành như sau:

- Lập một khung chọn mẫu chứa đựng tất cả các đơn vị mẫu.
- Sử dụng một quá trình ngẫu nhiên để chọn các cá thể vào mẫu. Có nhiều cách để chọn một mẫu ngẫu nhiên đơn từ quần thể như: Tung đồng xu,

tung súc sắc, bốc thăm... nhưng với các quần thể nghiên cứu lớn, người ta thường sử dụng bảng số ngẫu nhiên hoặc phần mềm máy tính để chọn.

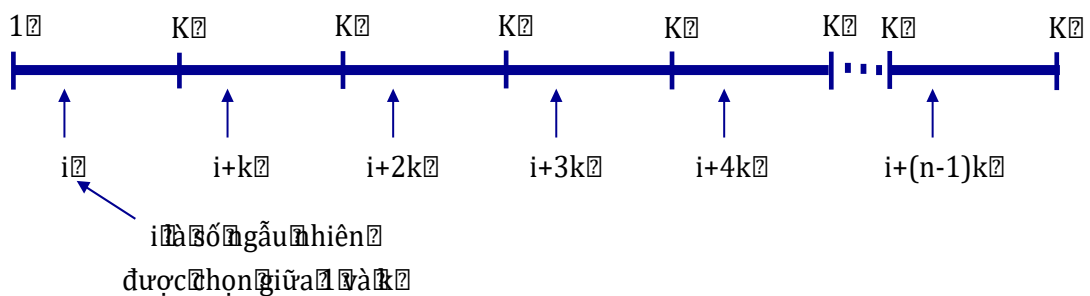
- Ưu điểm:
 - Cách làm đơn giản, tính ngẫu nhiên và tính đại diện cao.
 - Là kỹ thuật chọn mẫu xác suất cơ bản và có thể lồng vào tất cả các kỹ thuật chọn mẫu xác suất phức tạp khác.
- Nhược điểm:
 - Cần phải có một danh sách của các đơn vị mẫu để phục vụ cho chọn mẫu. Điều này thường không thể có được với một mẫu lớn hoặc mẫu dao động.
 - Các cá thể được chọn vào mẫu có thể phân bố tản mạn trong quần thể, do vậy, việc thu thập số liệu sẽ tốn kém và mất thời gian.
 - Do cách chọn này không quan tâm đến các tầng, các nhóm đối tượng nghiên cứu trong quần thể nên xác suất chọn vào mẫu có thể không đồng đều giữa các tầng dẫn đến kết quả chung có thể bị ảnh hưởng.

2.2. Chọn mẫu hệ thống (systematic sampling)

Trong chọn mẫu hệ thống, mỗi cá thể trong một danh sách được chọn bằng cách áp dụng một khoảng hằng định theo sau bởi một sự bắt đầu ngẫu nhiên.

- Các bước:
 - Tất cả các đơn vị mẫu (sampling unit) trong quần thể định nghiên cứu được ghi vào một danh sách hoặc trình bày trên bản đồ (gọi là khung chọn mẫu).
 - Xác định khoảng mẫu $k = N/n$ (N : cỡ của quần thể, n cỡ mẫu định chọn).
 - Một số (i) giữa 1 và k được chọn bằng cách chọn ngẫu nhiên đơn.
 - Các cá thể có số thứ tự $i + 1k$; $i + 2k$; $i + 3k$ v.v.... sẽ được chọn vào mẫu cho đến khi kết thúc danh sách hoặc bản đồ.
- Ưu điểm:
 - Nhanh và dễ áp dụng.
 - Nếu danh sách cá thể của quần thể được xếp ngẫu nhiên, chọn mẫu hệ thống tương tự như chọn ngẫu nhiên đơn.

- Nếu danh sách cá thể được xếp theo thứ tự tăng, đây là cách lựa chọn tương tự như mẫu tầng có tỷ lệ (proportionate stratified sample) tức là tầng có cỡ lớn hơn sẽ có nhiều cá thể được chọn vào mẫu hơn.
- Trong một số trường hợp, dù khung mẫu không có hoặc không biết tổng số cá thể trong quần thể nhưng chọn mẫu hệ thống vẫn có thể áp dụng được bằng cách xác định một quy luật phù hợp trước khi tiến hành chọn mẫu.



Hình 4: Chọn mẫu hệ thống với khoảng mẫu k và số bắt đầu i

Ví dụ: Để có thể thu thập được số liệu về sẹo lao ở trẻ em trong một cộng đồng nông thôn không biết danh sách các hộ gia đình, người nghiên cứu có thể xác định một quy luật chọn mẫu trước thu thập số liệu như sau:

- Hộ gia đình đầu tiên được điều tra là hộ thứ nhất nằm bên trái của ủy ban nhân dân xã.
- Các hộ tiếp theo sẽ được chọn bằng cách người nghiên cứu tiếp tục đi về bên trái và cứ cách 7 gia đình lại điều tra một gia đình (khoảng cách các hộ gia đình được chọn vào nghiên cứu được chọn bằng cách lấy số hộ gia đình trong cộng đồng chia cho số hộ gia đình dự kiến điều tra).
- Nếu gặp lối rẽ thì người nghiên cứu chỉ được rẽ trái.
- Tất cả các trẻ em trong các hộ gia đình được chọn đều được kiểm tra sẹo lao cho đến khi có đủ số trẻ cần được điều tra.
- Số hộ gia đình cần điều tra có thể được ước đoán từ số trẻ em trung bình trong một gia đình ở cộng đồng và số trẻ cần nghiên cứu (cỡ mẫu).

Trong một số trường hợp khác, các cá thể trong quần thể nghiên cứu có thể không cần lên danh sách để chọn, người nghiên cứu có thể đưa ra một quy luật trước khi chọn mẫu như:

- Tất cả các bệnh nhân đến phòng khám vào ngày thứ năm trong tuần sẽ được tham gia vào nghiên cứu.
- Bệnh nhân đến khám ngày lẻ sẽ vào nhóm 1, đến ngày chẵn sẽ vào nhóm hai để phục vụ cho một thử nghiệm nào đó.
- Các cá thể trong mẫu được chọn cho một nghiên cứu đứng vòng tròn, sau đó đếm lần lượt 1, 2, 3; 1, 2, 3 v.v... cho đến hết. Người đếm số 1 đầu tiên phải được chọn ngẫu nhiên. Các cá thể đếm số 1 sẽ vào nhóm 1, số 2 vào nhóm 2, số 3 vào nhóm 3. Như vậy, ta đã có ba nhóm được chọn ngẫu nhiên.

- **Nhược điểm:**

Khi việc sắp xếp khung mẫu có một quy luật nào đó tình cờ trùng với khoảng chọn mẫu hệ thống, các cá thể trong mẫu có thể thiếu tính đại diện.

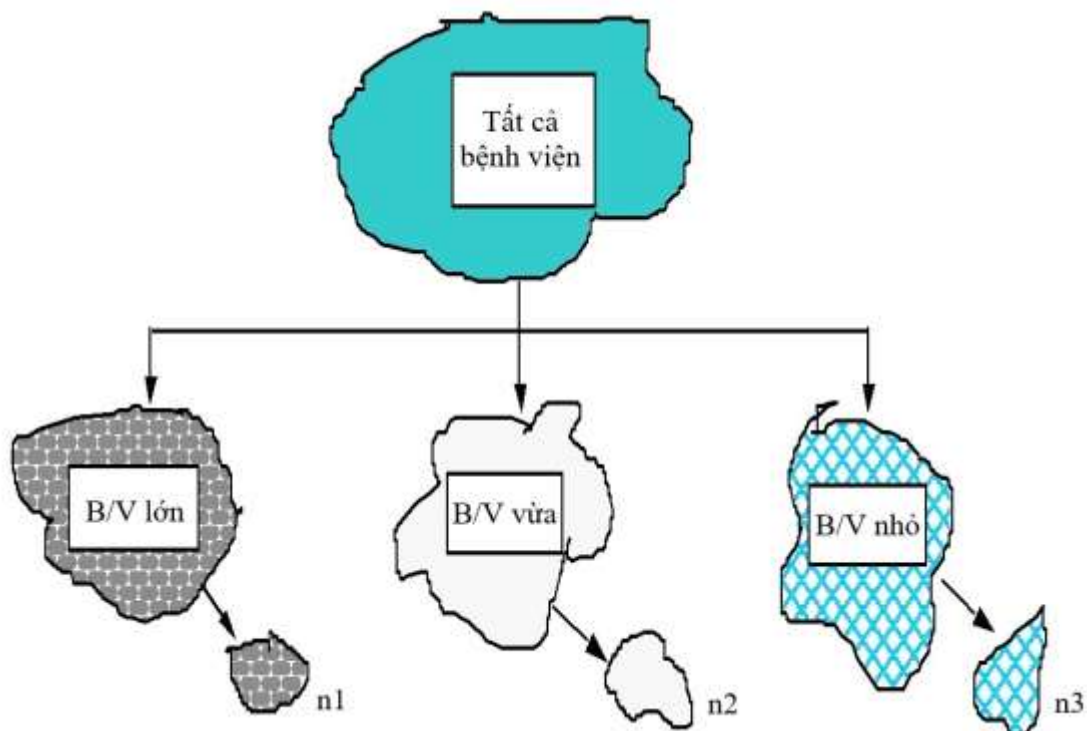
Ví dụ: Một cộng đồng được hình thành bởi nhiều dãy nhà, trong đó mỗi nhà đều có 10 gia đình. Nếu tình cờ hệ số k trong chọn mẫu hệ thống cũng bằng 10 và hộ gia đình đầu tiên được chọn vào mẫu là gia đình nằm ở đầu một dãy nhà thì tất cả các gia đình trong mẫu sẽ đều là các gia đình ở đầu các dãy nhà. Như vậy, mẫu sẽ thiếu tính đại diện về phân bố không gian.

2.3. Chọn mẫu ngẫu nhiên phân tầng (stratified random sampling)

Là kỹ thuật chọn mẫu được thực hiện bởi việc phân chia các cá thể của quần thể nghiên cứu thành các nhóm riêng rẽ được gọi là tầng. Đặc điểm của chọn mẫu kiểu này là tiêu chí nghiên cứu trong từng tầng tương đối đồng nhất còn giữa các tầng có sự khác biệt. Sau khi đã phân tầng xong ta vẫn có thể áp dụng cách chọn mẫu ngẫu nhiên đơn hoặc ngẫu nhiên hệ thống để chọn đối tượng của từng tầng vào nghiên cứu:

- **Các bước:**

- Phân chia quần thể nghiên cứu thành các tầng khác nhau dựa vào một hoặc vài đặc điểm nào đó như nhóm tuổi, giới, tầng lớp xã hội, dân tộc v.v... Giữa các tầng không có sự chồng chéo.
- Thực hiện việc chọn mẫu ngẫu nhiên đơn hoặc ngẫu nhiên hệ thống trong từng tầng.
- Các phân tích thống kê (như giá trị trung bình, độ lệch) được tính toán riêng cho mỗi tầng sau đó sẽ kết hợp lại trên cơ sở kích cỡ của từng tầng (cân bằng trọng: weighted) để cho kết quả của toàn bộ quần thể.



Hình 5: Chọn mẫu phân tầng trong một điều tra bệnh viện với tiêu thức phân tầng là cỡ của bệnh viện

- Ưu điểm:
 - Tạo ra trong mỗi tầng có một sự đồng nhất về yếu tố được chọn để nghiên cứu, do đó sẽ giảm sự chênh lệch giữa các cá thể.
 - Khi nguyên tắc mẫu tỷ lệ được sử dụng, tầng có kích cỡ lớn hơn sẽ có nhiều cá thể được chọn vào mẫu hơn. Khi đó cỡ mẫu cho một tầng i nào đó sẽ là:

$$n_i = n \frac{N_i}{N}$$

Trong đó:

n_i : Cỡ mẫu của tầng i .

N_i : Dân số tầng i .

n : Cỡ mẫu của tất cả các tầng.

N : Dân số của quần thể.

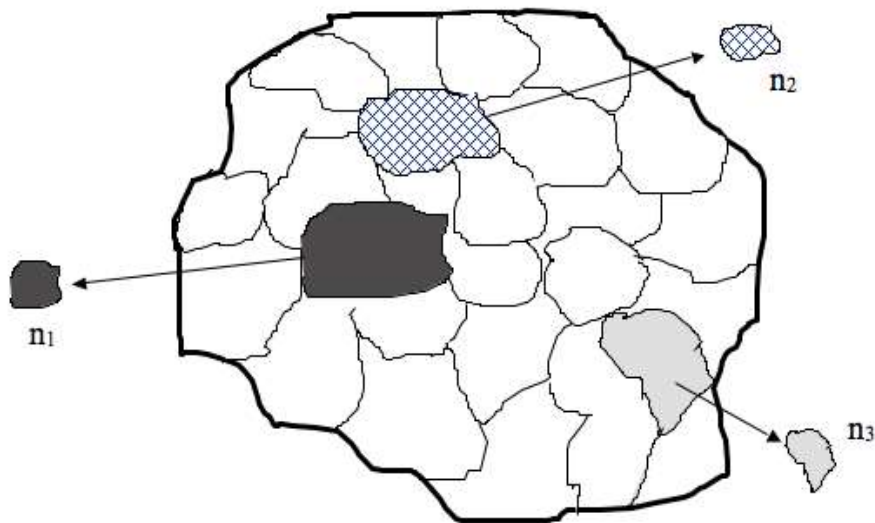
- Nguyên tắc mẫu không tỷ lệ cũng có thể được áp dụng trong mẫu tầng. Khi đó tỷ lệ mẫu trong các tầng sẽ khác nhau. Với những tầng có biến thiên lớn giữa các cá thể hoặc chi phí cho chọn mẫu thấp, người ta thường áp dụng tỷ lệ mẫu lớn. Cách này giúp cho người điều tra có đủ số cá thể trong mỗi tầng để có thể phân tích được.
- Mẫu chọn từ mỗi tầng có tính đại diện và khái quát hoá cao cho tầng đó.
- Nếu yếu tố được chọn để phân tầng có tính đồng nhất cao trong mỗi tầng nhưng lại thấp giữa các tầng thì kết quả nghiên cứu sẽ có độ chính xác cao hơn là mẫu chọn theo cách ngẫu nhiên đơn (2 cách phải có cùng cỡ mẫu).
- Ngoài việc tính được tham số mẫu (tỷ lệ, giá trị trung bình, tỷ suất chênh v.v...) cho toàn bộ mẫu nghiên cứu, với các chọn mẫu phân tầng ta có thể tính riêng được tham số mẫu cho từng tầng.
- Nhược điểm:
Cũng như chọn mẫu ngẫu nhiên đơn, danh sách tất cả các cá thể trong mỗi tầng phải được liệt kê và được gán số ngẫu nhiên. Điều này thường khó thực hiện trong thực tế.

2.4. Chọn mẫu chùm (cluster sampling)

Là kỹ thuật chọn mẫu trong đó việc lựa chọn ngẫu nhiên các nhóm cá thể (được gọi là chùm) từ nhiều chùm trong một quần thể nghiên cứu. Trong trường hợp này đơn vị mẫu là các chùm chứ không phải là các cá thể.

- Các bước:
 - Xác định các chùm thích hợp: Việc này thường được làm bởi người điều tra. Chùm được làm bởi tập hợp các cá thể gần nhau (làng, xã, trường học, khoa phòng, bệnh viện v.v...) do đó thường có chung một số đặc điểm. Các chùm thường không có cùng kích cỡ.

- Lập danh sách tất cả các chùm, chọn ngẫu nhiên một số chùm vào mẫu. Từ đây sẽ có hai cách chọn tiếp tùy theo ý tưởng của người nghiên cứu:
- + *Cách 1*: Tất cả các cá thể trong các chùm đã chọn sẽ được bao gồm vào nghiên cứu. Trong cách này đơn vị mẫu (sampling unit) chính là các chùm được chọn, trong khi yếu tố quan sát (observation element) lại là các cá thể trong chùm (ví dụ như các hộ gia đình trong thôn được chọn, trẻ em trong các hộ gia đình được chọn v.v...). Cách này được gọi là mẫu chùm một bậc và xác suất của một chùm được chọn vào mẫu bằng số chùm dự kiến chọn chia cho tổng số các chùm.
- + *Cách 2*: Liệt kê danh sách các cá thể trong các chùm đã chọn, sau đó áp dụng cách chọn mẫu ngẫu nhiên đơn hoặc ngẫu nhiên hệ thống trong mỗi chùm để chọn các cá thể vào mẫu. Trong trường hợp này đơn vị mẫu và đơn vị quan sát là trùng nhau (mẫu 2 bậc).



Hình 6: Mẫu hai giai đoạn, trong đó giai đoạn 1 là chọn mẫu chùm, giai đoạn 2 là chọn ngẫu nhiên đơn hoặc ngẫu nhiên hệ thống

- Ưu điểm:
 - Nó thường được áp dụng trong các nghiên cứu điều tra trong một phạm vi rộng lớn, độ phân tán cao, danh sách của tất cả các cá thể trong quần thể

không thể có được (do khó hoặc đắt), trong khi chỉ có danh sách hoặc bản đồ các chòm.

- Sự lựa chọn thường dễ hơn, chi phí cho nghiên cứu với mẫu chòm thường rẻ hơn nhiều do các cá thể trong một chòm thường gần nhau.
- Nhược điểm:
 - Tính đại diện cho quần thể hoặc tính chính xác (precision) của mẫu được chọn theo phương pháp mẫu chòm thường thấp hơn so với mẫu được chọn bằng phương pháp ngẫu nhiên đơn (nếu có cùng cỡ mẫu). Vì vậy, để tăng tính chính xác này người ta thường tăng cỡ mẫu bằng cách nhân cỡ mẫu (được tính theo công thức cho các chọn mẫu ngẫu nhiên đơn) với hệ số ảnh hưởng của thiết kế (design effect). Hệ số ảnh hưởng của thiết kế được tính theo công thức khá phức tạp nên người ta khuyên rằng nếu không tính chính xác được hệ số này thì ta có thể coi nó bằng 2.
 - Có một sự tương quan nghịch giữa cỡ của chòm và tính đại diện của mẫu, do vậy, cỡ chòm càng nhỏ càng tốt, tuy nhiên chi phí cho điều tra sẽ cao hơn. Số chòm được chọn vào nghiên cứu tốt nhất là phải > 30 .
 - Phân tích số liệu từ mẫu chòm thường phức tạp hơn các mẫu khác.

Việc lựa chọn chòm vào mẫu nghiên cứu cũng khá phức tạp, đặc biệt là khi cỡ chòm không đều nhau. Trong trường hợp này, người ta có thể áp dụng phương pháp chọn chòm theo phương pháp PPS (sẽ được trình bày trong phần sau).

2.5. Chọn mẫu nhiều giai đoạn (multistage sampling)

Trên thực tế, nhiều nghiên cứu với các quần thể lớn có cấu trúc phức tạp cần phải áp dụng nhiều kỹ thuật chọn mẫu trong các giai đoạn khác nhau. Có thể kết hợp cả mẫu xác suất và không xác suất. Với các thiết kế nghiên cứu trên diện rộng mà tiêu thức nghiên cứu biến đổi nhiều giữa các vùng sinh thái, các tầng lớp xã hội hoặc các loại nghề nghiệp khác nhau, thông thường người ta thực hiện bước thiết kế phân tầng trước, sau đó trong mỗi tầng sẽ thực hiện thiết kế chọn chòm (vì khi đó tiêu thức nghiên cứu giữa các chòm trong mỗi tầng là tương đối đồng đều). Tiếp đó trong mỗi chòm lại có thể chọn ngẫu nhiên đơn, ngẫu nhiên

hệ thống hoặc thậm chí có thể lại chọn mẫu chùm nếu cỡ chùm còn quá lớn (như tỉnh, huyện...) thì để hạn chế mức độ phân tán của mẫu, giảm thời gian và kinh phí cho việc thu thập số liệu.

Một điểm cần lưu ý là trong các nghiên cứu khi phải áp dụng nhiều lần chọn mẫu chùm thì cứ mỗi lần chọn mẫu chùm, cỡ mẫu lại phải điều chỉnh với hệ số thiết kế để đảm bảo tính đại diện của mẫu nghiên cứu so với quần thể.

3. Chọn mẫu không xác suất

3.1. Chọn mẫu thuận tiện (convenience or accidental sampling)

Đạt được trên cơ sở các cá thể có sẵn khi thu thập số liệu. (Ví dụ như tất cả các bệnh nhân đến khám tại phòng khám trong ngày). Phương pháp này không quan tâm đến việc sự lựa chọn có ngẫu nhiên hay không. Đây là cách chọn mẫu rất hay gặp trong các nghiên cứu lâm sàng.

3.2. Chọn mẫu chỉ tiêu (quota sampling)

Là phương pháp đảm bảo rằng một số nhất định các đơn vị mẫu từ các loại khác nhau của quần thể nghiên cứu với các tính chất đặc trưng sẽ có mặt trong mẫu. Nó gần giống như cách chọn mẫu tầng nhưng không ngẫu nhiên. Người nghiên cứu đặt kế hoạch là sẽ chọn bao nhiêu đối tượng cho mỗi tầng hoặc nhóm đối tượng và bằng cách chọn mẫu thuận tiện cho đến khi đủ số lượng từ mỗi tầng.

3.3. Chọn mẫu có mục đích (purposive sampling)

Người nghiên cứu đã xác định trước các nhóm quan trọng trong quần thể để tiến hành thu thập số liệu. Các nhóm khác nhau sẽ có tỷ lệ mẫu khác nhau. Đây là cách rất hay dùng trong các điều tra thăm dò, phỏng vấn sâu.

3.4. Các ứng dụng của các kỹ thuật chọn mẫu không xác suất

Các cách chọn mẫu không xác suất thường dễ làm, rẻ nhưng do lựa chọn không ngẫu nhiên nên tính đại diện cho quần thể nghiên cứu rất thấp. Nếu như mục đích của nghiên cứu là để đo lường các biến số và từ đó khái quát hoá cho một quần thể thì các kết quả thu được từ mẫu không xác suất thường không đủ cơ sở khoa học cho việc ngoại suy. Do đó, cần phải dè dặt khi đưa ra các kết luận.

Tuy nhiên với một số loại nghiên cứu được thiết kế với mục đích thăm dò hoặc muốn tìm hiểu sâu một vấn đề nào đó của quần thể (như kiến thức, động cơ, thái độ, niềm tin v.v...) thì khi đó việc chọn mẫu xác suất là không cần thiết và có thể áp dụng cách chọn mẫu không xác suất.

Ngoài ra, trong một số thử nghiệm lâm sàng, mẫu thường là những người tình nguyện, khi đó cách chọn mẫu xác suất không thể áp dụng được.

3.5. Sự kết hợp giữa chọn mẫu xác suất và không xác suất

Trong thực tế, có nhiều mẫu được chọn theo nhiều giai đoạn, khi đó cả hai cách chọn mẫu nêu trên có thể được kết hợp trong cùng một thiết kế nghiên cứu. Người điều tra có thể thực hiện một chọn mẫu xác suất trên cơ sở một mẫu không xác suất đã được chọn ở giai đoạn trước. Việc quyết định chọn mẫu xác suất hay không tùy thuộc vào tính chất của nghiên cứu, thời gian, kinh phí và nhiều vấn đề khác nhưng người làm nghiên cứu phải biết được ưu điểm và hạn chế trong các cách chọn mẫu đã sử dụng.

4. Chọn mẫu trong các thử nghiệm lâm sàng hoặc thực nghiệm

Thường được áp dụng khi tác giả muốn so sánh tác dụng của các phương pháp điều trị khác nhau, hiệu quả của các thuốc mới v.v... Trong trường hợp này cần phải có ít nhất hai nhóm nghiên cứu để so sánh. Khi nghiên cứu này được áp dụng trên người, sẽ có rất nhiều vấn đề về chọn mẫu và đạo đức trong nghiên cứu. Người nghiên cứu phải đảm bảo về tính an toàn của can thiệp. Mục đích và lợi ích của nghiên cứu cần phải được giải thích rõ ràng cho các đối tượng được chọn vào nghiên cứu và việc tham gia vào nghiên cứu là hoàn toàn tự nguyện. Do đặc tính này nên mẫu được chọn trong các nghiên cứu loại này thường thiếu tính đại diện cho quần thể chung, tuy nhiên nó có thể phần nào đại diện cho nhóm cá thể có cùng mọi tính chất như các cá thể được chọn vào thử nghiệm. Một điều cần lưu ý là mục đích ngoại suy ở đây thường không quan trọng bằng mục đích thử nghiệm, do vậy, điều quan trọng hơn là phân bố đối tượng vào các nhóm nghiên cứu sao cho ngẫu nhiên (randomization).

5. Đạo đức trong việc chọn mẫu nghiên cứu

Vấn đề đạo đức trong nghiên cứu cần phải được cân nhắc kỹ nhất là trong các thử nghiệm lâm sàng hoặc nghiên cứu can thiệp. Nếu can thiệp là tốt thì nhóm chúng sẽ mất cơ hội để được hưởng sự can thiệp, còn nếu can thiệp có tác dụng phụ nguy hiểm thì nhóm thử nghiệm có nguy cơ cao hơn.

Ngoài ra trong một số loại chọn mẫu điều tra cộng đồng, nếu áp dụng chọn mẫu ngẫu nhiên đơn, ngẫu nhiên hệ thống với các cộng đồng có mật độ dân số cao như các tổ, lớp, trường học thì việc đối tượng trong cùng tổ, lớp có người được chọn, người không được chọn cũng có thể phát sinh các thắc mắc. Trong trường hợp này ta có thể chọn chòm toàn bộ để nghiên cứu, tức là tổ lớp nào được chọn thì tất cả các cá thể trong đó đều được điều tra, nghiên cứu.

Chương III: TÍNH TOÁN CỠ MẪU NGHIÊN CỨU

GS.TS. Hoàng Văn Minh, BS. Khương Quỳnh Long, BS. Ong Phúc Thịnh,
PGS.TS. Võ Văn Thắng, PGS.TS. Kim Bảo Giang, PGS.TS. Nguyễn Thùy Dương

1. Tính toán cỡ mẫu cho nghiên cứu sử dụng thống kê suy luận dựa trên ước lượng khoảng

Tính toán cỡ mẫu nghiên cứu tùy thuộc vào phương pháp ngoại suy thống kê (thống kê suy luận) được áp dụng là ước lượng khoảng (*Confidence interval*) hay kiểm định giả thuyết (*Hypothesis testing*).

Khi nghiên cứu sử dụng thống kê suy luận dựa trên ước lượng khoảng (*Confidence interval*), chúng ta cần tính cỡ mẫu để đạt được độ chính xác (hay mức tin cậy) đã được lựa chọn.

Khoảng tin cậy 95% của biến định lượng được tính theo công thức:

$$95\% \text{ CI} = \text{Ước lượng điểm} \pm d$$

Trong đó, d là sai số biên (*Margin of error*) được tính theo công thức:

- Đối với biến định lượng:

$$d = Z_{1-\frac{\alpha}{2}} * \frac{\delta}{\sqrt{n}}$$

Trong đó:

- Z là giá trị từ phân bố chuẩn, được tính dựa trên mức ý nghĩa thống kê
- δ là độ lệch chuẩn
- n là cỡ mẫu

- Đối với biến định tính:

$$d = Z_{1-\frac{\alpha}{2}} * \sqrt{\frac{p(1-p)}{n}}$$

Trong đó:

- Z là giá trị từ phân bố chuẩn, được tính dựa trên mức ý nghĩa thống kê
- p là tỷ lệ
- n là cỡ mẫu

Như vậy, công thức tính cỡ mẫu sẽ là:

- *Đối với biến định lượng:*

$$n = \left[\frac{Z_{1-\frac{\alpha}{2}} * \sigma}{d} \right]^2 = \frac{Z_{1-\frac{\alpha}{2}}^2 * \sigma^2}{d^2}$$

Trong đó:

- n là cỡ mẫu
 - Z là giá trị được tính dựa trên mức ý nghĩa thống kê (Z = 1,96 nếu mức ý nghĩa thống kê = 5%)
 - σ là độ lệch chuẩn (Lấy từ nghiên cứu trước đây hoặc từ nghiên cứu thử)
 - d là mức sai số chấp nhận (Do nhà nghiên cứu quyết định tùy vào ý nghĩa thực tiễn của kết quả nghiên cứu và nguồn lực dành cho nghiên cứu). Sai số d có thể là sai số tuyệt đối hoặc sai số tương đối ($\epsilon * p$).
- *Đối với biến định tính:*

$$n = \left[\frac{Z_{1-\frac{\alpha}{2}} * p(1-p)}{d} \right]^2 = \frac{Z_{1-\frac{\alpha}{2}}^2 * p(1-p)^2}{d^2}$$

Trong đó:

- n là cỡ mẫu
- Z là giá trị được tính dựa trên mức ý nghĩa thống kê
- P là tỷ lệ (Lấy từ nghiên cứu trước đây hoặc từ nghiên cứu thử)
- d là mức sai số chấp nhận (Do nhà nghiên cứu quyết định tùy vào ý nghĩa thực tiễn của kết quả nghiên cứu và nguồn lực dành cho nghiên cứu). Sai số d có thể là sai số tuyệt đối hoặc sai số tương đối ($\epsilon * p$).

1.1. Nghiên cứu gồm 1 mẫu, xác định một trung bình

- ***Nghiên cứu gồm 1 mẫu, xác định một trung bình, sử dụng sai số tuyệt đối***

* Công thức:

$$n = \frac{Z_{1-\frac{\alpha}{2}}^2 \sigma^2}{d^2}$$

Trong đó:

- n là cỡ mẫu tối thiểu
- $Z_{1-\frac{\alpha}{2}}$ là giá trị từ phân bố chuẩn, được tính dựa trên mức ý nghĩa thống kê ($Z_{1-\frac{\alpha}{2}} = 1,96$ nếu mức ý nghĩa thống kê = 5%)
- σ là độ lệch chuẩn (Lấy từ nghiên cứu trước đây hoặc từ nghiên cứu thử)
- d là mức sai số tuyệt đối chấp nhận (Do nhà nghiên cứu quyết định tùy vào ý nghĩa thực tiễn của kết quả nghiên cứu và nguồn lực dành cho nghiên cứu).

* Ví dụ:

Tính toán cỡ mẫu cho một nghiên cứu ước tính trị số huyết áp tâm thu ở trẻ em. Biết rằng, nghiên cứu trước đây báo cáo độ lệch chuẩn của trị số huyết áp tâm thu là 20mmHg. Lấy mức ý nghĩa thống kê = 5% và sai số tuyệt đối là 5mmHg. Cỡ mẫu được tính toán theo công thức như sau:

$$n = \frac{Z_{1-\frac{\alpha}{2}}^2 \cdot \sigma^2}{d^2} = \frac{1,96^2 \cdot 20^2}{5^2} = 62$$

**Tính toán theo phần mềm HSS 1.0:*

Nghiên cứu gồm 1 mẫu, xác định 1 trung bình

Tính cỡ mẫu

Alpha (α)	Tỷ lệ không trả lời	
0.05	0	
$Z_{1-\alpha/2} = 1.96$	Hệ số thiết kế	
Độ lệch chuẩn (σ)	1	
20		
Chọn loại sai số	Cỡ mẫu	
<input checked="" type="radio"/> Sai số tuyệt đối		62
<input type="radio"/> Sai số tương đối		
Sai số tuyệt đối (d)		
5		

Ghi chú: Xem thêm phần hiệu chỉnh cỡ mẫu theo tỷ lệ không trả lời và hệ số thiết kế để tính toán cỡ mẫu cuối cùng

- **Nghiên cứu gồm 1 mẫu, xác định một trung bình, sử dụng sai số tương đối**

* Công thức:

$$n = \frac{Z_{1-\frac{\alpha}{2}}^2 \sigma^2}{\varepsilon^2 \mu^2}$$

Trong đó:

- n là cỡ mẫu tối thiểu
- $Z_{1-\frac{\alpha}{2}}$ là giá trị từ phân bố chuẩn, được tính dựa trên mức ý nghĩa thống kê ($Z_{1-\frac{\alpha}{2}} = 1,96$ nếu mức ý nghĩa thống kê = 5%)
- σ là độ lệch chuẩn (Lấy từ nghiên cứu trước đây hoặc từ nghiên cứu thử)
- ε là mức sai số tương đối chấp nhận (Do nhà nghiên cứu quyết định tùy vào ý nghĩa thực tiễn của kết quả nghiên cứu và nguồn lực dành cho nghiên cứu)
- μ là giá trị trung bình của quần thể (Lấy từ nghiên cứu trước đây hoặc từ nghiên cứu thử)

* Ví dụ:

Tính toán cỡ mẫu cho một nghiên cứu ước tính trị số huyết áp tâm thu ở trẻ em. Biết rằng, nghiên cứu trước đây báo cáo giá trị trung bình và độ lệch chuẩn của trị số huyết áp tâm thu lần lượt là 100mg và 20mmHg. Lấy mức ý nghĩa thống kê = 5% và sai số tương đối là 5%.

Cỡ mẫu được tính toán theo công thức như sau:

$$n = \frac{Z_{1-\frac{\alpha}{2}}^2 \sigma^2}{\varepsilon^2 \mu^2} = \frac{1,96^2 * 20^2}{0,05^2 * 100^2} = 62$$

**Tính toán theo phần mềm HSS 1.0:*

Nghiên cứu gồm 1 mẫu, xác định 1 trung bình

Tính cỡ mẫu

Alpha (α)	Tỷ lệ không trả lời
0.05	0
$Z_{1-\frac{\alpha}{2}} = 1.96$	Hệ số thiết kế
Độ lệch chuẩn (σ)	1
20	
Chọn loại sai số	
<input type="radio"/> Sai số tuyệt đối	
<input checked="" type="radio"/> Sai số tương đối	
Sai số tương đối (ϵ)	
0.05	
Trung bình quần thể (μ)	
100	

Cỡ mẫu
62

Ghi chú: Xem thêm phần hiệu chỉnh cỡ mẫu theo tỷ lệ không trả lời và hệ số thiết kế để tính toán cỡ mẫu cuối cùng

1.2. Nghiên cứu gồm 1 mẫu, xác định một tỷ lệ

- **Nghiên cứu gồm 1 mẫu, xác định một tỷ lệ, sử dụng sai số tuyệt đối**

* Công thức:

$$n = \frac{Z_{1-\frac{\alpha}{2}}^2 P(1 - P)}{d^2}$$

Trong đó:

- n là cỡ mẫu tối thiểu
- $Z_{1-\frac{\alpha}{2}}$ là giá trị từ phân bố chuẩn, được tính dựa trên mức ý nghĩa thống kê ($Z_{1-\frac{\alpha}{2}} = 1,96$ nếu mức ý nghĩa thống kê = 5%)
- p là tỷ lệ ước đoán (Lấy từ nghiên cứu trước đây hoặc từ nghiên cứu thử)
- d là mức sai số tuyệt đối chấp nhận (Do nhà nghiên cứu quyết định tùy vào ý nghĩa thực tiễn của kết quả nghiên cứu và nguồn lực dành cho nghiên cứu). Hiện nay các nhà thống kê đưa ra gợi ý tính d dựa trên giá trị của p như sau:

- $p < 0,1 \rightarrow d = p/2$;
- $p = 0,1-0,3 \rightarrow d = 0,05$;
- $p = 0,3-0,7 \rightarrow d = 0,1$;
- $p = 0,7-0,9 \rightarrow d = 0,05$;
- $p > 0,9 \rightarrow d = (1-p)/2$

**Ví dụ*

Tính toán cỡ mẫu cho một nghiên cứu ước tính tỷ lệ hút thuốc lá ở nam giới. Biết rằng, nghiên cứu trước đây báo cáo tỷ lệ hút thuốc lá ở nam giới là 50%. Lấy mức ý nghĩa thống kê = 5% và sai số tuyệt đối là 5%. Cỡ mẫu được tính toán theo công thức như sau:

$$n = \frac{Z_{1-\frac{\alpha}{2}}^2 P(1 - P)}{d^2} = \frac{1,96^2 \cdot 0,5(1 - 0,5)}{0,05^2} = 385$$

**Tính toán theo phần mềm HSS 1.0:*

Ghi chú: Xem thêm phần hiệu chỉnh cỡ mẫu theo tỷ lệ không trả lời và hệ số thiết kế để tính toán cỡ mẫu cuối cùng

- ***Nghiên cứu gồm 1 mẫu, xác định một tỷ lệ, sử dụng sai số tương đối***

** Công thức:*

$$n = \frac{Z_{1-\frac{\alpha}{2}}^2 (1 - P)}{\varepsilon^2 P}$$

Trong đó:

- n là cỡ mẫu tối thiểu
- $Z_{1-\frac{\alpha}{2}}$ là giá trị từ phân bố chuẩn, được tính dựa trên mức ý nghĩa thống kê ($Z_{1-\frac{\alpha}{2}} = 1,96$ nếu mức ý nghĩa thống kê = 5%)
- p là tỷ lệ ước đoán (Lấy từ nghiên cứu trước đây hoặc từ nghiên cứu thử)
- ε là mức sai số tương đối chấp nhận (Do nhà nghiên cứu quyết định tùy vào ý nghĩa thực tiễn của kết quả nghiên cứu và nguồn lực dành cho nghiên cứu).

**Ví dụ*

Tính toán cỡ mẫu cho một nghiên cứu ước tính tỷ lệ hút thuốc lá ở nam giới. Biết rằng, nghiên cứu trước đây báo cáo tỷ lệ hút thuốc lá ở nam giới là 50%. Lấy mức ý nghĩa thống kê = 5% và sai số tương đối là 10%.

Cỡ mẫu được tính toán theo công thức như sau:

$$n = \frac{Z_{1-\frac{\alpha}{2}}^2 P(1-P)}{\varepsilon^2 P} = \frac{1,96^2 * 0,5(1-0,5)}{0,1^2 * 0,5} = 385$$

**Tính toán theo phần mềm HSS 1.0:*

Nghiên cứu gồm 1 mẫu, xác định 1 tỷ lệ

Tính cỡ mẫu

Alpha (α) 0.05	Tỷ lệ không trả lời 0
$Z_{1-\alpha/2} = 1.96$	Hệ số thiết kế 1
Tỷ lệ ước lượng (P) 0.5	
Chọn loại sai số <input type="radio"/> Sai số tuyệt đối <input checked="" type="radio"/> Sai số tương đối	
Sai số tương đối (ϵ) 0.1	

Cỡ mẫu
385

Ghi chú: Xem thêm phần hiệu chỉnh cỡ mẫu theo tỷ lệ không trả lời và hệ số thiết kế để tính toán cỡ mẫu cuối cùng

1.3. Nghiên cứu gồm 2 mẫu độc lập, xác định sự khác biệt giữa 2 số trung bình

- Công thức:

$$n = 2 \left(\frac{Z_{1-\frac{\alpha}{2}}^2 \sigma^2}{d^2} \right)$$

Trong đó:

- n là cỡ mẫu tối thiểu cho mỗi nhóm
- Z là giá trị từ phân bố chuẩn, được tính dựa trên mức ý nghĩa thống kê ($Z = 1,96$ nếu mức ý nghĩa thống kê = 5%)
- σ là độ lệch chuẩn chung của 2 nhóm, tính bằng công thức

$$\sigma = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

- d là mức sai số chấp nhận (Do nhà nghiên cứu quyết định tùy vào ý nghĩa thực tiễn của kết quả nghiên cứu và nguồn lực dành cho nghiên cứu)

- Ví dụ

Tính toán cỡ mẫu cho một nghiên cứu so sánh mức tăng cân nặng ở 2 nhóm trẻ. Biết rằng, nghiên cứu trước đây tiến hành trên 2 nhóm, mỗi nhóm 100 trẻ, phát hiện được độ lệch chuẩn lần lượt là 8,4g và 7,7g. Lấy mức ý nghĩa thống kê = 5% cho kiểm định 2 phía và sai số là 3g. Cỡ mẫu được tính toán theo công thức như sau:

$$\sigma = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{(100 - 1)8,4^2 + (100 - 1)7,7^2}{100 + 100 - 2}} = 8,1$$

$$n = \frac{Z_{1-\frac{\alpha}{2}}^2 \sigma^2}{d^2} = 2\left(\frac{1,96 * 8,1}{3}\right)^2 = 56$$

* Tính toán theo phần mềm HSS 1.0

Nghiên cứu 2 mẫu độc lập, xác định sự khác biệt 2 trung bình

Tính cỡ mẫu

Alpha (α)
0.05

Tỷ lệ không trả lời
0

$Z_{1-\alpha/2} = 1.96$

Sai số (d)
3

Hệ số thiết kế
1

Chọn cách nhập
 Độ lệch chuẩn gộp
 Độ lệch chuẩn của 2 nhóm

Cỡ mẫu nhóm 1 (n_1)
100

Độ lệch chuẩn nhóm 1 (s_1)
8.4

Cỡ mẫu nhóm 2 (n_2)
100

Độ lệch chuẩn nhóm 2 (s_2)
7.7

Độ lệch chuẩn gộp (σ) tính được là:
8.06

Cỡ mẫu
56

Ghi chú: Xem thêm phần hiệu chỉnh cỡ mẫu theo tỷ lệ không trả lời và hệ số thiết kế để tính toán cỡ mẫu cuối cùng

1.4. Nghiên cứu gồm 2 mẫu ghép cặp, xác định sự khác biệt giữa 2 số trung bình

- Công thức:

$$n = \frac{Z_{1-\frac{\alpha}{2}}^2 \sigma_d^2}{d^2}$$

Trong đó:

- n là cỡ mẫu tối thiểu
 - Z là giá trị từ phân bố chuẩn, được tính dựa trên mức ý nghĩa thống kê (Z = 1,96 nếu mức ý nghĩa thống kê = 5% và kiểm định 2 phía, Z = 1,65 nếu mức ý nghĩa thống kê = 5% và kiểm định 1 phía)
 - σ_d là độ lệch chuẩn của sự khác biệt về số trung bình giữa 2 nhóm
 - d là mức sai số chấp nhận (Do nhà nghiên cứu quyết định tùy vào ý nghĩa thực tiễn của kết quả nghiên cứu và nguồn lực dành cho nghiên cứu)
- Ví dụ

Tính toán cỡ mẫu cho một nghiên cứu theo dõi dọc mức tăng cân nặng ở 1 nhóm trẻ qua thời gian. Biết rằng, nghiên cứu trước đây tiến hành trên 100 trẻ, phát hiện được độ lệch chuẩn của mức tăng cân nặng (sự khác biệt về cân nặng giữa 2 lần đo) là 9,1g. Lấy mức ý nghĩa thống kê = 5% cho kiểm định 2 phía và sai số là 3g.

Cỡ mẫu được tính toán theo công thức như sau:

$$n = \frac{Z_{1-\frac{\alpha}{2}}^2 \sigma_d^2}{d^2} = \left(\frac{1,96 * 9,1}{3} \right)^2 = 36$$

* Tính toán theo phần mềm HSS 1.0

Tính cỡ mẫu

Alpha (α) 0,05	Tỷ lệ không trả lời 0
$Z_{1-\frac{\alpha}{2}} = 1,96$ Độ lệch chuẩn của sự khác biệt trước-sau (σ_d) 9,1	Hệ số thiết kế 1
Sai số (d) 3	Cỡ mẫu 36

Ghi chú: Xem thêm phần hiệu chỉnh cỡ mẫu theo tỷ lệ không trả lời và hệ số thiết kế để tính toán cỡ mẫu cuối cùng

1.5. Nghiên cứu gồm 2 mẫu độc lập, xác định sự khác biệt giữa 2 tỷ lệ

- Công thức

$$n = \frac{Z_{1-\frac{\alpha}{2}}^2 [P_1(1 - P_1) + P_2(1 - P_2)]}{d^2}$$

Trong đó:

- n là cỡ mẫu tối thiểu cho mỗi nhóm
- Z là giá trị từ phân bố chuẩn, được tính dựa trên mức ý nghĩa thống kê (Z = 1,96 nếu mức ý nghĩa thống kê = 5% và kiểm định 2 phía, Z = 1,65 nếu mức ý nghĩa thống kê = 5% và kiểm định 1 phía)
- p1, p2 là tỷ lệ ước đoán trong từng nhóm
- d là mức sai số chấp nhận (Do nhà nghiên cứu quyết định tùy vào ý nghĩa thực tiễn của kết quả nghiên cứu và nguồn lực dành cho nghiên cứu)

- Ví dụ

Tính toán cỡ mẫu cho một nghiên cứu về tác động của hút thuốc đối với ung thư phổi. Biết rằng, nghiên cứu trước đây cho biết nhóm tỷ lệ bị ung thư phổi ở nhóm hút thuốc và không hút thuốc lần lượt là 34% và 17%.

Lấy mức ý nghĩa thống kê = 5% cho kiểm định 2 phía và sai số 5%. Cỡ mẫu được tính toán theo công thức như sau:

$$n = [0,34(1 - 0,34) + 0,17(1 - 0,17)]\left(\frac{1,96}{0,05}\right)^2 = 562$$

* *Tính toán theo phần mềm HSS 1.0*

Nghiên cứu 2 mẫu độc lập, xác định sự khác biệt 2 tỷ lệ

Tính cỡ mẫu

Alpha (α)	Tỷ lệ không trả lời
0.05	0
$Z_{1-\alpha/2} = 1.96$	Hệ số thiết kế
Tỷ lệ nhóm 1 (P_1)	1
0.34	
Tỷ lệ nhóm 2 (P_2)	
0.17	
Sai số (d)	
0.05	

Cỡ mẫu

562

Ghi chú: Xem thêm phần hiệu chỉnh cỡ mẫu theo tỷ lệ không trả lời và hệ số thiết kế để tính toán cỡ mẫu cuối cùng

2. Tính toán cỡ mẫu cho nghiên cứu sử dụng thống kê suy luận dựa trên kiểm định giả thuyết

Tính toán cỡ mẫu nghiên cứu tùy thuộc vào phương pháp ngoại suy thống kê (thống kê suy luận) được áp dụng là ước lượng khoảng (*Confidence interval*) hay kiểm định giả thuyết (*Hypothesis testing*).

Khi nghiên cứu sử dụng thống kê suy luận dựa trên kiểm định giả thuyết (*Hypothesis testing*), chúng ta cần tính cỡ mẫu để đạt được mức ý nghĩa thống kê (*Significance level*) phù hợp để bác bỏ sai lầm loại 1 (hay loại α) và hiệu lực thống kê (*Power*) phù hợp để bác bỏ sai lầm loại 2 (hay loại β).

Cỡ mẫu cho các nghiên cứu sử dụng thống kê suy luận dựa trên kiểm định giả thuyết (*Hypothesis testing*) được xác định dựa trên công thức chung sau:

$$n = \left(\frac{Z_{1-\frac{\alpha}{2}} + Z_{1-\beta}}{ES} \right)^2$$

Trong đó:

- n là cỡ mẫu
- $Z_{1-\frac{\alpha}{2}}$ là giá trị từ phân bố chuẩn, được tính dựa trên xác suất sai lầm loại 1 ($Z_{1-\frac{\alpha}{2}} = 1,96$ nếu xác suất sai lầm loại 1 = 5% và kiểm định 2 phía).
- $Z_{1-\beta}$ là giá trị từ phân bố chuẩn, được tính dựa trên hiệu lực thống kê ($Z_{1-\beta} = 0,842$ nếu lực thống kê là 80%). Giá trị $Z_{1-\beta}$ được trình bày ở bảng dưới.

Type II error (β)	(1- β)	Z score (1- β)
0.2	0.8	0.842
0.1	0.9	1.282
0.05	0.95	1.645
0.01	0.99	2.326

- ES là mức khác biệt (ES = Effect Size) được tính theo công thức:

- *Đối với biến định lượng*

$$ES = \frac{\mu_1 - \mu_0}{\sigma}$$

Trong đó:

- μ_1 là kết quả nghiên cứu theo giả thuyết H_a
- μ_0 là kết quả nghiên cứu theo giả thuyết H_0
- σ là độ lệch chuẩn

- *Đối với biến định tính*

$$ES = \frac{p_1 - p_0}{\sqrt{p_0(1 - p_0)}}$$

Trong đó:

- n là cỡ mẫu
- p_1 là kết quả theo giả thuyết H_a

- p_0 là theo giả thuyết H_0

Trên thực tế, tử số trong công thức tính toán cỡ mẫu nêu trên chính là hằng số (phụ thuộc vào giá trị cố định của $Z_{1-\frac{\alpha}{2}}$ và $Z_{1-\beta}$):

$$C = Z_{1-\frac{\alpha}{2}} + Z_{1-\beta}$$

Giá trị của C được liệt kê ở bảng dưới đây:

$\alpha =$	$\beta = 0.20$ (Power = 0.80)	$\beta = 0.10$ (Power = 0.90)	$\beta = 0.05$ (Power = 0.95)
0.10	6.15	8.53	10.79
0.05	7.85	10.51	13.00
0.01	13.33	16.74	19.84

2.1. Nghiên cứu gồm 1 mẫu, kiểm định một trung bình

- Công thức:

$$n = \left(\frac{Z_{1-\frac{\alpha}{2}} + Z_{1-\beta}}{ES} \right)^2$$

$$ES = \frac{\mu_1 - \mu_0}{\sigma}$$

Trong đó:

- n là cỡ mẫu
- μ_1 là kết quả theo giả thuyết H_a
- μ_0 là kết quả theo giả thuyết H_0
- $Z_{1-\frac{\alpha}{2}}$ là giá trị từ phân bố chuẩn, được tính dựa trên xác suất sai lầm loại 1 ($Z_{1-\frac{\alpha}{2}} = 1,96$ nếu xác suất sai lầm loại 1 = 5% và kiểm định 2 phía).
- $Z_{1-\beta}$ là giá trị từ phân bố chuẩn, được tính dựa trên lực thống kê ($Z_{1-\beta} = 0,842$ nếu lực thống kê là 80%).
- ES là mức khác biệt
- σ là độ lệch chuẩn

○ $H_0: \mu_1 = \mu_0$

○ $H_a: \mu_1 \neq \mu_0$

• Ví dụ

Tính toán cỡ mẫu cho một nghiên cứu cắt ngang xác định hàm lượng đường máu ở những người uống cà phê hàng ngày. Biết rằng nghiên cứu trước đây báo cáo, ở những người uống cà phê hàng ngày, đường máu trung bình là 95 mg/dl và độ lệch chuẩn là 9,8 mg/dl. Chọn $\alpha = 5\%$, $1 - \beta = 80\%$ và mức hàm lượng đường máu 100 mg/dl là có ý nghĩa lâm sàng. Cỡ mẫu được tính toán theo công thức như sau:

$$\text{Mức khác biệt} = ES = \frac{\mu_1 - \mu_0}{\sigma} = \frac{100 - 95}{9,8} = 0,51$$

$$\text{Cỡ mẫu} = n = \left(\frac{Z_{1-\frac{\alpha}{2}} + Z_{1-\beta}}{ES} \right)^2 = \left(\frac{1,96 + 0,842}{0,51} \right)^2 = 31$$

* Tính toán theo phần mềm HSS 1.0:

The screenshot shows the HSS 1.0 software interface for calculating sample size. The title bar reads "Nghiên cứu gồm 1 mẫu, kiểm định 1 trung bình". There are two tabs: "Tính cỡ mẫu" (selected) and "Tính lực thống kê".

Input fields and their values:

- Alpha (α): 0.05
- Tỷ lệ không trả lời: 0
- Lực thống kê ($1-\beta$): 0.8
- Hệ số thiết kế: 1
- Trung bình theo giả thuyết H_1 (μ_1): 100
- Trung bình theo giả thuyết H_0 (μ_0): 95
- Độ lệch chuẩn (σ): 9.8

Calculated values shown:

- $Z_{1-\alpha/2} = 1.96$
- $Z_{1-\beta} = 0.84$

The final result, "Cỡ mẫu", is displayed in a large grey box as **31**.

Ghi chú: Xem thêm phần hiệu chỉnh cỡ mẫu theo tỷ lệ không trả lời và hệ số thiết kế để tính toán cỡ mẫu cuối cùng

2.2. Nghiên cứu gồm 1 mẫu, kiểm định một tỷ lệ

- Công thức:

$$n = \frac{\left\{ Z_{1-\frac{\alpha}{2}} \sqrt{P_0(1-P_0)} + Z_{1-\beta} \sqrt{P_a(1-P_a)} \right\}^2}{(P_a - P_0)^2}$$

Trong đó:

- n là cỡ mẫu
 - P_a là kết quả theo giả thuyết H_a
 - P_0 là kết quả theo giả thuyết H_0
 - $Z_{1-\frac{\alpha}{2}}$ là giá trị từ phân bố chuẩn, được tính dựa trên xác suất sai lầm loại 1 ($Z_{1-\frac{\alpha}{2}} = 1,96$ nếu xác suất sai lầm loại 1 = 5% và kiểm định 2 phía).
 - $Z_{1-\beta}$ là giá trị được tính dựa trên lực thống kê ($Z_{1-\beta} = 0,842$ nếu lực thống kê là 80%)
- Ví dụ

Tính toán cỡ mẫu cho một nghiên cứu cắt ngang về tỷ lệ tăng huyết áp. Biết rằng tỷ lệ tăng huyết áp ước tính ở nghiên cứu này là 31% và nghiên cứu trước đây báo cáo tỷ lệ này là 26%. Chọn $\alpha = 5\%$, $1 - \beta = 80\%$.

Cỡ mẫu được tính toán theo công thức như sau:

$$\begin{aligned} n &= \frac{\left\{ Z_{1-\frac{\alpha}{2}} \sqrt{P_0(1-P_0)} + Z_{1-\beta} \sqrt{P_a(1-P_a)} \right\}^2}{(P_a - P_0)^2} \\ &= \frac{\left\{ 1,96 \sqrt{0,26 \times (1 - 0,26)} + 0,842 \sqrt{0,31 \times (1 - 0,31)} \right\}^2}{(0,31 - 0,26)^2} \\ &= 624 \end{aligned}$$

* Tính toán theo phần mềm HSS 1.0:

Nghiên cứu gồm 1 mẫu, kiểm định 1 tỷ lệ

Tính cỡ mẫu Tính lực thống kê

Alpha (α)	Tỷ lệ không trả lời
0.05	0
$Z_{1-\alpha/2} = 1.96$	Hệ số thiết kế
Lực thống kê ($1-\beta$)	1
0.8	
$Z_{1-\beta} = 0.84$	
Kết quả theo giả thuyết H_1 (P_1)	Cỡ mẫu 624
0.31	
Kết quả theo giả thuyết H_0 (P_0)	
0.28	

Ghi chú: Xem thêm phần hiệu chỉnh cỡ mẫu theo tỷ lệ không trả lời và hệ số thiết kế để tính toán cỡ mẫu cuối cùng

2.3. Nghiên cứu gồm 2 mẫu độc lập, kiểm định 2 số trung bình

- Công thức

$$n = 2 \left(\frac{Z_{1-\frac{\alpha}{2}} + Z_{1-\beta}}{ES} \right)^2$$

$$ES = \frac{\mu_1 - \mu_2}{\sigma}$$

Trong đó:

- n là cỡ mẫu cho mỗi nhóm
- μ_1 là số trung bình của nhóm 1 (nhóm không phơi nhiễm hoặc nhóm chứng)
- μ_2 là số trung bình của nhóm 2 (nhóm phơi nhiễm hoặc nhóm can thiệp)
- $Z_{1-\frac{\alpha}{2}}$ là giá trị từ phân bố chuẩn, được tính dựa trên xác suất sai lầm loại 1 ($Z_{1-\frac{\alpha}{2}} = 1,96$ nếu xác suất sai lầm loại 1 = 5% và kiểm định 2 phía).

- $Z_{1-\beta}$ là giá trị được tính dựa trên lực thống kê ($Z_{1-\beta} = 0,842$ nếu lực thống kê là 80%)
- ES là mức khác biệt
- σ là độ lệch chuẩn chung theo công thức dưới đây hoặc của nhóm 1 (nhóm không phơi nhiễm hoặc nhóm chứng)

$$\sigma = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

- $H_0: \mu_1 = \mu_2$
- $H_a: \mu_1 \neq \mu_2$
- Ví dụ

Tính toán cỡ mẫu cho một nghiên cứu ngang đánh giá hiệu quả điều trị tăng huyết áp của 2 thuốc (thuốc mới và thuốc cũ). Dự kiến thuốc cũ giảm được 5 mmHg và thuốc mới giảm được 10 mmHg sau 1 tháng điều trị. Độ lệch chuẩn về sự thay đổi huyết áp (giảm) là 19 mmHg. Chọn $\alpha = 5\%$, $1 - \beta = 80\%$.

Cỡ mẫu được tính toán theo công thức như sau:

$$\text{Mức khác biệt} = ES = \frac{\mu_1 - \mu_2}{\sigma} = \frac{10 - 5}{19} = 0,2632$$

$$\text{Cỡ mẫu} = n = 2\left(\frac{Z_{1-\frac{\alpha}{2}} + Z_{1-\beta}}{ES}\right)^2 = 2\left(\frac{1,96 + 0,842}{0,2632}\right)^2 = 227$$

Vậy cỡ mẫu cho mỗi nhóm là: $n = 227$

* Tính toán theo phần mềm HSS 1.0:

Nghiên cứu gồm 2 mẫu độc lập, kiểm định 2 trung bình

Tính cỡ mẫu Tính lực thống kê

Alpha (α)
0.05

$Z_{1-\alpha/2} = 1.96$

Lực thống kê ($1-\beta$)
0.8

$Z_{1-\beta} = 0.84$

Trung bình nhóm 1 (μ_1)
10

Trung bình nhóm 2 (μ_2)
5

Chọn cách nhập
 Độ lệch chuẩn gộp
 Độ lệch chuẩn của 2 nhóm

Độ lệch chuẩn gộp (σ)
19

Tỷ số 2 nhóm
1

Tỷ lệ không trả lời
0

Hệ số thiết kế
1

Cỡ mẫu nhóm 1
227

Cỡ mẫu nhóm 2
227

Ghi chú: Xem thêm phần hiệu chỉnh cỡ mẫu theo tỷ lệ không trả lời và hệ số thiết kế để tính toán cỡ mẫu cuối cùng

2.4. Nghiên cứu gồm 2 mẫu ghép cặp, kiểm định 2 số trung bình

- Công thức

$$n = \left(\frac{Z_{1-\frac{\alpha}{2}} + Z_{1-\beta}}{ES} \right)^2$$

$$ES = \frac{\mu_d}{\sigma_d}$$

Trong đó:

- n là cỡ mẫu cho mỗi nhóm
- μ_d là sự khác biệt giữa 2 lần đo
- $Z_{1-\frac{\alpha}{2}}$ là giá trị từ phân bố chuẩn, được tính dựa trên xác suất sai lầm loại 1 ($Z_{1-\frac{\alpha}{2}} = 1,96$ nếu xác suất sai lầm loại 1 = 5% và kiểm định 2 phía).

- $Z_{1-\beta}$ là giá trị được tính dựa trên lực thống kê ($Z_{1-\beta} = 0,842$ nếu lực thống kê là 80%)
- ES là mức khác biệt
- σ_d là độ lệch chuẩn của sự khác biệt giữa 2 nhóm

$$\sigma = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

- $H_0: \mu_d = 0$
- $H_a: \mu_d \neq 0$
- Ví dụ

Tính toán cỡ mẫu cho một nghiên cứu theo dõi dọc mức tăng cân nặng ở 1 nhóm trẻ qua thời gian. Biết rằng, nghiên cứu trước đây báo cáo mức tăng cân nặng (sự khác biệt về cân nặng giữa 2 lần đo) là 3g và độ lệch chuẩn là 9,1g. Chọn $\alpha = 5\%$, $1 - \beta = 80\%$.

Cỡ mẫu được tính toán theo công thức như sau:

$$\text{Mức khác biệt} = ES = \frac{\mu_d}{\sigma_d} = \frac{3}{9,1} = 0,3297$$

$$\text{Cỡ mẫu} = n = \left(\frac{Z_{1-\frac{\alpha}{2}} + Z_{1-\beta}}{ES} \right)^2 = \left(\frac{1,96 + 0,842}{0,3297} \right)^2 = 73$$

Vậy cỡ mẫu của mỗi nhóm là: $n = 73$

* Tính toán theo phần mềm HSS 1.0:

Nghiên cứu gồm 2 mẫu ghép cặp, kiểm định 2 trung bình

Tính cỡ mẫu

Tính lực thống kê

Alpha (α)
0.05

$Z_{1-\alpha/2} = 1.96$

Lực thống kê ($1-\beta$)
0.8

$Z_{1-\beta} = 0.84$

Sự khác biệt giữa 2 lần đo (μ_d)
3

Chọn cách nhập

- Độ lệch chuẩn gộp
- Độ lệch chuẩn của 2 nhóm

Độ lệch chuẩn gộp (σ)
9.1

Tỷ số 2 nhóm
1

Tỷ lệ không trả lời
0

Hệ số thiết kế
1

Cỡ mẫu nhóm 1

73

Cỡ mẫu nhóm 2

73

Ghi chú: Xem thêm phần hiệu chỉnh cỡ mẫu theo tỷ lệ không trả lời và hệ số thiết kế để tính toán cỡ mẫu cuối cùng

2.5. Nghiên cứu gồm 2 mẫu độc lập, kiểm định 2 tỷ lệ

- Công thức

$$n = 2 \left(\frac{Z_{1-\frac{\alpha}{2}} + Z_{1-\beta}}{ES} \right)^2$$

$$ES = \frac{p_1 - p_2}{\sqrt{p(1-p)}}$$

$$p = \frac{p_1 + p_2}{2}$$

Trong đó:

- n là cỡ mẫu
- p_1 là kết quả theo giả thuyết H_a
- p_0 là tỷ lệ từ nghiên cứu trước đây hay giá trị cần kiểm định (Giả thuyết H_0)

- $Z_{1-\frac{\alpha}{2}}$ là giá trị từ phân bố chuẩn, được tính dựa trên xác suất sai lầm loại 1 ($Z_{1-\frac{\alpha}{2}} = 1,96$ nếu xác suất sai lầm loại 1 = 5% và kiểm định 2 phía).
- $Z_{1-\beta}$ là giá trị được tính dựa trên lực thống kê ($Z_{1-\beta} = 0,842$ nếu lực thống kê là 80%)
- ES là mức khác biệt
- $H_0: p_1 = p_2$
- $H_a: p_1 \neq p_2$
- Ví dụ

Tính toán cỡ mẫu cho một nghiên cứu so sánh hiệu quả điều trị tỷ lệ tăng huyết áp của 2 thuốc (thuốc mới và thuốc cũ). Biết rằng nghiên cứu trước đây báo cáo tỷ lệ đáp ứng ở thuốc cũ là 24%. Nhà nghiên cứu kỳ vọng tỷ lệ đáp ứng ở thuốc mới là 30%. Chọn $\alpha = 5\%$, $1 - \beta = 80\%$ và sai số là 5%.

Cỡ mẫu được tính toán theo công thức như sau:

$$\text{Mức khác biệt} = ES = \frac{p_1 - p_0}{\sqrt{p(1-p)}} = \frac{0,30 - 0,24}{\sqrt{0,27(1-0,27)}} = 0,135$$

$$\text{Cỡ mẫu} = n = 2 \left(\frac{Z_{1-\frac{\alpha}{2}} + Z_{1-\beta}}{ES} \right)^2$$

$$n = 2 \left(\frac{1,96 + 0,842}{0,135} \right)^2 = 860$$

Vậy cỡ mẫu của mỗi nhóm là: $n = 860$

* Tính toán theo phần mềm HSS 1.0:

Nghiên cứu gồm 2 mẫu độc lập, kiểm định 2 tỷ lệ

Tính cỡ mẫu Tính lực thống kê

Alpha (α)
0.05

Tỷ lệ không trả lời
0

$Z_{1-\alpha/2} = 1.96$

Lực thống kê ($1-\beta$)
0.8

Hệ số thiết kế
1

$Z_{1-\beta} = 0.84$

Tỷ lệ ở nhóm can thiệp (p_1)
0.30

Tỷ số 2 nhóm
1

Tỷ lệ ở nhóm chứng (p_2)
0.24

Cỡ mẫu nhóm 1
860

Cỡ mẫu nhóm 2
860

Ghi chú: Xem thêm phần hiệu chỉnh cỡ mẫu theo tỷ lệ không trả lời và hệ số thiết kế để tính toán cỡ mẫu cuối cùng

2.6. Nghiên cứu gồm 2 mẫu ghép cặp, kiểm định 2 tỷ lệ (McNemar)

- Công thức

Kiểm định này thường được sử dụng trong nghiên cứu bệnh chứng ghép cặp. Các đối tượng nghiên cứu được phân làm 2 nhóm (có bệnh và không có bệnh) và được theo dõi 2 lần. Có sự chuyển đổi về tình trạng có bệnh-không có bệnh giữa 2 nhóm như dưới đây:

Bảng 2: Nghiên cứu bệnh chứng ghép cặp (n)

Lần 1	Lần 2		
	Có bệnh	Không bệnh	Tổng
Có bệnh	n11	n10	N1+
Không bệnh	n01	n00	N0+
Tổng	N+1	N+0	N

Bảng 3: Nghiên cứu bệnh chứng ghép cặp (p)

Lần 1	Lần 2		
	Có bệnh	Không bệnh	Tổng
Có bệnh	p11	p10	P1+
Không bệnh	p01	p00	P0+
Tổng	P+1	P+0	100%

$$n = \frac{\left\{ Z_{1-\frac{\alpha}{2}}(OR + 1) + Z_{1-\beta} \sqrt{(OR + 1)^2 - (OR - 1)^2 PD} \right\}^2}{(OR - 1)^2 PD}$$

Trong đó:

- $Z_{1-\frac{\alpha}{2}}$ là giá trị từ phân bố chuẩn, được tính dựa trên xác suất sai lầm loại 1 ($Z_{1-\frac{\alpha}{2}} = 1,96$ nếu xác suất sai lầm loại 1 = 5% và kiểm định 2 phía).
- $Z_{1-\beta}$ là giá trị được tính dựa trên lực thống kê ($Z_{1-\beta} = 0,842$ nếu lực thống kê là 80%)
- $OR = \frac{P10}{P01}$
- $PD = P10 + P01$
- Ví dụ

Chọn $\alpha = 5\%$, $1 - \beta = 80\%$. Nghiên cứu ghép cặp có sự chuyển đổi về tình trạng có bệnh-không có bệnh giữa 2 nhóm như dưới đây

Lần 1	Lần 2	
	Có bệnh	Không bệnh
Có bệnh	P11 = 27%	P10 = 32%
Không bệnh	P01 = 16%	P00 = 25%

Cỡ mẫu được tính toán theo công thức như sau:

$$OR = \frac{P10}{P01} = \frac{0,32}{0,16} = 2$$

$$PD = P10 + P01 = 0,32 + 0,16 = 0,48$$

$$n = \frac{\left\{ Z_{1-\frac{\alpha}{2}}(OR + 1) + Z_{1-\beta}\sqrt{(OR + 1)^2 - (OR - 1)^2 PD} \right\}^2}{(OR - 1)^2 PD}$$

$$= \frac{\left\{ 1,96(2 + 1) + 1,28\sqrt{(2 + 1)^2 - (2 - 1)^2 0,48} \right\}^2}{(2 - 1)^2 0,48} = 193$$

* *Tính toán theo phần mềm HSS 1.0:*

Nghiên cứu gồm 2 mẫu ghép cặp, kiểm định 2 tỷ lệ (McNemar)

Tính cỡ mẫu Tính lực thống kê

P10 0.32	Tỷ lệ không trả lời 0
P01 0.16	Hệ số thiết kế 1
Alpha (α) 0.05	Cỡ mẫu 193
Z _{1-α/2} = 1.96	
Lực thống kê (1-β) 0.9	
Z _{1-β} = 1.28	
OR = 2	
PD = 0.48	

Ghi chú: Xem thêm phần hiệu chỉnh cỡ mẫu theo tỷ lệ không trả lời và hệ số thiết kế để tính toán cỡ mẫu cuối cùng

2.7. Nghiên cứu bệnh chứng, kiểm định OR

- *Công thức*

Trong các nghiên cứu bệnh chứng (Case-control study), chúng ta nghiên cứu về mối liên quan giữa một yếu tố kết quả (Outcome) và một yếu tố phơi nhiễm nào đó (Exposure) và phải tính toán tỷ số chênh (OR) để thể hiện mối liên quan này. Cỡ mẫu của nghiên cứu bệnh chứng được tính toán theo công thức:

$$n = \frac{\left\{ Z_{1-\frac{\alpha}{2}} \sqrt{2P_2(1-P_2)} + Z_{1-\beta} \sqrt{P_1(1-P_1) + P_2(1-P_2)} \right\}^2}{(P_1 - P_2)^2}$$

Trong đó:

- n là cỡ mẫu chung
- $Z_{1-\frac{\alpha}{2}}$ là giá trị từ phân bố chuẩn, được tính dựa trên xác suất sai lầm loại 1 ($Z_{1-\frac{\alpha}{2}} = 1,96$ nếu xác suất sai lầm loại 1 = 5% và kiểm định 2 phía).
- $Z_{1-\beta}$ là giá trị được tính dựa trên lực thống kê ($Z_{1-\beta} = 0,842$ nếu lực thống kê là 80%)
- P_1 là tỷ lệ có yếu tố phơi nhiễm trong nhóm có bệnh
- P_2 là tỷ lệ có yếu tố phơi nhiễm trong nhóm không có bệnh
- Ví dụ

Tính toán cỡ mẫu cho một nghiên cứu bệnh chứng về mối quan giữa hút thuốc lá (phơi nhiễm) và ung thư phổi (bệnh). Biết rằng, nghiên cứu trước đây báo cáo tỷ lệ hút thuốc lá trong nhóm có ung thư phổi là 35% và tỷ lệ hút thuốc lá trong nhóm chứng là 17,5%. Chọn $\alpha = 5\%$ và $1 - \beta = 80\%$. Cỡ mẫu được tính toán theo công thức như sau:

$$\begin{aligned} n &= \frac{\left\{ Z_{1-\frac{\alpha}{2}} \sqrt{2P_2(1-P_2)} + Z_{1-\beta} \sqrt{P_1(1-P_1) + P_2(1-P_2)} \right\}^2}{(P_1 - P_2)^2} \\ &= \frac{\left\{ 1,96 \sqrt{2 \times 0,175 \times (1 - 0,175)} + 0,842 \sqrt{0,35 \times (1 - 0,35) + 0,175 \times (1 - 0,175)} \right\}^2}{(0,35 - 0,175)^2} \\ &= 81 \end{aligned}$$

Cỡ mẫu cần có cho mỗi nhóm $n = 81$

* Tính toán theo phần mềm HSS 1.0:

Nghiên cứu bệnh chứng, kiểm định OR

Tính cỡ mẫu Tính lực thống kê

Alpha (α)	Tỷ lệ không trả lời
0.05	0
$Z_{1-\alpha/2} = 1.96$	Hệ số thiết kế
Tỷ lệ phơi nhiễm trong nhóm chứng (P_2)	1
0.175	Tỷ số 2 nhóm
Lực thống kê ($1-\beta$)	1
0.8	
$Z_{1-\beta} = 0.84$	
Tỷ lệ phơi nhiễm trong nhóm bệnh (P_1)	
0.35	

Cỡ mẫu nhóm 1

81

Cỡ mẫu nhóm 2

81

Ghi chú: Xem thêm phần hiệu chỉnh cỡ mẫu theo tỷ lệ không trả lời và hệ số thiết kế để tính toán cỡ mẫu cuối cùng

2.8. Nghiên cứu thuần tập, kiểm định RR

- Công thức

Trong các nghiên cứu thuần tập (Cohort study), chúng ta nghiên cứu về mối liên quan giữa một yếu tố kết quả (Outcome) và một yếu tố phơi nhiễm nào đó (Exposure) và phải tính toán tỷ số tỷ lệ (hay còn gọi là nguy cơ tương đối RR) để thể hiện mối liên quan này. Cỡ mẫu của nghiên cứu thuần tập được tính toán theo công thức:

$$n = \frac{\left\{ Z_{1-\frac{\alpha}{2}} \sqrt{2\bar{P}(1-\bar{P})} + Z_{1-\beta} \sqrt{P_1(1-P_1) + P_2(1-P_2)} \right\}^2}{(P_1 - P_2)^2}$$

- n là cỡ mẫu chung
- $Z_{1-\frac{\alpha}{2}}$ là giá trị từ phân bố chuẩn, được tính dựa trên xác suất sai lầm loại 1 ($Z_{1-\frac{\alpha}{2}} = 1,96$ nếu xác suất sai lầm loại 1 = 5% và kiểm định 2 phía).

- $Z_{1-\beta}$ là giá trị được tính dựa trên lực thống kê ($Z_{1-\beta} = 0,842$ nếu lực thống kê là 80%)
- P_1 : Tỷ lệ mắc bệnh trong nhóm phơi nhiễm
- P_2 : Tỷ lệ mắc bệnh trong nhóm không phơi nhiễm
- $\bar{P} = (P_1+P_2)/2$: Tỷ lệ trung bình của 2 nhóm

- *Ví dụ*

Tính toán cỡ mẫu cho một nghiên cứu đoàn hệ về mối quan giữa hút thuốc lá (phơi nhiễm) và ung thư phổi (bệnh). Biết rằng, nghiên cứu trước đây báo cáo tỷ lệ ung thư phổi trong nhóm có hút thuốc lá là 35% và tỷ lệ ung thư phổi trong nhóm không hút thuốc lá là 17,5%. Chọn $\alpha = 5\%$ và $1 - \beta = 80\%$. Cỡ mẫu được tính toán theo công thức như sau:

$$\bar{P} = \frac{P_1 + P_2}{2} = \frac{0,35 + 0,175}{2} = 0,2625$$

$$n = \frac{\left\{ Z_{1-\frac{\alpha}{2}} \sqrt{2\bar{P}(1-\bar{P})} + Z_{1-\beta} \sqrt{P_1(1-P_1) + P_2(1-P_2)} \right\}^2}{(P_1 - P_2)^2}$$

$$= \frac{\left\{ 1,96 \sqrt{2 \times 0,2625 \times (1 - 0,2625)} + 0,842 \sqrt{0,35 \times (1 - 0,35) + 0,175 \times (1 - 0,175)} \right\}^2}{(0,35 - 0,175)^2}$$

$$= 99$$

Cỡ mẫu cần có cho mỗi nhóm $n = 99$

- *Tính toán theo phần mềm HSS 1.0:*

Nghiên cứu thuần tập, kiểm định RR

Tính cỡ mẫu Tính lực thống kê

Alpha (α)
0.05

Tỷ lệ không trả lời
0

$Z_{1-\alpha/2} = 1.96$

Lực thống kê ($1-\beta$)
0.8

Hệ số thiết kế
1

$Z_{1-\beta} = 0.84$

Tỷ lệ phơi nhiễm trong nhóm bệnh (P_1)
0.175

Tỷ số 2 nhóm
1

Tỷ lệ phơi nhiễm trong nhóm chứng (P_2)
0.35

Cỡ mẫu nhóm 1
99

Cỡ mẫu nhóm 2
99

Ghi chú: Xem thêm phần hiệu chỉnh cỡ mẫu theo tỷ lệ không trả lời và hệ số thiết kế để tính toán cỡ mẫu cuối cùng

2.9. Nghiên cứu sống còn (survival analysis study)

- Công thức

Trong nghiên cứu sống còn, nếu thời gian theo dõi đối tượng được xác định, cỡ mẫu được tính toán theo công thức:

$$n = \frac{(HR + 1)^2 (Z_{1-\frac{\alpha}{2}} + Z_{1-\beta})}{(HR - 1)^2 (2 - P_1 - P_2)}$$

Trong đó:

- n là cỡ mẫu
- $Z_{1-\frac{\alpha}{2}}$ là giá trị từ phân bố chuẩn, được tính dựa trên xác suất sai lầm loại 1 ($Z_{1-\frac{\alpha}{2}} = 1,96$ nếu xác suất sai lầm loại 1 = 5% và kiểm định 2 phía).
- $Z_{1-\beta}$ là giá trị được tính dựa trên lực thống kê ($Z_{1-\beta} = 0,842$ nếu lực thống kê là 80%)

○ $HR = \frac{\ln(p_1)}{\ln(p_2)}$ là tỷ số nguy hại (Hazard Ratio)

○ P_1 và P_2 là tỷ lệ biến cố ở nhóm can thiệp và nhóm đối chứng

• Ví dụ

Tính toán cỡ mẫu cho một nghiên cứu sống còn trong 2 năm. Biết rằng, nghiên cứu trước đây báo cáo tỷ lệ sống còn ở nhóm chứng (P_1) là 25% và các nhà nghiên cứu kỳ vọng sẽ kéo dài thời gian sống ở nhóm can thiệp cao gấp 1,5 lần so với nhóm chứng ($P_2 = 39,7\%$). Chọn $\alpha = 5\%$ và $1 - \beta = 80\%$. Cỡ mẫu được tính toán theo công thức như sau:

$$n = \frac{(HR + 1)^2 (Z_{1-\frac{\alpha}{2}} + Z_{1-\beta})}{(HR - 1)^2 (2 - p_1 - p_2)} = \frac{(1,5 + 1)^2 (1,96 + 0,842)}{(1,5 - 1)^2 (2 - 0,25 - 0,397)} = 52$$

* Tính toán theo phần mềm HSS 1.0:

Nghiên cứu sống còn

Tính cỡ mẫu Tính lực thống kê

Tỷ lệ ở nhóm can thiệp (p_1)
0.25

Chọn loại tham số
 Tỷ số nguy hại
 Tỷ lệ ở nhóm chứng

Tỷ số nguy hại (HR)
1.5

Tỷ lệ ở nhóm chứng (p_2) = 0.397

Alpha (α)
0.05

$Z_{1-\alpha/2} = 1.96$

Lực thống kê ($1-\beta$)
0.8

$Z_{1-\beta} = 0.84$

Tỷ lệ không trả lời
0

Hệ số thiết kế
1

Tỷ số 2 nhóm
1

Cỡ mẫu nhóm 1
52

Cỡ mẫu nhóm 2
52

Ghi chú: Xem thêm phần hiệu chỉnh cỡ mẫu theo tỷ lệ không trả lời và hệ số thiết kế để tính toán cỡ mẫu cuối cùng

3. Nghiên cứu nghiệm pháp chẩn đoán

- Công thức

Nghiên cứu về nghiệm pháp chẩn đoán thường đề cập đến độ nhạy (Sensitivity) và đặc hiệu (Specificity) như trình bày ở bảng dưới đây

Bảng 4: Khái niệm trong nghiên cứu về nghiệm pháp chẩn đoán

Nghiệm pháp chẩn đoán	Thực tế	
	Có bệnh	Không bệnh
Có bệnh	Độ nhạy (Dương tính thật)	Dương tính giả
Không bệnh	Âm tính giả	Độ đặc hiệu (Âm tính thật)

Một nghiệm pháp chẩn đoán được xem là có giá trị và có thể sử dụng trong thực hành lâm sàng cần phải đạt độ nhạy và đặc hiệu tối thiểu 0.75 (hay tốt hơn nữa là 0.80). Giá trị của nghiệm pháp chẩn đoán còn tùy thuộc vào tỉ lệ hiện mắc (prevalence) của bệnh trong một quần thể. Do đó, phương pháp các định cỡ mẫu nghiên cứu nghiệm pháp chẩn đoán phải dựa vào các chỉ số này.

Cụ thể, các tham số cần có như sau:

- Sen: Độ nhạy (Xác suất dương tính thật)
- Spe: Độ đặc hiệu (Xác suất âm tính thật)
- d: Mức sai số của Sen và Spe
- P: Tỉ lệ hiện mắc (prevalence) của bệnh trong một quần thể

Cỡ mẫu được tính toán theo công thức:

- Cỡ mẫu cho độ nhạy

$$n(\text{sen}) = \frac{TP + FN}{P}$$

Trong đó:

- TP là số dương tính thật – true positive
- FN là số âm tính giả - false negative

- P là tỷ lệ hiện mắc

TP + FN (Số có bệnh trên thực tế) được tính bằng công thức

$$TP + FN = \frac{Z_{1-\frac{\alpha}{2}}^2 * Sen * (1 - Sen)}{d^2}$$

Trong đó:

- $Z_{1-\frac{\alpha}{2}}$ là giá trị từ phân bố chuẩn, được tính dựa trên mức ý nghĩa thống kê ($Z_{1-\frac{\alpha}{2}} = 1,96$ nếu mức ý nghĩa thống kê = 5% và kiểm định 2 phía)
- Sen là độ nhạy
- d là mức sai số
- *Cỡ mẫu cho độ đặc hiệu*

$$n(spe) = \frac{TN + FP}{1 - P}$$

Trong đó:

- TN là số âm tính thật – true negative
- FP là số dương tính giả - false positive
- P là tỷ lệ hiện mắc

TN + FP (Số không có bệnh trên thực tế) được tính bằng công thức

$$TN + FP = \frac{Z_{1-\frac{\alpha}{2}}^2 * Spe * (1 - Spe)}{d^2}$$

Trong đó:

- $Z_{1-\frac{\alpha}{2}}$ là giá trị từ phân bố chuẩn, được tính dựa trên mức ý nghĩa thống kê ($Z_{1-\frac{\alpha}{2}} = 1,96$ nếu mức ý nghĩa thống kê = 5% và kiểm định 2 phía)
- Spe là độ đặc hiệu
- d là mức sai số
- *Ví dụ*

Tính toán cỡ mẫu cho một nghiên cứu cho một nghiệm pháp chẩn đoán. Các nhà nghiên cứu độ nhạy là khoảng 0,8 và độ đặc hiệu là khoảng 0.9. Biết rằng tỷ lệ hiện mắc là 20%. Chọn $\alpha = 5\%$. Cỡ mẫu được tính như sau:

Để tính toán độ nhạy

$$TP + FN = \frac{Z_{1-\frac{\alpha}{2}}^2 * Sen * (1 - Sen)}{d^2} = \frac{1,96^2 * 0,8 * (1 - 0,8)}{0,05^2} = 246$$

$$n(sen) = \frac{TP + FN}{P} = \frac{246}{0,2} = 1230$$

* Tính toán theo phần mềm HSS 1.0:

Nghiên cứu tính toán độ nhạy

Tính cỡ mẫu

Alpha (α)	0.05
$Z_{1-\alpha/2} = 1.96$	
Độ nhạy ước tính (Sens)	0.8
Sai số tuyệt đối (d)	0.05
Tỷ lệ bệnh hiện hành (P)	0.2

Cỡ mẫu

1230

Để tính toán độ đặc hiệu

$$TN + FP = \frac{Z_{1-\frac{\alpha}{2}}^2 * Spe * (1 - Spe)}{d^2} = \frac{1,96^2 * 0,9 * (1 - 0,9)}{0,05^2} = 138$$

$$n(spe) = \frac{138}{1 - 0,2} = 173$$

* Tính toán theo phần mềm HSS 1.0:

Nghiên cứu tính toán độ đặc hiệu

Tính cỡ mẫu

Alpha (α)
0.05

$Z_{1-\alpha/2} = 1.96$

Độ đặc hiệu ước tính (Spec)
0.9

Sai số tuyệt đối (d)
0.05

Tỷ lệ bệnh hiện hành (P)
0.2

Cỡ mẫu
173

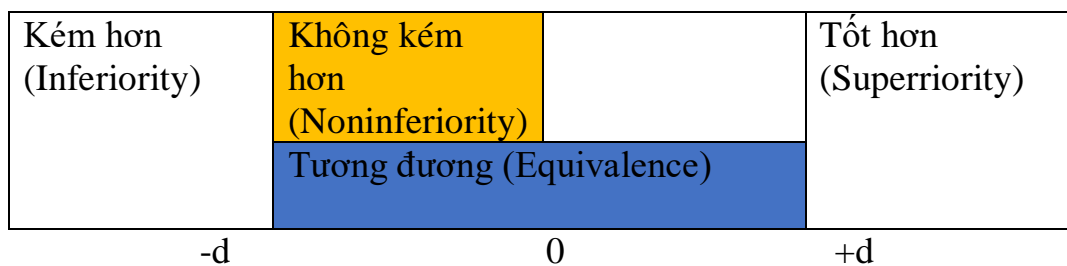
Chúng ta cần xác định từ đâu là ưu tiên hay độ nhạy hay độ đặc hiệu để quyết định tính cỡ mẫu.

4. Nghiên cứu tương đương (Equivalence trial)

- Công thức

Các nghiên cứu tương đương thường áp dụng nhằm mục tiêu chứng minh hiệu quả của 2 giải pháp điều trị là tương đương nhau (hay nói cách khác là không có sự khác biệt nhiều). Các nhà nghiên cứu cần đưa ra ngưỡng (d) và nếu sự khác biệt về hiệu quả điều trị nằm trong khoảng từ -d đến +d thì có thể kết luận là hiệu quả của 2 giải pháp điều trị là tương đương nhau. Lưu ý, đây là kiểm định 2 phía.

- H_0 : Có sự khác biệt có ý nghĩa lâm sàng giữa 2 phương pháp điều trị:
 $\mu_1 - \mu_2 \leq -d$ và $\mu_1 - \mu_2 \geq +d$
- H_a : Không có sự khác biệt có ý nghĩa lâm sàng giữa 2 phương pháp điều trị:
 $-d < |\mu_1 - \mu_2| < +d$



- *Đối với biến định lượng:*

$$n = \frac{2(Z_{1-\frac{\alpha}{2}} + Z_{1-\beta})}{H^2}$$

$$H = \frac{(|\mu_1 - \mu_2|) - d}{\sigma}$$

Trong đó:

- n là cỡ mẫu
- $Z_{1-\frac{\alpha}{2}}$ là giá trị từ phân bố chuẩn, được tính dựa trên xác suất sai lầm loại 1 ($Z_{1-\frac{\alpha}{2}} = 1,96$ nếu xác suất sai lầm loại 1 = 5% và kiểm định 2 phía).
- $Z_{1-\beta}$ là giá trị được tính dựa trên lực thống kê ($Z_{1-\beta} = 0,842$ nếu lực thống kê là 80%)
- H là mức độ khác biệt
- μ_1 là số trung bình của nhóm can thiệp
- μ_2 là số trung bình của nhóm đối chứng
- d là ngưỡng khác biệt (điểm cắt)
Nếu $-d < |\mu_1 - \mu_2| < +d$, chúng ta có thể kết luận hiệu quả của nhóm can thiệp và nhóm chứng tương đương nhau
- σ là độ lệch chuẩn chung của 2 nhóm (xem công thức tính độ lệch chuẩn chung)
- *Đối với biến định tính:*

$$n = \frac{2(Z_{1-\frac{\alpha}{2}} + Z_{1-\beta})}{H^2}$$

$$H = \frac{(|p_1 - p_2|) - d}{\sqrt{p_1(1-p_1) + p_2(1-p_2)}}$$

Trong đó:

- n là cỡ mẫu cho mỗi nhóm
- $Z_{1-\frac{\alpha}{2}}$ là giá trị từ phân bố chuẩn, được tính dựa trên xác suất sai lầm loại 1 ($Z_{1-\frac{\alpha}{2}} = 1,96$ nếu xác suất sai lầm loại 1 = 5% và kiểm định 2 phía).

- $Z_{1-\beta}$ là giá trị được tính dựa trên lực thống kê ($Z_{1-\beta} = 0,842$ nếu lực thống kê là 80%)
- H là mức độ khác biệt
- p1 là tỷ lệ ở nhóm can thiệp
- p2 là tỷ lệ ở nhóm nhóm đối chứng
- d là ngưỡng khác biệt (điểm cắt)
- Nếu $p1-p2 < d$, chúng ta có thể kết luận hiệu quả của nhóm can thiệp và nhóm chứng tương đương nhau
- Ví dụ
 - Đối với biến định lượng:

Tính toán cỡ mẫu cho nghiên cứu chứng minh hiệu quả 2 thuốc điều trị loãng xương. Hiệu quả kỳ vọng của thuốc mới và thuốc cũ lần lượt là 7 và 4 điểm. Độ lệch chuẩn của mức tăng là 10 điểm. Các nhà nghiên cứu quyết định nếu sự khác biệt < 2 điểm thì coi như 2 thuốc có hiệu quả tương đương nhau. Chọn $\alpha = 5\%$ và $1 - \beta = 80\%$. Cỡ mẫu được tính như sau

$$H = \frac{(|\mu_1 - \mu_2|) - d}{\sigma} = \frac{(|7 - 4|) - 2}{10} = 0,1$$

$$n = \frac{2(Z_{1-\frac{\alpha}{2}} + Z_{1-\beta})^2}{H^2} = \frac{2(1,96 + 0,842)^2}{0,1^2} = 561$$

Cỡ mẫu cho mỗi nhóm $n = 561$

* Tính toán theo phần mềm HSS 1.0:

Nghiên cứu tương đương với biến định lượng

Tính cỡ mẫu Tính lực thống kê

Alpha (α)
0.05
 $Z_{1-\alpha/2} = 1.96$

Lực thống kê ($1-\beta$)
0.8
 $Z_{1-\beta} = 0.84$

Trung bình của nhóm can thiệp (μ_1)
7

Trung bình của nhóm đối chứng (μ_2)
4

Ngưỡng khác biệt (d)
2

Độ lệch chuẩn (σ)
10

Tỷ lệ không trả lời
0

Hệ số thiết kế
1

Tỷ số 2 nhóm
1

Cỡ mẫu nhóm can thiệp
561

Cỡ mẫu nhóm đối chứng
561

Ghi chú: Xem thêm phần hiệu chỉnh cỡ mẫu theo tỷ lệ không trả lời và hệ số thiết kế để tính toán cỡ mẫu cuối cùng

o *Đối với biến định tính:*

Tính toán cỡ mẫu cho nghiên cứu chứng minh hiệu quả 2 thuốc điều trị loãng xương. Tỷ lệ đáp ứng kỳ vọng của thuốc mới và thuốc cũ lần lượt là 40% và 30%. Các nhà nghiên cứu quyết định nếu sự khác biệt <5% thì coi như 2 thuốc có hiệu quả tương đương nhau. Chọn $\alpha = 5\%$ và $1 - \beta = 80\%$. Cỡ mẫu được tính như sau:

$$H = \frac{(|p_1 - p_2|) - d}{\sqrt{p_1(1-p_1) + p_2(1-p_2)}} = \frac{(|0,40 - 0,3|) - 0,05}{\sqrt{0,4(1-0,4) + 0,3(1-0,3)}} = 0,0745$$

$$n = \frac{2(Z_{1-\frac{\alpha}{2}} + Z_{1-\beta})}{H^2} = \frac{2(1,96 + 0,842)}{0,0745^2} = 1009$$

Cỡ mẫu cho mỗi nhóm $n = 1009$

* *Tính toán theo phần mềm HSS 1.0:*

Nghiên cứu tương đương với biến định tính

Tính cỡ mẫu Tính lực thống kê

Alpha (α)	Tỷ lệ không trả lời
0.05	0
$Z_{1-\alpha/2} = 1.96$	Hệ số thiết kế
Lực thống kê ($1-\beta$)	1
0.8	Tỷ số 2 nhóm
$Z_{1-\beta} = 0.84$	1
Tỷ lệ của nhóm can thiệp (p_1)	
0.4	
Tỷ lệ của nhóm đối chứng (p_2)	
0.3	
Ngưỡng khác biệt (d)	
0.05	

Cỡ mẫu nhóm can thiệp

1009

Cỡ mẫu nhóm đối chứng

1009

Ghi chú: Xem thêm phần hiệu chỉnh cỡ mẫu theo tỷ lệ không trả lời và hệ số thiết kế để tính toán cỡ mẫu cuối cùng

5. Nghiên cứu không kém hơn (Non-inferiority trial)

- *Công thức*

Các nghiên cứu không kém hơn thường áp dụng nhằm mục tiêu chứng minh hiệu quả của 1 giải pháp điều trị là không kém hơn (không kém hơn nhiều) so với 1 giải pháp điều trị khác. Các nhà nghiên cứu cần đưa ra ngưỡng (d) và nếu sự khác biệt về hiệu quả điều trị lớn hơn $-d$ thì có thể kết luận là hiệu quả của giải pháp điều trị này là không kém hơn (nhiều) so với giải pháp điều trị khác Lưu ý, đây là kiểm định 1 phía.

- H_0 : Giải pháp điều trị 1 kém hơn giải pháp điều trị 2 ($\mu_1 - \mu_2 \leq -d$)
- H_a : Giải pháp điều trị 1 không kém hơn giải pháp điều trị 2 ($\mu_1 - \mu_2 > -d$)

Kém hơn (Inferiority)	Không kém hơn (Noninferiority)	Tốt hơn (Superiority)
	Tương đương (Equivalence)	

Công thức tính cỡ mẫu của nghiên cứu không kém hơn giống như công thức tính cỡ mẫu của nghiên cứu của nghiên cứu tương đương, ngoại trừ việc dùng $Z_{1-\alpha}=1,645$ (kiểm định một phía) thay cho dùng $Z_{1-\frac{\alpha}{2}} = 1,96$ trong nghiên cứu tương đương (kiểm định 2 phía)

o *Đối với biến định lượng:*

Tính toán cỡ mẫu cho nghiên cứu chứng minh hiệu quả 2 thuốc điều trị loãng xương. Hiệu quả kỳ vọng của thuốc mới và thuốc cũ lần lượt là 7 và 4 điểm. Độ lệch chuẩn của mức tăng là 10 điểm. Các nhà nghiên cứu quyết định nếu sự khác biệt <2 điểm thì coi như 2 thuốc có hiệu quả tương đương nhau. Chọn $\alpha = 5\%$ và $1 - \beta = 80\%$. Cỡ mẫu được tính như sau

$$H = \frac{(|\mu_1 - \mu_2|) - d}{\sigma} = \frac{(|7 - 4|) - 2}{10} = 0,1$$

$$n = \frac{2(Z_{1-\alpha} + Z_{1-\beta})}{H^2} = \frac{2(1,64 + 0,842)}{0,1^2} = 498$$

Cỡ mẫu cho mỗi nhóm $n = 498$

* *Tính toán theo phần mềm HSS 1.0:*

Tính cỡ mẫu Tính lực thống kê

Alpha (α) 0.05 $Z_{1-\alpha} = 1.64$	Tỷ lệ không trả lời 0
Lực thống kê ($1-\beta$) 0.8 $Z_{1-\beta} = 0.84$	Hệ số thiết kế 1
Trung bình của nhóm can thiệp (μ_1) 7	Tỷ số 2 nhóm 1
Trung bình của nhóm đối chứng (μ_2) 4	<p>Cỡ mẫu nhóm can thiệp</p> <p>498</p> <p>Cỡ mẫu nhóm đối chứng</p> <p>498</p>
Ngưỡng khác biệt (d) 2	
Độ lệch chuẩn (σ) 10	

○ *Đối với biến định tính:*

Tính toán cỡ mẫu cho nghiên cứu chứng minh hiệu quả 2 thuốc điều trị loãng xương. Tỷ lệ đáp ứng kỳ vọng của thuốc mới và thuốc cũ lần lượt là 40% và 30%. Các nhà nghiên cứu quyết định nếu sự khác biệt <5% thì coi như 2 thuốc có hiệu quả tương đương nhau. Chọn $\alpha = 5\%$ và $1 - \beta = 80\%$. Cỡ mẫu được tính như sau:

$$H = \frac{(|p_1 - p_2|) - d}{\sqrt{(p_1(1-p_1) + p_2(1-p_2))}} = \frac{(|0,40 - 0,3|) - 0,05}{\sqrt{(0,4(1-0,4) + 0,3(1-0,3))}} = 0,0745$$

$$n = \frac{2(Z_{1-\alpha} + Z_{1-\beta})}{H^2} = \frac{2(1,64 + 0,842)}{0,0745^2} = 896$$

* *Tính toán theo phần mềm HSS 1.0:*

Nghiên cứu không kém hơn với biến định tính

Tính cỡ mẫu Tính lực thống kê

Alpha (α) 0.05	Tỷ lệ không trả lời 0
$Z_{1-\frac{\alpha}{2}} = 1.64$	Hệ số thiết kế 1
Lực thống kê ($1-\beta$) 0.8	Tỷ số 2 nhóm 1
$Z_{1-\beta} = 0.84$	
Tỷ lệ của nhóm can thiệp (p_1) 0.4	
Tỷ lệ của nhóm đối chứng (p_2) 0.3	
Ngưỡng khác biệt (d) 0.05	

Cỡ mẫu nhóm can thiệp

896

Cỡ mẫu nhóm đối chứng

896

6. Nghiên cứu thử nghiệm lâm sàng theo cụm (Cluster randomized design)

- Công thức

Nghiên cứu thử nghiệm lâm sàng theo cụm (CRCT = Cluster Randomize Controlled Trial hay Group Randomized Controlled Trial) là loại hình nghiên cứu can thiệp trong đó toàn bộ các cá thể trong 1 đơn vị hành chính (hay còn gọi là cụm = cluster) được lựa chọn ngẫu nhiên vào nhóm can thiệp hay nhóm chứng (thay vì từng cá thể lựa chọn ngẫu nhiên vào nhóm can thiệp hay nhóm chứng trong các nghiên cứu can thiệp truyền thống). CRCT có sai số cao hơn RCT do sự tương đồng của các cá thể trong từng cụm (ICC = Intraclass Correlation).

$$n = \frac{2\sigma^2 [Z_{1-\frac{\alpha}{2}} + Z_{1-\beta}]^2 VIF}{\gamma\delta^2}$$

Trong đó:

- n là số cụm
- $Z_{1-\frac{\alpha}{2}}$ là giá trị từ phân bố chuẩn, được tính dựa trên xác suất sai lầm loại 1 ($Z_{1-\frac{\alpha}{2}} = 1,96$ nếu xác suất sai lầm loại 1 = 5% và kiểm định 2 phía).

- $Z_{1-\beta}$ là giá trị được tính dựa trên lực thống kê ($Z_{1-\beta} = 0,842$ nếu lực thống kê là 80%)
- σ là độ lệch chuẩn (σ^2 là phương sai)
- γ (Gamma) là số lượng cá thể trong từng cụm
- δ (Delta) là mức khác biệt về hiệu quả điều trị giữa nhóm can thiệp và nhóm chứng
- VIF là hệ số phóng đại phương sai (Variation inflation factor)
- $VIF = 1 + (\gamma - 1) * ICC$
- ICC là hệ số tương quan nội cụm (Intraclass Correlation). ICC được tính dựa trên mô hình hồi quy đa tầng và được hiểu đơn giản là tỷ trọng của sự khác biệt giữa các cụm trong tổng số sự khác biệt giữa các cụm và sự khác biệt nội cụm.

$$ICC = \frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}$$

Trong đó

- σ_u^2 là phương sai sự khác biệt giữa các cụm (tầng 2)
- σ_e^2 là phương sai sự khác biệt giữa các cá thể trong từng cụm (tầng 1).
Nếu các cá thể trong từng cụm càng giống nhau thì σ_e^2 càng nhỏ và ICC càng lớn và cỡ mẫu đòi hỏi càng lớn.

- *Ví dụ*

Tính toán cỡ mẫu cho một nghiên cứu thử nghiệm lâm sàng theo cụm so sánh hiệu quả 2 thuốc điều trị huyết áp. Mức khác biệt về hiệu quả điều trị giữa thuốc mới và thuốc cũ là 10 mmHg. Độ lệch chuẩn gộp của huyết áp là 25, số người tham gia nghiên cứu ở mỗi cụm là 20. Hệ số tương quan nội cụm (ICC) = 0,03. Chọn $\alpha = 5\%$ và $1 - \beta = 80\%$.

$$VIF = 1 + (20 - 1) * 0,03 = 1.57$$

$$n = \frac{2 * 25^2 * (1.96 + 0.84)^2 * 1.57}{20 * 10^2} = 8$$

* *Tính toán theo phần mềm HSS 1.0:*

Nghiên cứu thử nghiệm lâm sàng theo cụm

Tính cỡ mẫu Tính lực thống kê

Độ lệch chuẩn (σ)
25

Alpha (α)
0.05

$Z_{1-\alpha/2} = 1.96$

Lực thống kê ($1-\beta$)
0.8

$Z_{1-\beta} = 0.84$

Số lượng cá thể trong từng cụm (Y)
20

Mức khác biệt về hiệu quả điều trị giữa nhóm bệnh và nhóm chứng (δ)
10

Lựa chọn
 Nhập ICC
 Tính ICC

Hệ số tương quan nội cụm (ICC)
0.03

Kết quả hệ số phóng đại phương sai (VIF) = 1.57

Số cụm
8

7. Cỡ mẫu cho mô hình hồi quy

Đối với cỡ mẫu cho mô hình hồi quy đa biến. Hiện tại có rất nhiều ý kiến khác nhau về cỡ mẫu cần thiết cho mô hình đa biến. Một số tác giả đưa ra quy ước dựa vào số lượng tham số trong mô hình. Một trong những quy ước phổ biến đó là tỉ số cỡ mẫu/số lượng tham số ít nhất từ 10-30, tức là cần ít nhất 10-30 mẫu cho mỗi tham số đưa vào mô hình. Ví dụ một mô hình đa biến có 3 biến số, thì cần ít nhất 30-90 mẫu. Một quy tắc khác của tác giả Green đưa ra năm 1991 là cỡ mẫu cho mô hình đa biến là $N > 50 + 8m$ (với m là số lượng biến độc lập trong mô hình) khi, và cỡ mẫu $N > 104 + m$ khi muốn kiểm tra mức độ quan trọng của biến trong mô hình.

8. Một số hiệu chỉnh

8.1. Hiệu chỉnh đối với quần thể hữu hạn

Các công thức tính cỡ mẫu đều mặc định quá trình lấy mẫu là một quá trình có hoàn trả và kích cỡ dân số đích đủ lớn để coi là lấy mẫu có hoàn trả. Trong trường hợp kích cỡ dân số đích nhỏ, chúng ta có thể hiệu chỉnh cho phù hợp. Việc hiệu chỉnh này giúp cho giảm sự hao phí nguồn lực do phải chọn cỡ mẫu lớn hơn cần thiết mà vẫn giữ được tính đại diện của mẫu. Một số quy ước nếu số lượng

mẫu tính được chiếm hơn 10% kích cỡ dân số đích thì chúng ta cần hiệu chỉnh theo kích cỡ dân số đích theo công thức

$$n_{\text{hiệu chỉnh}} = (n_{\text{tính được}} * n_{\text{dân số đích}}) / (n_{\text{tính được}} + n_{\text{dân số đích}})$$

8.2. Hiệu chỉnh theo hệ số thiết kế (Design effect)

Việc tính cỡ mẫu phải phù hợp với thiết kế nghiên cứu. Các công thức tính cỡ mẫu thường được áp dụng với giả định đối tượng nghiên cứu được chọn một cách ngẫu nhiên. Trong các nghiên cứu lấy mẫu cụm, thường có thêm tác động của yếu tố cụm, ảnh hưởng tới tính ngẫu nhiên của các đối tượng nghiên cứu. Ví dụ nghiên cứu chọn đơn vị cụm là lớp học hoặc xã, huyện thì những học sinh trong cùng một lớp hoặc những người dân trong cùng một xã, một huyện thường có những đặc điểm giống nhau và khác so với học sinh lớp khác hoặc cư dân trong xã, huyện khác. Để giải quyết vấn đề này, đối với các nghiên cứu chọn mẫu cụm, công thức tính cỡ mẫu cần phải được nhân với hệ số thiết kế (design effect).

Tuy nhiên để tính được hệ số thiết kế phù hợp, trước tiên cần hiểu tác động của việc chọn mẫu cụm là như thế nào. Một cách tổng quát, những đối tượng nghiên cứu trong một cụm sẽ có sự tương đồng nhất định so với đối tượng nghiên cứu ở cụm khác. Một chỉ số đo lường mức độ tương đồng của các cá thể trong một cụm là chỉ số tương quan nội cụm – ICC (Intra-class Correlation Coefficient). Khi đó hệ số thiết kế sẽ được tính bằng công thức:

$$DE = 1 + ICC * (k - 1)$$

Trong đó, k là số lượng đối tượng nghiên cứu được chọn trong mỗi cụm.

Có thể thấy rằng, hệ số thiết kế phụ thuộc vào 2 yếu tố: i) mức độ tương quan của các cá thể trong một cụm, được đo lường bằng hệ số ICC và ii) số lượng cá thể được chọn trong mỗi cụm - k. Hệ số ICC có giá trị từ 0 đến 1, với 0 có nghĩa là các cá thể trong cùng một cụm không có tương đồng gì với nhau, hay độc lập hoàn toàn, trường hợp này hệ số thiết kế sẽ bằng 1 và công thức tính cỡ mẫu không có gì thay đổi so với việc chọn mẫu ngẫu nhiên đơn hay ngẫu nhiên hệ thống. Với hệ số ICC bằng 1 có nghĩa là các cá nhân trong cùng 1 cụm là hoàn toàn tương đồng, hoàn toàn giống nhau khi đó hệ số thiết kế chính bằng số lượng

cá thể được chọn trong mỗi cụm. Hệ số ICC có thể tính được trực tiếp từ các nghiên cứu trước thông qua công thức

$$ICC = \sigma^2_{\text{giữa các cụm}} / (\sigma^2_{\text{trong cụm}} + \sigma^2_{\text{giữa các cụm}})$$

Trong đó $\sigma^2_{\text{trong cụm}}$ là phương sai (mức độ giao động) của giá trị đo lường của các cá nhân trong cùng cụm $\sigma^2_{\text{giữa các cụm}}$ là phương sai giữa các cụm với nhau. Các giá trị này có thể được tính thông qua phân tích phương sai.

Tuy nhiên, trong thực tế thường không đủ thông tin để tính trực tiếp ICC từ công thức trên, để dễ dàng tính toán hệ số thiết kế, hệ số ICC được phân chia thành các mức độ: tương quan rất thấp ($0 < ICC \leq 0.01$), tương quan trung bình ($0.01 < ICC \leq 0.03$), tương quan mạnh ($0.03 < ICC \leq 0.05$) và tương quan rất mạnh ($ICC > 0.05$). Ví dụ một nghiên cứu sử dụng kỹ thuật chọn mẫu cụm, và 35 người được chọn từ mỗi cụm ($k = 35$), với giả định rằng các cá nhân trong mỗi cụm có tương quan với nhau ở mức cao, nhà nghiên cứu lựa chọn ICC bằng 0.05, hệ số thiết kế sẽ bằng $1 + 0.05 \cdot (35 - 1) = 2.7$. Như vậy cỡ mẫu ban đầu phải nhân thêm cho 2.7 lần để phù hợp với việc chọn mẫu cụm đối với nghiên cứu này.

Một điều chú ý khác, dựa vào công thức tính hệ số thiết kế, nếu muốn cỡ mẫu nhỏ hơn (DE nhỏ hơn) thì cần giảm số lượng cá thể cần lấy trong mỗi cụm, dẫn tới việc cần phải chọn nhiều cụm hơn. Ngược lại, nếu chọn nhiều cá thể trong mỗi cụm thì số lượng cụm cần lấy sẽ giảm nhưng cỡ mẫu tổng cộng sẽ tăng lên.

8.3. Hiệu chỉnh tỷ lệ không trả lời, bỏ cuộc

Ước lượng cỡ mẫu từ các công thức tính cỡ mẫu sẽ cho chúng ta cỡ mẫu cần có (cuối cùng) trong nghiên cứu. Tuy nhiên, việc tính cỡ mẫu cần phải dự trù cho khả năng đối tượng nghiên cứu từ chối tham gia, hoặc bỏ ngang trong quá trình nghiên cứu để đảm bảo có được cỡ mẫu cuối cùng cần thiết. Tùy vào tình huống cụ thể mà có thể lựa chọn tỉ lệ mất mẫu cho phù hợp. Thông tin này thường được tham khảo từ những nghiên cứu trước đó trên đối tượng nghiên cứu tương tự. Số lượng cỡ mẫu sau khi dự trù mất mẫu được tính bằng công thức:

$$n_{\text{bao gồm dự trù mất mẫu}} = n_{\text{ban đầu}} / (1 - \text{tỉ lệ mất mẫu})$$

Ví dụ nghiên cứu tính ra cần cỡ mẫu bằng 500, với dự trù mất mẫu khoảng 10% thì số đối tượng nghiên cứu cần tiếp cận là $500/0,9 = 556$ mẫu. Chú ý là dự trù mất mẫu là lấy số lượng mẫu ban đầu chia cho $1 -$ tỉ lệ mất mẫu, không phải lấy số mẫu ban đầu cộng thêm 10%, vì khi lấy thêm 10% tức là 50 mẫu, tổng cộng là 550 mẫu, nếu thực sự tỉ lệ mất mẫu là 10% thì cỡ mẫu thu được cuối cùng sau nghiên cứu là $550 - 550*0.1 = 495$ mẫu. Khi đó số lượng mẫu thu được cuối cùng ít hơn 5 mẫu so với cần thiết.

Tỉ lệ mất mẫu cũng cần được suy xét cẩn thận vì nếu chọn tỉ lệ mất mẫu quá thấp sẽ không thu được số lượng mẫu cần thiết cho nghiên cứu. Ngược lại, nếu chọn tỉ lệ mất mẫu quá cao sẽ làm hao phí nguồn lực nghiên cứu do phải tuyển chọn nhiều đối tượng nghiên cứu hơn mức cần thiết.

8.4. Hiệu chỉnh khi cỡ mẫu 2 nhóm không bằng nhau

Trong các nghiên cứu so sánh giữa các nhóm, thường các nhóm sẽ có cỡ mẫu giống nhau. Tuy nhiên, tùy vào hoàn cảnh nghiên cứu mà nhà nghiên cứu có thể thay đổi tỉ số mẫu giữa các nhóm. Ví dụ nghiên cứu thử nghiệm lâm sàng so sánh hiệu quả của một loại thuốc mới so với giả dược thì không nhất thiết phải chọn tỉ số mẫu là 1:1 mà có thể lựa chọn tỉ lệ một ca sử dụng thuốc và nhiều ca đối chứng bằng giả dược. Việc tăng tỉ số mẫu giữa các nhóm làm tăng số lượng mẫu ở nhóm chứng dẫn tới tăng cỡ mẫu chung. Lưu ý rằng các nghiên cứu can thiệp, nếu lấy tỉ số mẫu khác 1:1 thì số lượng mẫu ở nhóm chứng phải lớn hơn nhóm can thiệp (vì tránh ảnh hưởng không mong muốn của can thiệp). Việc tăng tỉ số mẫu ở các

$$N' = \frac{N(1 + k)^2}{4k}$$

nhóm này đặc biệt hữu ích trong nghiên cứu bệnh chứng với khả năng xảy ra bệnh rất thấp, khi đó tăng tỉ số mẫu giữa các nhóm giúp tăng độ chính xác của các phép kiểm định thống kê. Công thức hiệu chỉnh cho tỉ số các nhóm khác nhau như sau:

Trong đó N' là cỡ mẫu hiệu chỉnh, N là cỡ mẫu ban đầu, k là tỉ số giữa 2 nhóm. Ví dụ một nghiên cứu bệnh chứng tính ra được cỡ mẫu ban đầu là 400 (200 ở mỗi nhóm, bệnh và chứng). Tuy nhiên vì khó khăn trong việc tìm người trong nhóm bệnh, nhà nghiên cứu quyết định tăng tỉ lệ giữa các nhóm thành 1 ở nhóm bệnh và 2 ở nhóm chứng. Như vậy cỡ mẫu sau khi hiệu chỉnh là $400 \cdot 9/8 = 450$, trong đó 150 ở nhóm bệnh và 300 ở nhóm chứng.

Dựa vào công thức trên, ta có thể thấy rằng lựa chọn tỉ số cỡ mẫu ở nhóm chứng cũng không nên cao quá 4 lần so với nhóm can thiệp/bệnh, vì khi đó độ chính xác không tăng lên mà sẽ tiêu tốn nguồn lực nghiên cứu.

Tài liệu tham khảo

1. Julious SA (2009), Sample sizes for clinical trials, CRC Press, Boca Raton.
2. Machin D (2009), Sample size tables for clinical studies, Blackwell Science, Oxford.
3. Shein-Chung, Shao J, Wang H (2008). Sample size calculations in clinical research, second edition, Chapman & Hall/CRC, Taylor & Francis Group.
4. Jones SJ, Carley S, Harrison M (2003). An introduction to power and sample size estimation. Emerg Med J; 20:453-458
5. Nguyễn Văn Tuấn: Phương pháp ước tính cỡ mẫu cho một nghiên cứu y học.
6. Maxwell, S. E (2000). Sample size and multiple regression analysis. Psychological Methods, 5(4), 434.
7. Lemeshow S, Hosmer D, Klar J, Lwanga S (1990). Adequacy of Sample Size in Health Studies, World Health Organization, John Wiley & Sons Ltd.

PHỤ LỤC: CHỌN MẪU CHO QUẦN THỂ KHÓ TIẾP CẬN

ThS. Trần Hùng Minh, PGS.TS. Phạm Minh Khuê

1. Giới thiệu

Quần thể nghiên cứu khó tiếp cận (*hard to reach population*) trong lĩnh vực y tế công cộng thường được nhắc đến như: nhóm đối tượng tiêm chích ma túy, phụ nữ mại dâm, người vô gia cư, nam quan hệ tình dục đồng giới hoặc nhóm chuyển giới. Chọn mẫu trong nhóm quần thể “khó tiếp cận” luôn gặp khó khăn do không có khung mẫu (không có danh sách đối tượng của quần thể) đồng thời các nhóm này chiếm tỷ lệ nhỏ trong quần thể chung và người bên ngoài khó có thể “xâm nhập” do nhiều nguyên nhân, trong đó có kỳ thị xã hội. Kỹ thuật chọn mẫu theo chuỗi (*Chain Referral Sampling*- một dạng của chọn mẫu Snowball) giúp cho việc tiếp cận đối tượng tốt hơn thông qua mạng lưới của chính những người trong nhóm, tuy nhiên cách chọn mẫu này vẫn mang tính “thuận tiện” hơn là chọn mẫu xác suất [Leo A. Goodman 2011]. Năm 1997, Heckathorn đã phát triển kỹ thuật chọn mẫu “Dây chuyền có kiểm soát” (Respondent-driven sampling - RDS) để khắc phục hạn chế của chọn mẫu thuận tiện [Heckathorn, D. D. 1997]. Phương pháp này sau đó đã được áp dụng trong nhiều nghiên cứu với các quần thể khác nhau như: người tiêm chích ma túy, mại dâm, nam quan hệ tình dục đồng giới và chuyển giới trên nhiều quốc gia, trong đó có Việt Nam [Family Health International 2005; Lang 2004; Des Jarlais D et al. 2016; Michel L et al 2020]. Thậm chí một nghiên cứu trên đối tượng tiêm chích ma túy tại Việt Nam còn khuyến cáo RDS không chỉ là kỹ thuật chọn mẫu cho nghiên cứu mà rộng hơn có thể là một phương án hiệu quả thu nhận đối tượng cho các can thiệp cộng đồng cho những đối tượng khó tiếp cận này [Des Jarlais D et al. 2016].

Bên cạnh phương pháp RDS, kỹ thuật chọn mẫu theo Thời gian-Địa điểm (*Time Location Sampling*- TLS) là chọn mẫu chùm (*cluster sampling*) được sử dụng cho những quần thể di động (*Floating population*) [PSI, 2007]. Nghiên cứu viên có thể tiếp cận đối tượng tại một số địa điểm mà họ thường xuất hiện, nhưng

không cố định ở một địa điểm, ví dụ như: điem tiêm chích ma túy, điem bắt khách của nhóm mại dâm, tụ điem quán bar.

Mỗi phương pháp sẽ có ưu nhược điem khác nhau đòi hỏi sự lựa chọn phù hợp cho từng loại hình thiết kế. Một nghiên cứu tại Hà Nội và Hải Phòng đã sử dụng cùng lúc hai phương pháp RDS và phương pháp TLS và tiến hành so sánh hai phương pháp này trong việc thu nhận đối tượng nghiên cứu khó tiếp cận là phụ nữ bán dâm tại Việt Nam cho thấy cả hai phương pháp đều có thể được áp dụng tốt trong thu nhận đối tượng này mặc dù phương pháp RDS có thể tuyển chọn được nhiều phụ nữ mại dâm hoạt động ngầm hơn [Johnston LG et al. 2006]. Một nghiên cứu khác tiến hành thu nhận đối tượng nghiên cứu là người tiêm chích ma túy tại Hải Phòng cho thấy tỷ lệ từ chối tham gia là cao hơn khi áp dụng phương pháp RDS nhưng phương pháp TLS đòi hỏi ít nhân lực và chi phí cho thu nhận đối tượng nghiên cứu hơn [Tran HV et al. 2015]

Bài viết này nhằm giới thiệu các nguyên tắc cơ bản và qui trình chọn mẫu theo phương pháp RDS và TLS

2. Chọn mẫu RDS

Kỹ thuật chọn mẫu truyền thống Snowball có sai số chọn mẫu do yếu tố “tự nguyện tham gia” dẫn đến chọn mẫu quá nhiều (*over sampling*) một nhóm nào đó và bỏ sót (*under sampling*) một số nhóm khác của quần thể. Kỹ thuật chọn mẫu RDS khắc phục được hạn chế của phương pháp Snowball thông qua hình thức chi trả “kép”: trả tiền cho đối tượng tham gia nghiên cứu và trả tiền khi họ mời thêm đối tượng tham gia nghiên cứu đạt chất lượng [Heckathorn, D. D 1997]. Đồng thời, đối tượng nghiên cứu cũng chỉ được mời thêm các đối tượng nghiên cứu mới với số lượng hạn chế tùy theo thiết kế của nghiên cứu, thường là mời từ 3-5 đối tượng. Với phương thức tuyển dụng này và qua một số vòng tuyển dụng đủ lớn (5-7 vòng) thì mẫu được chọn vào nghiên cứu sẽ đạt tính đại diện cho quần thể nghiên cứu và không còn phụ thuộc vào đặc tính của những cá nhân được tuyển dụng ở vòng đầu tiên, còn gọi là các hạt giống [Heckathorn 2002, FHI 2003].

Các bước triển khai:

Bước 1: Tính toán cỡ mẫu nghiên cứu

Trong bài viết này không đi sâu vào phương pháp tính cỡ mẫu vì nội dung này đã được trình bày trong bài khác. Tuy nhiên, việc tính cỡ mẫu cho nghiên cứu chọn mẫu RDS cũng không khác biệt so với các nghiên cứu áp dụng kỹ thuật chọn mẫu thông thường khác. Việc sử dụng công thức tính cỡ mẫu nào tùy thuộc vào mục tiêu nghiên cứu và loại biến số chính mà nghiên cứu dự định đo lường, ví dụ: nghiên cứu ước tính giá trị trung bình hay ước tính một tỷ lệ hoặc so sánh 2 tỷ lệ hay so sánh 2 giá trị trung bình. Do chọn mẫu RDS không phải là chọn mẫu ngẫu nhiên đơn nên hệ số thiết kế (*design effect*) sẽ được sử dụng trong tính cỡ mẫu. Sau khi tính được cỡ mẫu nghiên cứu thì chúng ta mới quyết định sẽ chọn bao nhiêu đối tượng hạt giống và sẽ mỗi đối tượng tham gia nghiên cứu sẽ được mời thêm bao nhiêu người khác cùng tham gia nghiên cứu.

Ví dụ: một nghiên cứu xác định nhu cầu chăm sóc sức khỏe của người chuyển giới nữ tại Hà Nội và Tp HCM sử dụng công thức tính cỡ mẫu ước tính tỷ lệ với độ chính xác tuyệt đối và các giả định như sau:

$$n = DE \times \frac{Z_{1-\alpha/2}^2 P(1-P)}{d^2}$$

- $\alpha = 0,05$
 - P (Tỷ lệ sử dụng phương pháp hormon trong nhóm chuyển giới nữ) = 0,49
 - d (độ chính xác tuyệt đối) = 0,1
 - Hệ số thiết kế (*Design effect*) = 2
- ⇒ n = 192 (cho mỗi thành phố)

Bước 2: Thiết kế “Phiếu giới thiệu” đối tượng nghiên cứu (coupon) và sử dụng phần mềm quản lý coupon

Thông thường một phiếu coupon sẽ có 3 liên, liên số 1 sẽ do nghiên cứu viên giữ lại. Liên số 2 và 3 giao cho đối tượng tham gia phỏng vấn để họ mang về mời

thêm người khác tham gia nghiên cứu. Khi mời được người tham nghiên cứu mới thì sẽ giao Liên số 3 cho người mới và Liên số 2 đối tượng giữ lại để quay lại gặp nhóm nghiên cứu lĩnh tiền. Khách hàng mới sẽ mang Liên số 3 đến địa điểm phỏng vấn để đăng ký tham gia nghiên cứu. Nghiên cứu viên sẽ nhập mã số Người giới thiệu và Mã số khách hàng được giới thiệu vào phần mềm quản lý Coupon để quản lý chuỗi, nhằm biết được tiến độ tuyển mộ và thanh toán tiền cho đối tượng nghiên cứu.

Ví dụ minh họa: coupon trong một nghiên cứu nhóm đối tượng chuyển giới nữ

Coupon (mặt 1)

Mã số người giới thiệu:

Mã số KH **được** giới thiệu:

Ngày phát thẻ:

...../...../.....

Thời gian hạn:

.....-.....

Chữ ký của Phòng vấn
viên:

Nhóm nghiên cứu

Giữ lại phần này

Liên 1: PVV giữ lại

Mã số người giới thiệu:

Mã số KH **được** giới thiệu:

**CẢM ƠN BẠN ĐÃ THAM GIA NGHIÊN
CỨU CỦA CHÚNG TÔI!**

Hãy giới thiệu bạn của bạn đến tham gia
nghiên cứu của chúng tôi (xem mặt sau để
biết bạn cần giới thiệu như thế nào)

Địa điểm:

.....
.....
.....

Thời gian hạn:-.....

Điện thoại liên hệ:

Chữ ký của Phòng vấn viên:

Thẻ này CHỈ có giá trị đến:...../...../.....

Liên 2: Người giới thiệu giữ lại để nhận tiền hỗ trợ

Mã số người giới thiệu:

Mã số KH **được** giới thiệu:

**HÃY ĐẾN VÀ THAM GIA NGHIÊN CỨU
CỦA CHÚNG TÔI!**

Địa điểm:

.....
.....
.....
.....
.....

Thời gian hạn:-.....

Điện thoại liên hệ:

Chữ ký của Phòng vấn viên:

Thẻ này CHỈ có giá trị đến:...../...../.....

Liên 3: Đưa cho khách hàng để đến phỏng vấn

Coupon (mặt 2)

Liên 2:

Người được bạn giới thiệu cần:

1. Là Chuyên giới nữ
2. Tuổi ≥ 18
3. Chưa tham gia vào nghiên cứu này

Liên 3:

Hãy đến tham gia nghiên cứu với chúng tôi nếu bạn:

1. Là Chuyên giới nữ
2. Tuổi ≥ 18
3. Chưa tham gia vào nghiên cứu này

Phần mềm quản lý coupon:

Phần mềm thường được sử dụng là RDSCM [Heckathorn, D. D. 1997]. (RDS Coupon Manager) phiên bản tiếng Việt, có thể tải miễn phí từ đường link <http://www.respondentdrivensampling.org/>. [Des Jarlais D et al. 2016]

Bảng 5: Quy trình sử dụng phần mềm

Thời điểm	Công việc	Tính năng cần sử dụng trên phần mềm
Khi 1 người đến phỏng vấn	<ul style="list-style-type: none"> ○ Kiểm tra coupon ○ Hỏi các thông tin sàng lọc (phần đầu trong bộ câu hỏi) 	
	<ul style="list-style-type: none"> ○ Nhập mã số coupon và thông tin sàng lọc vào phần mềm ○ Kiểm tra sự trùng lặp đối tượng* 	Nhập số liệu và kiểm tra trùng lặp đối tượng
Khi hoàn tất phỏng vấn 1 người	Bổ sung các thông tin: <ul style="list-style-type: none"> ○ Có phỏng vấn không 	Tìm kiếm, chỉnh sửa, bổ sung số liệu

	○ Mã số các coupon được giao	
Cuối mỗi ngày	○ Vẽ sơ đồ tuyển người	Tra cứu chuỗi tuyển người
	○ Kiểm tra tình hình phân phát coupon và phỏng vấn	Tình hình phân phát coupon và phỏng vấn
Khi 1 người đến nhận tiền giới thiệu người	○ Kiểm tra coupon được mang đến xem có phải coupon giả không	
	Kiểm tra xem: ○ người có mã số coupon đó đã được Phỏng vấn chưa ○ đã trả tiền cho người giới thiệu chưa	Kiểm tra tình hình trả tiền giới thiệu người

**Kiểm tra trùng lặp đối tượng nghiên cứu:* Phần mềm RDSCM sử dụng một số biến số sinh trắc để sàng lọc trùng lặp đối tượng như: chiều dài khủy tay, chu vi cổ tay và một số nhận dạng đặc biệt trên khuôn mặt (nốt ruồi, sẹo...). Trong bối cảnh công nghệ phát triển rất nhanh, hiện nay việc xác định trùng lặp đối tượng nghiên cứu có thể sử dụng máy quét vân tay. Máy quét vân tay không lưu hình ảnh vân tay mà lưu dữ liệu dưới dạng dãy số duy nhất và không gắn kết với thông tin cá nhân, do vậy không có nguy cơ lộ thông tin của đối tượng nghiên cứu.

Bước 3: Chọn hạt giống

- Xác định các đặc điểm quan trọng của quần thể nghiên cứu mà có thể ảnh hưởng đến biến số nghiên cứu: Nghiên cứu viên cần thảo luận với những đại diện của quần thể đích nhằm xác định xem quần thể đích thường được chia thành các nhóm nhỏ với những đặc điểm gì. Ví dụ, khi nghiên cứu hành vi

nguy cơ nhiễm HIV trong quần thể MSM thì chúng ta cần tìm hiểu xem quần thể MSM tại địa bàn nghiên cứu thường được phân ra những nhóm nào, họ có đặc điểm gì quan trọng có thể liên quan đến hành vi nguy cơ mà nghiên cứu đang quan tâm. Ví dụ, trong quần thể MSM có một số nhóm nhỏ như: nhóm mại dâm nam, nhóm sử dụng thuốc kích thích khi quan hệ tình dục, nhóm tuổi trẻ hoặc lớn tuổi.

- Xây dựng tiêu chuẩn chọn hạt giống: Hạt giống là người thuộc cộng đồng đích và là người có mối quan hệ tương đối rộng trong cộng đồng đích, có khả năng mời được nhiều đối tượng khác tham gia nghiên cứu. Tuy nhiên, tránh tuyển các hạt giống là nhân viên của các tổ chức cộng đồng (CBO) hoặc là cán bộ tuyên truyền, đồng đảng viên của các dự án hoặc đã tham gia nhiều nghiên cứu trong lĩnh vực liên quan. Về mặt lý thuyết, nếu phân cộng đồng đích thành 3 nhóm thì có thể phân bổ hạt giống theo bảng bên dưới. Tuy nhiên, trên thực tế có thể không tìm được hạt giống hoàn toàn đạt tiêu chí như vậy (ví dụ: không tìm được hạt giống ở độ tuổi 30-40 tuổi thuộc nhóm mại dâm) hoặc nhóm đối tượng đó chiếm tỷ lệ nhỏ trong quần thể thì cần phân bổ nhiều hạt giống hơn cho nhóm đông hơn. Như vậy, nguyên tắc cơ bản là: nếu quần thể nghiên cứu được chia làm các nhóm có một số đặc điểm khác nhau thì các đặc điểm đó phải được xuất hiện ở các hạt giống được chọn.

Hành vi	Nhóm tuổi		Cộng
	<i>Nhóm trẻ tuổi</i> (18-30)	<i>Nhóm lớn tuổi</i> (30-40)	
Mại dâm	1	1	2
Dùng thuốc kích thích khi quan hệ tình dục	1	1	2

- Xác định số hạt giống cần tuyển và số coupon cần phát cho mỗi đối tượng: Việc xác định số hạt giống cần tuyển dụng còn phụ thuộc vào cỡ mẫu và số vòng tuyển dụng (wave) mong đợi. Công thức ước tính số hạt giống cần tuyển như sau:

$$a = n/(b^1+b^2+b^3+\dots b^c)$$

Trong đó:

- Số hạt giống = a
- Cỡ mẫu = n
- Số coupon dự kiến phát cho một đối tượng = b
- Số vòng tuyển dụng mong đợi = c

Ví dụ: cỡ mẫu n= 192, số coupon dự kiến phát cho một đối tượng b = 3, số vòng tuyển dụng mong đợi c = 3

$$a = 192/(3^1+3^2+3^3) = 192/39 = 4.9$$

Như vậy, chúng ta có thể tuyển dụng 5 hạt giống với mong đợi là mỗi đối tượng tham gia nghiên cứu có thể tuyển đủ 3 người mới và chuỗi tuyển dụng của mỗi hạt giống sẽ đi đủ cả 3 vòng tuyển dụng.

Có một số quan điểm khác nhau về việc ưu tiên số hạt giống hay ưu tiên số vòng tuyển dụng hay số coupon phát ra cho một đối tượng trong khi chúng ta chỉ có một cỡ mẫu giới hạn. Nếu dùng nhiều hạt giống và phát nhiều coupon cho một đối tượng thì số vòng tuyển dụng sẽ ít và có thể chưa đạt được điểm cân bằng (equilibrium) [Heckathorn, D 2002]. Thực tế triển khai cho thấy, số coupon phát cho mỗi đối tượng nghiên cứu thường là 3 hoặc 4 là phù hợp để giảm thiểu nguy cơ chuỗi tuyển người của một hạt giống nào đó bị dừng lại sớm. Thông thường nghiên cứu viên cần dự trù thêm các hạt giống để thay thế khi chuỗi tuyển người của một hạt giống nào đó bị dừng ở vòng 1 hoặc vòng 2 [Des Jarlais D et al. 2016].

Bước 4: Tổ chức và triển khai chọn đối tượng nghiên cứu trên thực địa

Tùy đặc điểm của đối tượng nghiên cứu mà nghiên cứu viên quyết định chọn một hoặc nhiều địa điểm khác nhau để bố trí nơi phỏng vấn đối tượng nghiên cứu. Ví dụ, với đối tượng nghiên cứu là người tiêm chích ma túy hoặc mại dâm thì nghiên cứu viên cần thực hiện mapping địa bàn, sau đó quyết định chọn bao nhiêu địa điểm, mỗi địa điểm làm trong bao nhiêu ngày để phù hợp với cỡ mẫu, khả năng đi lại của đối tượng và thời gian triển khai nghiên cứu. Đối với đối tượng

nhà nghiên cứu như người chuyên giới nữ, do số lượng không nhiều và không xác định được địa bàn thì có thể chỉ cần chọn một địa điểm sao cho thuận tiện cho các đối tượng nghiên cứu tham gia. Một câu hỏi đặt ra là: có thể triển khai nghiên cứu cùng một lúc trên nhiều địa điểm trong cùng một thành phố không? Hoặc nhà nghiên cứu viên có thể đi đến tận nơi mà đối tượng nghiên cứu hẹn gặp để phỏng vấn không? Việc triển khai nghiên cứu theo hình thức cuốn chiếu (làm xong một địa điểm thì mới chuyển sang điểm tiếp theo) giúp đảm bảo: sàng lọc tránh đối tượng tham gia nhiều lần và quản lý chuỗi tuyển người cũng như chi trả tiền bằng phần mềm trên máy tính.

**Ghi chú:* Trong nghiên cứu chọn mẫu RDS, khi phân tích số liệu cần sử dụng phần mềm phân tích số liệu RDSAT [Des Jarlais D et al. 2016] để hiệu chỉnh số liệu. Nếu dùng phần mềm phân tích thống kê thông dụng như SPSS hoặc STATA thì cần dùng hệ số hiệu chỉnh thông qua các biến số đo lường kích thước mạng lưới khi phỏng vấn đối tượng [Carballo-Die'guez A, et al, 2011].

3. Chọn mẫu TLS

Chọn mẫu TLS còn được biết đến với một số tên gọi khác như: Chọn mẫu Địa điểm-ngày- thời gian (*Venue-Day-Time Sampling*) hoặc Chọn mẫu Thời gian-địa điểm (*Time-space sampling*) được sử dụng từ những năm 1980s trong nghiên cứu các cộng đồng khó tiếp cận [Watters JK, Biernacki P.1989]. Chọn mẫu TLS được coi như là chọn mẫu chùm ngẫu nhiên, trong đó các cá nhân trong quần thể đích có cơ hội được chọn vào nghiên cứu xấp xỉ như nhau do các địa điểm mà họ thường xuất hiện được chọn ngẫu nhiên theo cả biến số: các ngày trong tuần và khoảng thời gian trong một ngày [H. Fisher Raymond, M.P.H 2007].

Nhiều nghiên cứu tại Việt Nam cũng đã sử dụng kỹ thuật chọn mẫu TLS để chọn các loại hình đối tượng nghiên cứu khác nhau vào điều tra của mình như người tiêm chích ma túy tại Hải Phòng [Tran HV et al. 2015] phụ nữ bán dâm tại Hải Phòng [Nguyen TN et al. 2009] và Thành phố Hồ Chí Minh [Johnston LG et

al. 2006] hay cả khách hàng của phụ nữ bán dâm tại Hà Nội và tại Thành phố Hồ Chí Minh [Nadol P et al 2017].

Các bước triển khai:

Bước 1: Khảo sát định tính (*formative assessment*)

Mục đích của hoạt động này là tìm hiểu sâu hơn về quần thể đích để từ đó xây dựng danh sách các địa điểm mà đối tượng đích thường xuyên đến, nơi mà nghiên cứu viên có thể tiếp cận được họ tại đó.

Một số công việc mà nghiên cứu viên cần thực hiện như sau:

- Xác định người cung cấp tin chính: Người cung cấp thông tin chính thường bao gồm 2 nhóm: nhân viên y tế cộng đồng hoặc cán bộ dự án phụ trách địa bàn và đối tượng đích (ví dụ như: phụ nữ mại dâm, người tiêm chích ma túy...). Người cung cấp tin chính nên là người hiểu rõ về nhóm đối tượng đích, đặc biệt là họ có thể cung cấp thông tin về các địa điểm mà đối tượng đích thường đến và làm thế nào để có thể tiếp cận được đối tượng đích một cách thuận lợi nhất.
- Tiến hành phỏng vấn sâu hoặc thảo luận nhóm với người cung cấp tin chính với các câu hỏi:
 - Đối tượng đích có được phân ra thành các nhóm với các đặc điểm như: nhóm tuổi, giới, mức kinh tế, học vấn hoặc hành vi nguy cơ không?
 - Mức độ “ẩn danh” của đối tượng thế nào? Có những kỳ thị xã hội gì hoặc yếu tố pháp lý gì liên quan đến mức độ “ẩn danh” của họ không? Nếu có, thì điều đó ảnh hưởng như thế nào đến khả năng tiếp cận đối tượng của nghiên cứu viên?
 - Đối tượng đích được phân bố thế nào trên địa bàn điều tra: chia thành từng khu vực hay phân tán lẻ tẻ?
 - Đối tượng đích sống, làm việc và giao lưu với nhau thế nào?

- Đối tượng đích thường tập trung, xuất hiện ở những địa điểm nào? Thời gian và tần suất xuất hiện như thế nào?
 - Làm thế nào để hợp tác với những người chủ, người quản lý các địa điểm nêu trên để nghiên cứu viên có thể tiến hành nghiên cứu tại đó?
- Dựa trên kết quả thảo luận nhóm và phỏng vấn sâu, nghiên cứu viên tổng hợp danh sách địa điểm khảo sát và tiến hành quan sát thực địa để bổ sung thông tin chi tiết. Trong quá trình quan sát thực địa, nghiên cứu viên có thể cập nhật thêm danh sách mới chưa có trong bảng danh sách ban đầu.

Ví dụ: Danh sách địa điểm vui chơi giải trí trong nghiên cứu khách hàng nam giới quan hệ tình dục với phụ nữ mại dâm

STT	Loại hình	Tên quán	Địa chỉ	Ngày mở cửa và khung giờ làm việc							Số khách hàng trung bình/buổi
				T2	T3	T4	T5	T6	T7	CN	
1	Bar	Thiên Đường							Trưa	Trưa	100
				Tối	Tối	Tối		Tối	Tối	Tối	150
2	Karaoke	Ngân Hà		Trưa	Trưa	Trưa	Trưa	Trưa	Trưa	Trưa	20
				Tối	Tối	Tối	Tối	Tối	Tối	Tối	50
3	Karaoke	Hải ngoại							Trưa	Trưa	25
				Tối	Tối	Tối	Tối	Tối	Tối	Tối	50

Sau khi có danh sách các địa điểm, nhóm nghiên cứu viên cần thảo luận thống nhất xem có cần loại bỏ những địa điểm nào không phù hợp. Ví dụ: số lượng khách quá ít hoặc chủ quán không hợp tác. Từ đó, nghiên cứu viên chốt danh sách các địa điểm/quán đang hoạt động để đưa vào danh sách chọn mẫu.

Bước 2: Xây dựng khung mẫu cho nghiên cứu và chọn mẫu

- Chọn ngẫu nhiên địa điểm (quán) để khảo sát: tùy theo số lượng quán trên danh sách và cỡ mẫu của nghiên cứu mà nghiên cứu viên quyết định chọn bao nhiêu địa điểm/quán trong danh sách. Về nguyên tắc, nếu nguồn lực cho phép

thì hạn chế số lượng khách hàng điều tra tại một địa điểm và tăng số địa điểm điều tra để đảm bảo tính đại diện. Có thể áp dụng phương pháp điều tra 30 cụm ngẫu nhiên với 7 ca trong 1 cụm của khảo sát tiêm chủng do WHO hướng dẫn. Tuy nhiên, trong một số trường hợp, nghiên cứu viên có thể tăng cỡ mẫu cho từng địa điểm dựa trên số lượng khách hàng trung bình của địa điểm đó.

- Nếu địa điểm/quán được chọn vào danh sách điều tra có nhiều Ngày và nhiều Khung giờ mở cửa, nghiên cứu viên cần đánh mã số “Ngày và Khung giờ” và chọn ngẫu nhiên 1 Khung giờ trong 1 ngày để tiến hành điều tra. Ví dụ: quán Thiên Đường có tổng cộng 8 Khung giờ để điều tra trong 1 tuần (6 buổi tối và 2 buổi trưa). Nghiên cứu viên đánh mã số từ 1 đến 8 cho 8 khung giờ này và chọn ngẫu nhiên một Khung giờ để điều tra. Trong tình huống nhiều địa điểm được chọn vào cùng một khung giờ của một ngày, dẫn đến tình trạng không có đủ nghiên cứu viên đi phỏng vấn thì tiếp tục chọn ngẫu nhiên những khung giờ còn lại của địa điểm bị trùng.
- Nghiên cứu viên nên chọn dư thêm danh sách một số địa điểm và Khung giờ điều tra để đề phòng tình huống không thể tiến hành điều tra tại những điểm đã được chọn vào danh sách trước đó.

Bước 3: Chọn mẫu tại thực địa

Tại mỗi địa điểm nghiên cứu/quán, nghiên cứu viên tiến hành các công việc sau:

- Giới thiệu nhóm nghiên cứu với chủ quán: thông thường, giám sát viên sẽ tiếp cận địa điểm nghiên cứu và trao đổi với người quản lý địa điểm/chủ quán trước khi triển khai. Khi nhóm nghiên cứu viên tập hợp tại địa điểm/quán sẽ được giám sát viên giới thiệu với chủ quán và nhân viên của quán.
- Giám sát viên quan sát địa bàn và cách phân bố chỗ ngồi cho khách ngồi trong quán và phân công nghiên cứu viên phụ trách các khu vực khác nhau của quán sao cho phù hợp với số lượng nghiên cứu viên và lưu lượng khách hàng.

- Đếm/ước lượng tổng số khách trong quán: tùy theo yêu cầu của nghiên cứu mà có những quy định về cách đếm khách hàng khác nhau. Với những nghiên cứu chọn số lượng khách hàng để phỏng vấn bằng nhau tại các quán sẽ cần đếm chính xác hơn số lượng khách để có thể sử dụng cho việc hiệu chỉnh (weight) số liệu khi phân tích [H. Fisher Raymond, M.P.H.2007].
- Tiếp cận khách hàng theo nguyên tắc chọn mẫu hệ thống: khách hàng vào trước sẽ được tiếp cận trước. Trong tình huống, khi nhóm nghiên cứu viên tới quán đã có nhiều khách hàng ngồi trong quán thì mỗi nghiên cứu viên sau khi được phân công quản lý một số bàn sẽ áp dụng quy tắc tiếp cận khách hàng chung, ví dụ: bắt đầu từ bàn số 1 bên tay phải, đi tiếp đến hết các bàn bên tay phải thì sẽ tiếp cận sang các bàn tiếp theo bên tay trái. Trong một bàn, nếu có nhiều khách hàng phù hợp tiêu chuẩn thì cũng áp dụng quy tắc lần lượt từ phải sang trái.
- Sàng lọc khách hàng phù hợp, đạt sự đồng thuận tham gia của khách hàng và tiến hành phỏng vấn. Việc tiếp cận khách hàng và phỏng vấn dừng lại khi đã đạt được cỡ mẫu mong đợi cho nghiên cứu.

KẾT LUẬN

Hai kỹ thuật chọn mẫu được mô tả ở trên được tối ưu hoá giúp cho chọn mẫu nghiên cứu trên các quần thể khó tiếp cận (*hard to reach population*). Mỗi phương pháp sẽ có ưu nhược điểm khác nhau đòi hỏi sự lựa chọn phù hợp cho từng loại hình thiết kế.

TÀI LIỆU THAM KHẢO

1. S. K. Lwanga and S. Lemeshow: SAMPLE SIZE DETERMINATION IN HEALTH STUDIES A Practical Manual World Health Organization 1991
2. Essentials of Biostatistics in Public Health Third Edition Lisa M. Sullivan, PhD, 2018
3. Nguyen Văn Tuấn: Phương pháp ước tính cỡ mẫu cho một nghiên cứu y học.

4. Carballo-Die'guez A, Balan I, Marone R, Pando MA, Dolezal C, et al. 2011. Use of Respondent Driven Sampling (RDS) Generates a Very Diverse Sample of Men Who Have Sex with Men (MSM) in Buenos Aires, Argentina. PLoS ONE 6(11): e27447.
5. Chris LS Coryn et al, 2007. Needs assessment in hidden populations using respondent-driven sampling. Evaluation Journal of Australia, Vol 1.7, No.2, 2007
6. Cyprian Wejnert, Huong Pham, Nevin Krishna, Binh Le, Elizabeth DiNenno 2012. Estimating Design Effect and Calculating Sample Size for Respondent-Driven Sampling Studies of Injection Drug Users in the United States. AIDS Behav (2012) 16:797–806.
7. Des Jarlais D, Duong HT, Pham Minh K, Khuat OH, Nham TT, Arasteh K, Feelemyer J, Heckathorn DD, Peries M, Moles JP, Laureillard D, Nagot N; (The Drive Study Team). Integrated respondent-driven sampling and peer support for persons who inject drugs in Haiphong, Vietnam: a case study with implications for interventions. AIDS Care 2016 Oct;28(10):1312-5
8. Des Jarlais DC, Thi Huong D, Thi Hai Oanh K, Khuê Pham M, Thi Giang H, Thi Tuyet Thanh N, Arasteh K, Feelemyer J, Hammett T, Peries M, Michel L, Vu Hai V, Roustide MJ, Moles JP, Laureillard D, Nagot N; DRIVE Study Team. 2016. Prospects for ending the HIV epidemic among persons who inject drugs in Haiphong, Vietnam. Int J Drug Policy 2016 Jun;32:50-6
9. Family Health International 2005, Family Health International, viewed 3 March 2006, <<http://www.fhi.org/en/index.htm>>.
10. H. Fisher Raymond, M.P.H; Theresa Ick; Michael Grasso; Jason Vaudrey, M.P.H; Willi McFarland, M.D., Ph.D., M.P.H. 2007. Resource guide: Time Location Sampling
11. Heckathorn Douglas D., Christopher Cameron, and Yongren Shi, 2013. Respondent-Driven Sampling: Origins, Current Developments, and Alternative Estimators

12. Heckathorn, D. D. (2007). Extensions of Respondent-Driven Sampling: Analyzing Continuous Variables
13. and Controlling for Differential Recruitment Using Dual-Component Sampling Weights. Center for the Study of Economy and Society, Working Paper #38. Ithica, NY: CSES, Cornell University.
14. Heckathorn, D. (2002). Respondent-driven sampling II: Deriving valid population estimates from chain-referral samples of hidden populations. *Social Problems*, 49(1), 11–34.
15. Heckathorn, D. D. (1997). Respondent-driven sampling: A new approach to the study of hidden populations. *Social problems*. Ewing: University of California Press.
16. Johnston LG, Sabin K, Prybylski D. Update for sampling most-at-risk and hidden populations for HIV biological and behavioral surveillance. *jHASE* 2010, 2(1):2. <http://www.ieph.org/ojs/index.php/jHASE/issue/archive> [DOI: pending | predoi ver.04.09.2010]
17. Johnston LG, Sabin K, Mai TH, Pham TH. Assessment of respondent driven sampling for recruiting female sex workers in two Vietnamese cities: reaching the unseen sex worker. *J Urban Health*. 2006;83(6 Suppl):i16-i28.
18. Kipke M. et al. (2007). Residential status as a risk factor for drug use and HIV risk among young men who have sex with men. *AIDS Behav*. 11:S56-S69.
19. Lang, SS 2004, Cornell news: Cornell sociologist's statistical sampling method to be used by CDC to track HIV-risk behavior among drug users in 25 cities, viewed 3 March 2006, <<http://www.news.cornell.edu/releases/Nov04/CDC.HIV.ssl.html>>.
20. Leo A. Goodman, 2011. Comment: On Respondent-Driven Sampling and Snowball Sampling in Hard-to-Reach Populations and Snowball Sampling Not in Hard-to-Reach Populations.

21. Mansergh, G. et al. (2006). Adaptation of venue day time sampling in southeast Asia to access men who have sex with men for HIV assessment in Bangkok. *Field Methods*. 18 (2) 135-152.
22. Michel L, Nguyen LT, Nguyen AK, et al. Exposure to HIV risks among young people who use drugs (YPUD) in three cities in Vietnam: time to develop targeted interventions. *Harm Reduct J*. 2020;17(1):13. Published 2020 Feb 24.
23. Nadol P, Hoang TV, Le LV, Nguyen TA, Kaldor J, Law M. High HIV Prevalence and Risk Among Male Clients of Female Sex Workers in Hanoi and Ho Chi Minh City, Vietnam. *AIDS Behav*. 2017;21(8):2381-2393.
24. Nguyen NT, Nguyen HT, Trinh HQ, Mills SJ, Detels R. Clients of female sex workers as a bridging population in Vietnam. *AIDS Behav*. 2009;13(5):881-891.
25. Parsons JT. et al. (2006). Differences in club drug use between heterosexual and lesbian / bisexual females. *Addictive Behaviors*. 31: 2344-2349.
26. Path, CCIHP 2018. Transgender women survey in Hanoi and HCM city.
27. PSI 2009. Male clients survey
28. PSI, 2007. Sampling hard to reach populations
29. Rassul Nala´, Beverley Cummings, Roberta Horth, Celso Inguane, Marcos Benedetti, Marcos Chissano, Isabel Sathane, Peter Young, Danilo da Silva, Joy Mirjahangir, Mike Grasso, H. Fisher Raymond, Willi McFarland, Tim Lane. 2015. Men Who Have Sex with Men in Mozambique: Identifying a Hidden Population at High-risk for HIV. *AIDS Behav* (2015) 19:393–404.
30. Tran HV, Le LV, Johnston LG, et al. Sampling Males Who Inject Drugs in Haiphong, Vietnam: Comparison of Time-Location and Respondent-Driven Sampling Methods. *J Urban Health*. 2015;92(4):744-757.
31. Watters JK, Biernacki P, 1989. Targeted sampling: options for the study of hidden populations. *Sociological Problems*, 1989, 36, 416–430.

32. WHO 2011. Guidelines on surveillance among populations most at risk for HIV
33. Wirtz AL, Mehta SH, Latkin C, Zelaya CE, Galai N, Peryshkina A, et al. (2016) Comparison of Respondent Driven Sampling estimators to Determine HIV Prevalence and Population Characteristics among Men Who Have Sex with Men in Moscow, Russia. PLoS ONE 11(6): e0155519.