

Odds Ratio와 Risk Ratio의 차이점

정재영*

동 자료는 2008년 Southern Medical Journal에 등재된 “Odds Ratios and Risk Ratios: What's the Difference and Why Does It Matter?” 을 요약 · 정리한 것으로 Odds Ratio와 Risk Ratio의 차이점을 분석한 내용임

I. 개 요

- 임상의 입장에서는 환자에게 존재하는 리스크 요소 또는 리스크 노출정도(exposure)에 따라 해당 환자에게 특정 질병이 발생할 확률이 얼마나 증가(감소)하는지가 중요함
- 연구원들은 리스크 요소에 노출된 집단과 노출되지 않은 집단을 비교하여 임상의에게 연구결과를 제공
 - 두 집단을 수년간에 걸쳐 관찰한 후, 연구원들은 리스크에 노출된 집단이 그렇지 않은 집단에 비해 얼마나 많이 발병하는지를 추정할 수 있음
 - 이 때의 비율을 Risk Ratio, 즉 “Relative Risk”라고 함
 - 또한 연구의 설계 또는 분석 기술에 따라 Relative Risk가 직접 산출되지 않는 연구도 존재함
 - 이 때 연구원들은 Relative Odds, 즉 “Odds Ratio”를 측정
 - 그러나, 연구 논문 작성자 중에서도 Risk Ratio와 Odds Ratio를 부정확하게 해석하는 경우가 있음
- 본고에서는 Odds Ratio와 Risk Ratio의 의미를 이해 · 해석하고, 이에 따른 차이점을 분석하고자 함

* 생명보험서비스실 담당역

II. Risk Ratio와 Odds Ratio

□ 리스크의 종류

- 일반적인 코호트 연구(cohort study)*에서 연구원들은 두 집단을 리스크 노출여부(예, 흡연)에 따라 정의함
- 두 집단을 추적 관찰하여 각 그룹에서 연구 대상 질병(폐암 등)의 발생 건수를 추적

<자료1> 코호트 연구(흡연과 폐암)

구 분	폐암 발생	폐암 미발생
흡연	a	b
비흡연	c	d

○ 발생률

- 한 집단의 발생자를 해당 집단 전체 대상자로 나눴을 때의 값을 발생률이라고 함 : $\frac{a}{a+b}$
- 발생률은 우리가 흔히 부르는 리스크의 의미로서, 질병 발병 리스크를 말함

○ Risk Ratio(RR)

- 이 때 흡연 집단의 폐암 발생률을 비흡연 집단의 폐암 발생률로 나눈 값을 Risk Ratio 또는 Relative Ratio 라고 함

* 특정 요인에 노출된 집단과 노출되지 않은 집단을 추적하고 연구 대상 질병의 발생률을 비교하여 요인과 질병 발생 관계를 조사하는 연구 방법. 요인 대조 연구(factor-control study)라고도 불림.

$$: RR = \frac{q_{\text{흡연자}}}{q_{\text{비흡연자}}} = \frac{a/(a+b)}{c/(c+d)}$$

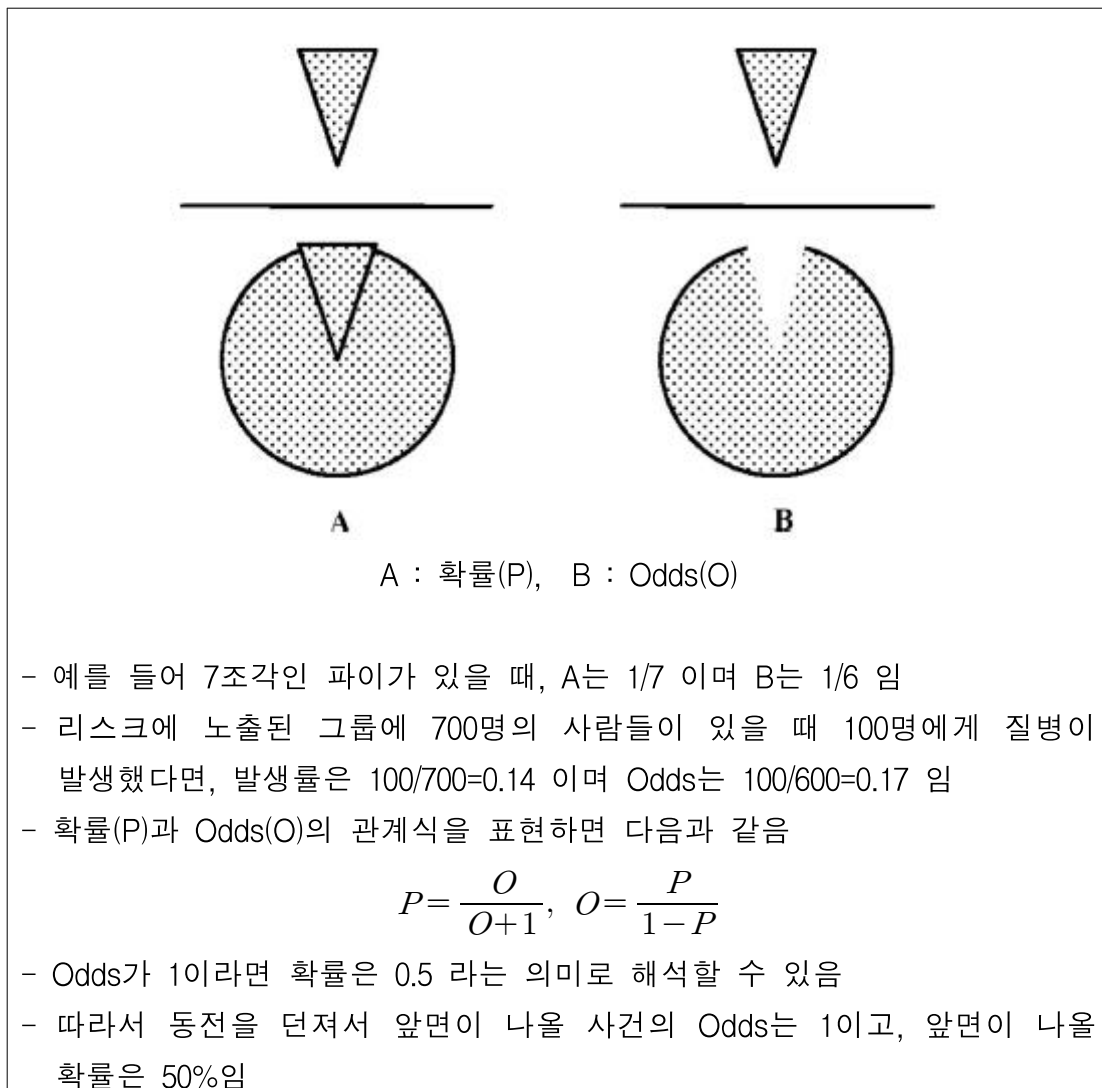
- Relative Ratio는 리스크 노출여부에 따라 연구 대상 질병이 몇 배나 발생하는지를 보여줌
- RR이 1이라면 두 그룹의 리스크가 동일하다고 볼 수 있으며, RR이 1보다 크면 흡연 집단의 리스크가 비흡연 집단에 비해 크다는 것을 의미함
- 예를 들어, 흡연과 폐암간의 코호트 연구에서 RR이 15로 나타났다면 비흡연자에 비해 흡연자의 폐암 발생률은 15배 크다는 의미로 볼 수 있음

○ Odds Ratio(OR)

- 앞서 살펴본 리스크 계산(발생률, Relative Ratio)은 수학적으로 모두 분모가 분자를 포함하는 방식으로서, 이 분수는 모두 확률(어떤 사건의 '발생' 개념)임
- 이때 분모에서 분자를 제외할 경우 해당 분수를 "odds" 라고 부름 : $\frac{a}{b}$
- 따라서 odds는 확률보다 큰 값을 나타냄 : $\frac{a}{b} > \frac{a}{a+b}$
- 집단간 odds 비율을 Odds Ratio라고 함 :

$$OR = \frac{odds_{\text{흡연자}}}{odds_{\text{비흡연자}}} = \frac{a/b}{c/d} = \frac{ad}{bc}$$

<자료2> 확률과 Odds의 차이



□ 코호트 연구에서 Odds Ratio의 의미

- 코호트 연구에서 리스크 노출 집단에서 질병 발생자수를 미발생자수로 나눈 값을 리스크 보유자의 Odds라고 하며, 마찬가지로 리스크 비보유자의 Odds를 구할 수 있음
- 이 때 Odds Ratio는 리스크 보유자의 Odds를 비보유자의 Odds로 나눈 값을 말하는데, 이러한 OR은 리스크 요소에 노출된 사람과 아닌 사람의 질병 발생 위험도를 나타냄

- OR이 1이면 두 집단의 Odds는 같다는 의미이며, OR이 1보다 크다면 리스크 보유 집단의 Odds가 비보유 집단의 Odds보다 크다는 의미임
- 예를 들어, 흡연과 폐암간의 코호트 연구에서 OR이 15라면 흡연자의 폐암 발생 Odds가 비흡연자의 Odds에 비해 15배 높다는 것을 말함

□ 환자군-대조군 연구에서 Odds Ratio의 의미

- 환자군-대조군 연구(case control study)*에서는 발생자 집단과 미발생자 집단을 구분하여 추적 관찰함
- 연구원들은 차트나 인터뷰 등을 통하여 과거를 추적하여 질병 발생전의 잠재적인 리스크 요소 또는 노출여부를 측정함

<자료3> 환자군-대조군 연구

구 분	폐암 발생	폐암 미발생
흡연 이력	a	b
비흡연 이력	c	d

- 본 연구에서는 각 집단에서 질병 발생률을 산출할 수 없음
 - 전체적인 유병률을 알 수 없기 때문
- 그러나 연구원들은 질병 발생자의 리스크 보유 Odds를 계산할 수 있음 : $\frac{a}{c}$

- 또한 미발생자의 리스크 보유 Odds를 계산 : $\frac{b}{d}$

* 어떤 특정 질환이나 문제를 가진 집단(환자군 등으로 불린다)과 그런 질환이나 문제를 가지지 않은 집단(정상군 또는 대조군 등으로 불린다)을 비교하여 질병이나 문제와 연관된 특정한 위험 요소를 밝히는 연구 방법

- 이 때 두 집단의 Odds Ratio는 발생자 집단의 리스크 노출 정도가 얼마나 미발생자 집단에 비해 높은지를 의미함
- 예를 들어, 폐암 발생자 집단과 미발생자 집단간 OR이 15로 나타났다면 폐암 발생자의 흡연 Odds가 미발생자의 흡연 Odds에 비해 15배 높다는 의미로 해석할 수 있음 :

$$OR = \frac{odds_{\text{발생자}}}{odds_{\text{미발생자}}} = \frac{a/c}{b/d} = \frac{ad}{bc}$$

- 환자군-대조군 연구에서의 OR이 코호트 연구에서처럼 리스크 노출 집단의 질병 발생 Odds 비율을 계산한 것은 아니지만, 수학적으로는 동일한 산식을 보임 :

$$OR_{\text{질병발생여부}} = \frac{ad}{bc} = OR_{\text{리스크노출여부}}$$

- 따라서, 환자군-대조군 연구에서의 OR이 15라는 것은 흡연자의 폐암 발생 Odds가 비흡연자의 폐암 발생 Odds에 비해 15배 높다는 것을 의미

III. Risk Ratio와 Odds Ratio의 관계

□ Risk Ratio에 근사되는 Odds Ratio

- 발생가능성이 희박한 사건을 연구하는 데 있어서는 환자군-대조군 연구가 유용함
 - 이러한 희귀 질병에 관한 코호트 연구를 수행하기 위해서는 많은 시간과 비용이 소모되기 때문
 - 따라서 이미 질병이 발병한 환자들을 대상으로 환자군-대조군 연구를 수행하는 것이 보다 효율적
- 다만, 환자군-대조군 연구에서는 Risk Ratio가 직접 계산되지 않는다는 단점이 있음
- 그러나 희박한 사건에 있어서는 Odds Ratio가 Risk Ratio에 대해 매우 근접한 추정치가 될 수 있음
 - 이는 매우 작은 결과값 수치가 계산에 미치는 영향이 미미하기 때문임(RR 계산과정상 분모에 미치는 영향이 작음)
 - 일반적으로 발생 확률이 낮은 사건에 있어서 a와 c는 매우 작은 값을 보일 것이므로, $a/(a+b)$ 는 a/b 에 수렴하며 마찬가지로 $c/(c+d)$ 는 c/d 에 수렴함
 - 따라서, OR은 RR에 근사될 수 있음

$$\text{Odds Ratio} = \frac{A/B}{C/D} \doteq \frac{\frac{A}{A+B}}{\frac{C}{C+D}} = \text{Risk Ratio}$$

□ 일반적으로 사용되는 Odds Ratio

- Risk Ratio는 코호트 연구에서만 직접 계산이 가능하므로 코호트 연구와 환자군-대조군 연구 모두에서 계산이 가능한 Odds Ratio를 보다 널리 사용함
 - 또한 횡단면적 연구(cross-sectional study, 유병률 연구 등)에서도 Odds Ratio는 산출 가능
- 또한 일반적으로 리스크 요소와 발생 결과와의 상관관계를 파악하는 연구 설계상 리스크 요소와 발생 결과 사이의 관계를 왜곡하는 교란 변수(confounder)가 있을 수 있는데, Odds Ratio를 통해 이러한 교란 변수를 조정할 수 있음

□ Odds Ratio의 Risk Ratio 근사의 잠재적인 문제점

- OR이 RR에 근사된다는 가정 하에 OR이 널리 사용되지만, OR은 RR에 비해 결과를 과장할 수 있는 잠재적인 문제점도 있음
 - OR이 3.5라고 하더라도, 리스크 노출 집단이 그렇지 않은 집단에 비해 발병할 위험이 3.5배 많다고 할 수는 없음
 - 다만 OR이 3.5배 많다고 할 수 있음
- 예를 들어, 혈압을 낮추기 위해서는 식습관을 조절해야 한다고 조언을 한 후 실제 식습관을 변화시킨 집단과 그렇지 않은 집단을 비교했을 때
 - 해당 조언이 식습관을 변화시키는 데 유의한 효과가 있었다고 해석할 수 있으나(OR:4.2, 95% 신뢰구간:3.8~4.7), 조언을 들은 집단이 그렇지 않은 집단에 비해 4.2배나 많이 식습관을 변화시켰다고 말할 수는 없음

- 식습관을 변화시킨 집단 중 조언을 들었다고 응답한 사람은 82%였고 그렇지 않은 집단 중 조언을 들었다고 응답한 사람은 51%였음(82%가 51%에 비해 4.2배나 크지 않음)
- OR만을 가지고 해석한다면 조언이 식습관 개선에 현저한 영향을 미치는 것으로 이해할 수 있으나, 이는 일반적인 회상 오류(recall bias)의 효과로서 식습관을 변화시킨 집단에서 조언을 기억하기 쉽고 그렇지 않은 집단에서는 조언을 들었더라도 기억하지 못할 가능성이 높음

□ Risk Ratio의 추정

- 일반적으로 발생할 확률이 희박할 때 OR은 RR에 근사된다고 하였는데 이러한 가정을 '희귀성 가정(rarity assumption)'이라고 함
- 실제 데이터를 통해 사건 발생 빈도를 변화시킴으로써 OR과 RR의 관계를 파악할 수 있음

<자료4> 발생 빈도에 따른 Risk Ratio

Risk Ratio		Odds Ratio				
		0.5	1.5	2.0	2.5	3.0
비 흡연 집단의 폐암 발병 비율	0.001	0.50	1.50	2.00	2.50	2.99
	0.01	0.50	1.49	1.98	2.46	2.94
	0.05	0.51	1.46	1.90	2.33	2.73
	0.1(10%)	0.53	1.43	1.82	2.17	2.50
	0.2	0.56	1.36	1.67	1.92	2.14
	0.3	0.59	1.30	1.54	1.72	1.88
	0.4	0.63	1.25	1.43	1.56	1.67

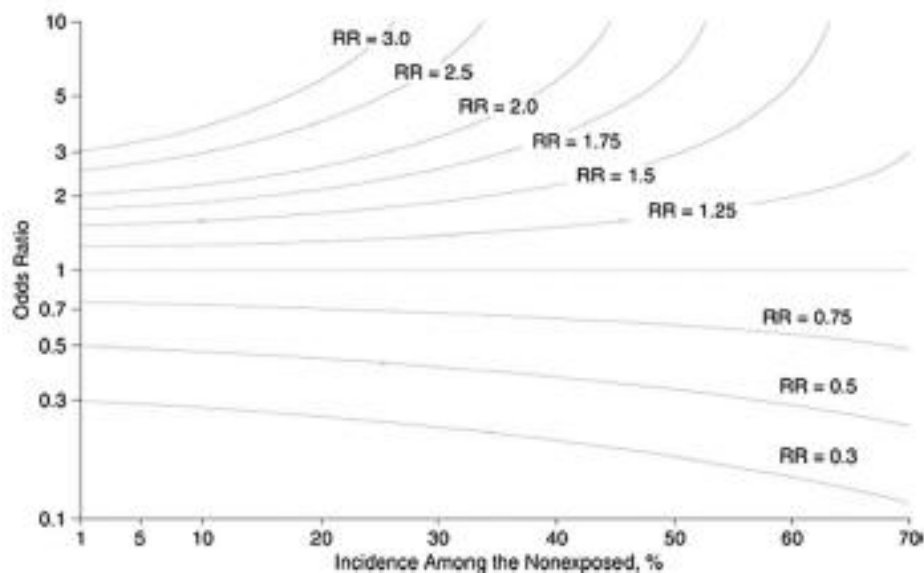
- 일반적으로 리스크에 노출되지 않은 집단에서 발생자 비율이 10% 이하라면 이러한 희귀성 가정을 충족함
- <자료4>에 의하면, OR이 0.5 이하거나 2.5 이상일 때 RR에 비해 OR은 상이한 분포를 보임

○ OR을 통해 RR을 추정하는 방법 중 다음과 같은 산출식이 있음

$$RR = \frac{OR}{(1 - P_0) + (P_0 \times OR)}$$

- P_0 : 리스크에 노출되지 않은 집단의 발생자 비율

<자료5> 발생 빈도에 따른 RR과 OR의 관계



○ 따라서 연구 결과를 해석하기 이전에 앞서 리스크 노출 집단과 그렇지 않은 집단의 발생자 비율을 체크하는 것이 선행되어야 할 것임

□ 효과 크기(Effect Size)

○ 두 변수 사이의 연관성의 정도를 효과 크기라고 함

- OR과 RR은 리스크 노출여부가 결과에 미치는 영향을 설명하는데, RR(OR)이 클수록 효과 크기는 커짐
- 효과 크기가 작을수록 리스크 요소와 결과와의 상관관계를 설명하기가 어려움(통계적으로 유의적이지 않을 수 있음)

- 또한 OR과 RR은 상대적인 척도일 뿐 절대적인 리스크 척도가 아니므로 실제 절대적인 리스크가 작더라도 RR은 크게 나타날 수 있음
- 예를 들어, NSAID라는 약을 5년이상 복용했을 때 유방암을 막을 수 있다는 주장을 할 수 있음(RR은 0.81로 산출됨)
 - 즉, 상대적인 리스크가 19% 정도 감소된다고 주장
 - 그러나 절대적인 리스크 감소량은 0.09%에 불과(NSAID 미복용 집단의 발생자 비율은 0.004, NSAID 복용 집단의 발생자 비율은 0.0049로 나타남)

□ 신뢰구간(Confidence Intervals)

- 리스크 척도는 항상 신뢰구간과 함께 제시됨(보통 95% CI)
 - 신뢰구간은 독자들에게 통계적 유의성과 리스크 척도의 정확성에 대한 정보를 제공함
 - 95% CI가 1을 포함한다면 이는 통계적으로 유의하지 않을 수 있음(1은 실제로 존재할 수 없는 값임)
- 예를 들어, 임상역학 강의를 수강한 레지던트는 그렇지 않은 레지던트에 비해 회진 중에 더 많이 졸음을 느낀다는 주장이 있음(OR : 3.72, 95% CI : 1.02-6.40)
 - 해당 CI는 1을 포함하지 않으므로 OR은 통계적으로 유의하다고 할 수 있음(p-value가 0.05 이하라는 의미)
 - 또한 연구원들이 본 연구를 100회 반복한다면, 95회 이상의 연구에서 OR의 추정치는 1.02와 6.40 사이에 존재할 것임
 - 즉, 임상역학 강의와 회진 중 졸음의 실제 연관성은 1.02만큼 작을 수도 있고, 6.40만큼 클 수도 있음

□ 상대적인 리스크 측정에 있어 기준이 되는 참조 집단

- OR과 RR을 해석할 때 참조 집단(referent group)이 무엇인지 파악하는 것이 필수적임
 - 예를 들어, 임상역학 강의를 들은 집단과 아닌 집단에 관한 연구에서 참조 집단은 강의를 수강하지 않은 집단을 말함
 - 만약 참조 집단을 임상역학 강의를 수강한 집단으로 할 때, OR은 0.27임(기존 OR은 3.72)
 - 즉, 임상역학 강의를 수강하지 않은 집단이 수강한 집단에 비해 0.27배만큼 더 많이 깨어있다는 주장이 됨
- 리스크 노출여부가 이분법적이지 않다면, 참조 집단은 여러 집단 중에서 한 집단이 될 것임
 - 인종 또는 민족은 주요 리스크 요소가 될 수 있는데 이는 3가지 이상의 분류가 가능함(흑인, 백인, 히스패닉, 아시아인 등)
 - 이 때 저자는 반드시 참조 집단을 명시하여 독자들이 OR 또는 RR의 의미를 해석할 수 있게 해야함
 - 참조 집단은 리스크 계수(OR 또는 RR)를 '1.0'으로 표현하기도 함

IV. 요약

- ☐ 논문을 해석하는 데 있어서 Odds Ratio와 Risk Ratio에 대한 차이점을 이해하는 것은 매우 중요함
 - 본고는 Odds Ratio의 취약점에 대한 설명이 아니라 Risk Ratio와의 차이점을 분석한 것임
- ☐ Odds Ratio는 Risk Ratio의 추정치가 될 수 있음
 - Odds Ratio는 집단간 Odds의 비율로서 확률의 비율이 아니지만 희귀성 가정을 충족한다면 Risk Ratio에 근사됨
- ☐ 또한 Odds Ratio 또는 Risk Ratio 해석에 있어서 신뢰구간을 바탕으로 해당 리스크 척도가 통계적으로 유의한 수치인지 명확히 할 필요가 있음