

# t 검정 실험 설계 시 표본 크기 결정에 대한 논의\*

강 정 희<sup>1)</sup> · 방 경 숙<sup>2)</sup> · 고 성 희<sup>1)</sup>

## 서 론

### 연구의 필요성

타당도가 의심스러운 표본은 연구의 외적 타당도를 위협한다. 표본이 모집단을 대표하지 않으면 그 표본으로부터 얻은 연구결과를 모집단에게 일반화시키기 어렵기 때문에, “표본의 타당도 문제를 적게 하려면 무작위로 가능한 한 많은 대상자를 선정하고 표본이 모집단을 대표할 수 있게 하는 표집방법을 이용”(이혜경, 양영희, 구미옥, 은영, 2005, p. 142)하는 것이 해결책으로 제시되기도 한다. 그러나 표본의 타당도, 즉 모집단을 대표하는 정도를 증대시키기 위해 대상자를 가능한 한 많이 표집 하다 보면 연구자가 제1종 오류를 일으킬 위험성도 증가한다. 연구의 외적 타당도를 높이는 것도 중요하지만, 자료 분석 시 제1종 오류를 범한다면 그 결과를 모집단에 일반화시키는 것 자체가 무의미하다. 또한 많은 대상자를 표집 하다 보면 연구자는 필요 이상의 시간과 경비와 노력을 허비할 수도 있다. 그러므로 표본은 너무 크지도 않고 너무 작지도 않게 적절한 수준에서 결정되어야 한다.

검정력 분석은 적절한 표본 크기를 결정하기 위해서 연구자가 연구 계획 단계에서 사용하는 방법으로서, 유의수준, 검정력, 효과크기, 자료 분석 계획 등을 이용하여 표집 대상자수를 결정하는 방법이다(이은옥, 임난영, 박현애, 1998). 유의수준은 특이한 상황이 아니라면 통상적으로 .05로 정한다. 검정력은 귀무가설을 기각할 확률을 의미한다. 검정력을 높게

설정하면 그만큼 대상자도 더 많이 모집해야 한다. 간호학과 같은 행동과학에 적합한 크기의 검정력은 .80정도라고 한다(Cohen, 1988). 효과크기는 두 집단 간 차이를 말한다. 효과크기가 클수록 적은 수의 대상자라도 그 차이를 측정해내기가 용이하다. 보통은 중간 정도의 효과크기를 검정력 분석에 이용하는데, 그 보다는 선행연구에서 효과크기를 계산해내어 적용하는 것이 바람직하다(Cohen, 1988).

최근 문헌에는 검정력 분석으로 설정한 표본 크기에 따라 대상자를 표집 했다고 보고한 실험연구논문이 많다. 이때 검정력 분석 방법으로는 Cohen(1988)의 방법을 가장 많이 이용하고 있다. 그러나 발표된 실험연구논문들의 검정력 분석 과정을 자세히 살펴보면 연구자들이 Cohen(1988)의 검정력 분석방법을 정확하게 사용했다고 보기 어려운 경우가 발견되곤 한다. 연구자들이 제시한 유의수준, 검정력, 효과크기 등으로 표본 크기를 계산해보면 연구자들이 제시한 표본 크기보다 훨씬 큰 수치가 나오곤 한다. 이러한 문제는 Polit과 Sherman(1990)이 예전에 이미 지적했던 문제가 여전히 존재함을 보여 준다. Polit과 Sherman은 Nursing Research와 Research in Nursing and Health에 발표된 62편의 실험연구논문을 분석한 결과, 단 한편의 논문만이 충분한 크기의 표본으로 실험을 진행했다고 보고하였다. 표본 크기 결정 방법을 설명하는 다양한 문헌(이재원, 박미라, 이정복, 이숙자, 박은숙, 박영주, 1998; Cohen, 1988; Devane, Begley, & Clarke, 2004; Kraemer & Thiemann, 1987)과 표본 크기 결정을 도와주는 다양한 컴퓨터프로그램이 있음(Thomas & Krebs, 1997)에도 불구하고

주요어 : 실험 설계, t 검정, 검정력 분석, 표본 크기

\* 본 연구는 전북대학교 간호과학연구소로부터 지원받은 연구임

1) 전북대학교 간호대학 교수, 2) 아주대학교 간호대학 교수(교신저자 E-mail: ksbang@ajou.ac.kr)

투고일: 2009년 4월 27일 심사완료일: 2009년 5월 20일 게재확정일: 2009년 6월 12일

연구자들이 표집 할 대상자 수를 부적절하게 계산해내는 문제가 여전히 지속되고 있는 것이다.

표본 크기의 적절성은 연구 결과의 타당성과도 관련이 있는 중요한 문제이다. 그러나 국내 문헌에서는 검정력 분석, 즉 표본 크기 결정과정이 어느 정도로 적절하게 사용되고 보고되고 있는지를 파악한 연구가 거의 없는 실정이다. 그러므로 본 연구는 검정력 분석 과정 실태를 조사하고 정확한 적용법을 모색하고자 한다.

## 연구 목적

구체적인 연구 목적은, 일 간호학술지에 게재된 실험연구논문에 나타난 검정력 분석 보고 횟수와 보고 내용을 조사하고, 검정력 분석의 적절한 적용에 대해 기술하고, 검정력 분석을 통해 결정한 표본 크기보다 적은 수의 대상자로 진행된 실험 연구 결과 해석 시에 그 연구 결과의 일반화와 관련하여 고려해야 할 사항을 논의하고자 한다.

## 연구 방법

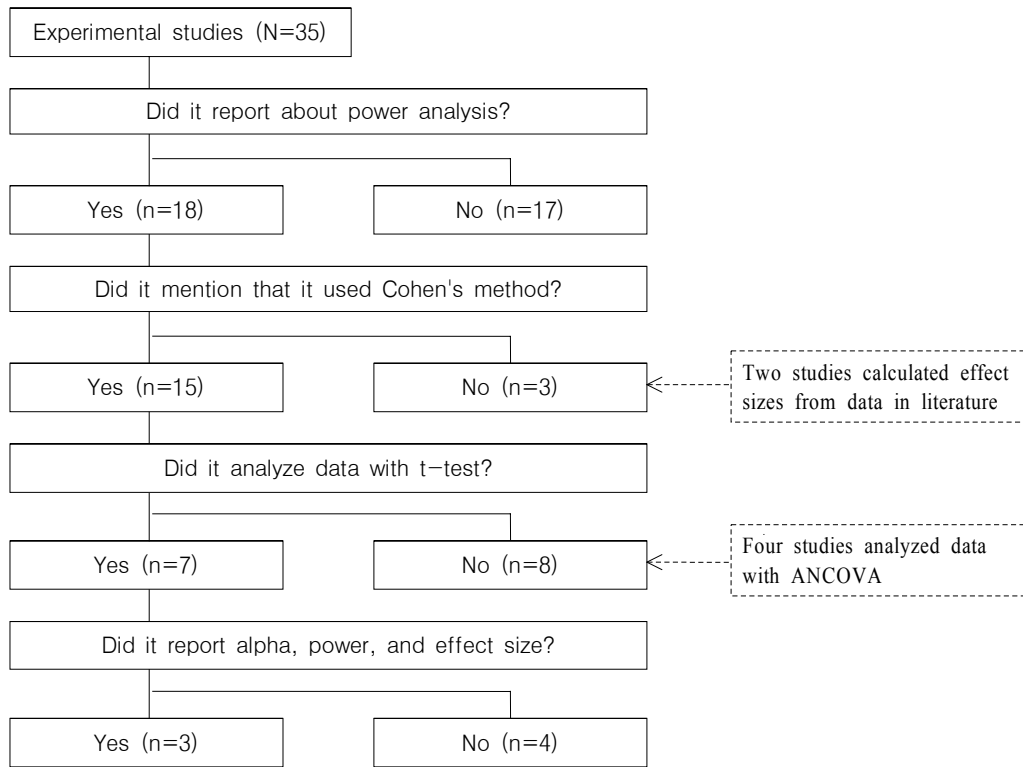
검정력 분석 방법의 사용 및 보고 실태를 파악하기 위한 논문의 출처로 대한간호학회지를 선정하였다. 대한간호학회지

<Table 1> Experimental studies and their power analysis uses and reports

(N=35)

Issue	Author <sup>a</sup>	Power analysis	Alpha	Power	Effect size	Analysis	Sample size <sup>b</sup>
1	Choe	NA	NA	NA	NA	ANOVA	10/group
1	Ahn	NA	NA	NA	NA	paired t test	40 total
1	Kim	NA	.10	.80	1.7 <sup>c</sup>	ANCOVA	15/group
2	Oh	NA	NA	NA	NA	t test	52 total
2	Kim	NA	NA	NA	NA	Mann-Whitney U test	26 total
2	Roh	NA	NA	NA	NA	ANOVA	35 total
2	Song	NA	NA	NA	NA	t test	43 total
3	Lee, Park	NA	.05	.80	.40	repeated measures ANOVA	20/group
3	Sok	Cohen	NA	NA	NA	t test	30/group
3	Lee	NA	NA	NA	NA	t test, repeated measures ANOVA	42/group
4	So	Cohen	.05	.70	.40	repeated measures ANOVA	20/group
4	Sung	Cohen	.05	.70	.40	repeated measures ANOVA	17/group
4	An	Cohen	.05	.70	NA	t test, ANCOVA	25/group
4	Ahn	NA	.05	.80	.80 <sup>c</sup>	ANCOVA	26/group
4	Park	NA	NA	NA	NA	paired t test	90 total
4	Kim, Kim	NA	NA	NA	NA	t test, X <sup>2</sup> test	32 total
4	Kwon	NA	NA	NA	NA	ANCOVA, t test	39 total
4	Kim, Park	Cohen	.05	.80	.60	ANCOVA	12/group
5	Ko	NA	NA	NA	NA	paired t test, ANOVA	90 total
5	Lee	NA	NA	NA	NA	paired t test	37 total
5	Yoo	Cohen	NA	NA	NA	t test & paired t test	16/group
5	Shin	Cohen	.05	.90	.50	t test & paired t test	16-22/group
5	Kim	Cohen	.05	.70	.35	repeated measures ANOVA	26/group
6	Han	NA	NA	NA	NA	two-way repeated ANOVA	32 total
6	Tak	Cohen	.05	.80	.50	ANCOVA	17/group
6	Lee, Shin	Cohen	.05	.70	.40	ANCOVA, t test	20/group
6	Lee, Lim	NA	NA	NA	NA	ANCOVA, t test	84 total
6	Shin	NA	NA	NA	NA	t test	36 total
7	Jo	Cohen	.05	.70	NA	t test	20/group
7	Chu	Cohen	.05	.80	.50	paired t test	17/group
7	Park, Yoon	Cohen	.05	.80	.50	ANCOVA	25/group
7	Park, Kang	Cohen	.05	.80	.63	ANCOVA	19/group
7	Hwang	Cohen	.05	.70	.60	Mann-Whitney U-test	10/group
7	Seo, Kim	NA	NA	NA	NA	paired t test	NA
7	Seo, Suh	NA	NA	NA	NA	repeated measures ANOVA	24/group

Note. NA = not available. <sup>a</sup> 2nd author's last name is presented only when needed to differentiate two articles with the same 1<sup>st</sup> author's last name within the same issue. <sup>b</sup> When the estimated sample size was not reported in an article, the total number of participants was presented in the table. <sup>c</sup> The effect size was calculated from a similar study in the literature.



<Figure 1> Summary of power analysis uses and reports

는 국내에서 가장 많은 회원을 보유하고 있고, 2008년 7월부터는 SCI-expanded와 SSCI 등재지로도 선정되어 국제적으로도 가치를 인정받기 시작한 학술지이다. 그러므로 본 연구는 국제적인 학술지로서의 등재를 받기 바로 전인 2007년에 대한간호학회지에 발표된 실험연구논문을 대상으로 검정력 분석 방법 사용 실태를 조사하였다.

조사자들은 본 논문 저자들로서, 먼저 대한간호학회지 37권 1호부터 7호까지에 게재된 실험연구논문을 각자 파악한 후 서로 일치하였음을 확인하였다. 다음 단계로 조사자들은 분석 대상인 실험연구논문을 각자 읽고 <Table 1>과 <Figure 1>을 작성하였고, 모두 모여서 내용이 일치함을 확인하였다.

## 연구 결과

대한간호학회지 37권에 출판된 총 126편의 논문 중에서, 실험연구논문은 35편(28%)이었다. 실험연구논문을 대상으로 조사한 검정력 분석 방법의 사용 및 보고 실태는 <Table 1>에 자세히 제시되어 있고, <Figure 1>에 요약된 내용이 제시되어 있다. 분석대상인 35편의 실험연구논문 중에서 검정력 분석 내용을 보고한 논문은 18편으로, 전체 실험논문의 51%였다. 검정력 분석에 대해 보고한 실험연구논문 18편 중에서 Cohen의 방법을 사용했다고 기술한 논문은 15편(83%)이었다.

Cohen의 방법을 사용했다고 밝힌 15편의 논문 중에서 t 검정으로 자료를 분석한 논문은 7편(47%)이었고, 나머지 8편 중에서 2편은 ANCOVA를 사용하여 자료를 분석하였다. 자료를 t 검정으로 분석한 논문 중에서 검정력 분석에 필요한 유의수준, 검정력, 효과크기를 모두 보고한 논문은 단 3편(43%)이었다.

또한, 총 35편의 실험논문 중에서 선행연구결과를 토대로 효과 크기를 결정한 논문은 단 두 편 이었다. 그러나 이들 두 편의 논문은 유의수준, 검정력, 효과크기는 보고하였으나 Cohen의 검정력 분석 방법을 사용했다고 구체적으로 밝히지는 않았고 ANCOVA로 자료를 분석하였다.

## 논 의

대한간호학회지 37권 1호부터 7호에 발표된 실험연구논문을 분석한 결과, 검정력 분석 과정을 보고한 논문이 전체 실험연구논문의 반 정도였고, Cohen의 검정력 분석 방법이 가장 많이 쓰이고 있었다. 그러나 본 연구자의 판단에 의하면 분석논문에 발표된 표본 크기가 Cohen이 제시한 것에 비해 적은 경우가 많았다. 그러므로 본 논의에서는 Cohen의 검정력 분석 방법을 적절하게 적용하는 과정을 다루고자 한다. 먼저 부적절하게 시행된 검정력 분석 결과를 예로 들면서 적절하게 적용하는 과정을 설명하고, 실제로 Cohen의 공식에 의

한 표본 크기를 구해볼 것이다. 마지막으로 본 조사에서 밝혀졌듯이 대부분의 실험연구가 충분치 않은 수의 표본을 대상으로 시행되었는데, 결과 해석 시 고려해야 할 사항이 무엇인지 논의할 것이다.

**검정력 분석 방법의 적용**

**부적절한 적용 예:** 이해를 도모하기 위해 검정력 분석 방법을 부적절하게 적용한 예를 들면서 적절한 적용 방법을 모색하고자 한다. 아래에 인용한 부분은 표본 크기 결정 과정이 부적절한 예이다.

표본의 크기는 Cohen(1988)의 Power Analysis법에 의해 효과크기 .60, 유의수준 .05, 검정력 .80로 정하고 Cohen표를 읽어 그룹당 12명으로 확인되었다.

위에 제시된 인용 부분을 포함한 실험연구(이하 ‘인용연구’라 함)는 비동등성 대조군 사전사후 유사실험설계였고, 독립 표본 t 검정으로 자료를 분석하였으며, 가설은 “...가 감소할 것이다”로 지시적이어서 단측검정의 필요성을 시사하였다. 인용연구논문에는 각 그룹에 필요한 대상자 수가 12명이라고 제시되어 있는데, 각 그룹당 표본 크기 12라는 수치에 도달한 과정을 추측해 보면 다음과 같다.

이은옥 등(1998, pp. 669-674)의 저서에 부록으로 첨가되어 있는 Cohen(1988)의 표 일부가 <Table 2>에 제시되어 있다. 이은옥 등이 “두 개 이상의 평균값을 비교할 때 필요한 대상자 수는 Cohen이 제시한 표를 이용할 수 있다”(p. 282)고 기술한 것처럼 유의수준 .05와 자유도 1인 경우에 참고하는 <Table 2>를 보면, 효과크기 .60과 검정력 .80에 해당하는 부분을 찾아보면 한 그룹당 필요한 대상자 수가 12라고 표기되어 있다. 위의 인용연구에서도 동일한 표를 참고하여 12라는 표본 크기를 얻은 듯하다. 그렇지만 이것은 Cohen의 공식을 잘못 적용한 결과이다.

그 이유는 바로 인용연구 연구자가 참고한 표 자체가 인용연구 상황에 부합하지 않는 표이기 때문이다. <Table 2>에 제시한 Cohen의 표는 분산분석으로 자료를 분석할 예정인 실험 연구에 필요한 대상자 수를 구할 때 참고하는 표이다. 이은옥

등도 ‘u’가 “분자의 자유도(그룹의 수 - 1)”(p. 282)를 의미한다고 설명하여, 일원분산분석에서의 사용 가능성을 지적하였다. 그러나 정확하게 말하면 <Table 2>는 다원분산분석인 경우에도 사용할 수 있는 표이다. 다원분산분석 상황에서는 교호작용의 효과를 보는 것이 연구자들의 주요 관심사이기 마련이고, 이 때 필요한 대상자 수 예측을 위해 Cohen의 표를 참고하는 경우 ‘u’는 교호작용의 자유도를 의미한다. 그러므로 ‘u’가 1인 경우는 두 그룹 일원분산분석 상황을 의미할 수도 있고, 독립변수가 두 개 이면서 2\*2의 이원분산분석인 상황을 의미할 수도 있는 것이다.

만약 인용연구 연구자가 ‘u’를 1로 놓고 효과크기와 검정력 등은 동일하게 유지한 채로 <Table 2>를 참고하여 표본 크기를 채택하고자 한다면 이것도 아주 틀린 방법은 아니다. 그러나 그것은 <Table 2>에서 적절하게 효과 크기를 선택할 경우 가능한 것이다. 두 개 이상의 평균값을 비교하는 분산분석 상황에서는, 효과크기를 f라 하고, f값은 아주 커도 .40정도일 뿐이며, 중간 정도의 효과크기는 .25, 작은 효과크기는 .10정도이다. 반면 t 검정 상황에서의 효과크기는 d라 하고, d값은 아주 큰 것이 .80, 중간 크기는 .50, 작은 효과크기는 .20정도이다(Cohen, 1988; 이은옥, 임남영, 박현애, 1998, p. 282에 재인용). 그렇다면 인용연구 상황에 제시된 유의수준 .05와 검정력 .80은 동일하게 유지하고 자유도는 1로 설정하고, 효과크기만 바꾸어서 동일한 표에서 효과크기를 구하여보자. 위 인용연구에서 연구자가 사용한 효과크기 .60은 Cohen이 제시한 t 검정인 경우에 사용하는 중간 정도의 효과크기보다 약간 크다. 일단은 일원분산분석에서의 중간 정도 효과크기인 .25를 사용하여 일원분산분석 시에 참고하는 <Table 2>에서 표본 크기를 구하면 각 그룹당 64명이라는 수치가 얻어진다. 인용연구에 제시된 그룹당 대상자 수 12는 본 연구자가 새롭게 구한 표본 크기 64에 비해, 앞서 얘기한 것처럼, 충분치 않다. 그러나 그룹당 64명이라는 대상자 수도 효과크기에 대한 논의를 하기 위해서 시험적으로 구해본 것일 뿐, 여전히 인용연구 상황에 적절한 표본 크기는 아니다. 바로 t 검정이 지닌 특징 때문이다.

분산분석과는 달리 t 검정은 연구 상황에 따라 적용하는 방

<Table 2> Part of Cohen's table for estimating sample size in F tests when U=1 and alpha=.05

Power	Effect size (f)											
	.05	.10	.15	.20	.25	.30	.35	.40	.50	.60	.70	.80
.10	84	22	10	6	5	4	3	3	2	-	-	-
.50	769	193	86	49	32	22	17	13	9	7	5	4
.70	1235	310	138	78	50	35	26	20	13	10	7	6
.80	1571	393	175	99	64	45	33	26	17	12	9	7

Note. From *Statistical power analysis for the behavioral sciences (2nd ed.)* (p. 384), by J. Cohen, 1988, Lawrence Erlbaum Associates. Copyright 1988 by Lawrence Erlbaum Associates. Adapted and reprinted with permission. U = degree of freedom for numerator of F test.

식이 다양하므로, 그러한 다양성을 고려하여 검정력 분석을 시행해야 한다. 먼저, 연구에 참가한 그룹 수에 따라 적용하는 t 검정이 달라진다. 한 그룹만을 가지고 연구를 진행하는 경우에는, 표본 평균을 이미 알고 있는 모집단 평균에 비교하여 t 검정을 시행하는데, 이는 인용연구에서 채택한 검정방법이 아니다. 두 그룹으로 연구를 진행한다면, 연구에 참여한 대상자들간의 관계에 따라 대응표본 t 검정과 독립표본 t 검정 중에서 선택하여 자료를 분석한다. 이 때 대응표본 t 검정을 실시하는 경우에는 짝이 되는 두 자료간의 관계를 나타내는 상관관계 계수( $\rho$ )를 알아야 검정력 분석에 의한 표본 크기 결정이 가능하다. 마지막으로, 가설을 지시적으로 기술하였느냐 아니면 비지시적으로 기술하였느냐에 따라 t 검정 시 단측 검정과 양측검정 중에서 선택하여 적용한다. 그러므로 표본 크기를 결정할 때에도 이러한 다양한 상황을 고려해야 함은 당연하다. 요약하면, 두 그룹으로 실험연구를 진행하고 t 검정 분석 예정인 연구자가 <Table 2>에 제시된 분산분석 시 참고하는 표에서 표본 크기를 구하려면, 독립표본 t 검정이면서 양측검정으로 두 그룹 간의 차이를 분석하는 상황에서만 가능하다. 단, 그것도 t 검정 시의 효과크기 d를 분산분석 상황에 적절한 효과크기 f로 전환시킬 때에만 가능하다.

**적절한 적용 예:** 앞선 논의에서 부적절한 표본 크기 결정 예를 다루면서, t 검정 시에는 표본 크기를 결정하기 위해 고려하여야 할 사항이 무엇인지도 동시에 다루었다. 기본적으로 유의수준, 검정력, 효과크기를 고려해야 함은 물론, t 검정 상황에서는 그룹 수, 대상자들간의 관계, 가설의 방향성 등을 더 고려해야 한다. 그렇다면, 인용연구 상황에 적절한 표본 크기를 구해보자. 독립 표본 t 검정이고 단측검정이면서 유의수준 .05, 검정력 .80, 효과크기 .60인 상황에 해당하는 Cohen (1988, pp. 54-55)의 표를 찾아보면 각 그룹당 필요한 대상자 수는 35명이 나온다. 독립표본 t 검정 상황에 해당하는 Cohen(1988)의 표는 Munro(2006, p. 144)의 저서에서도 찾아볼 수 있다. 연구자들이 가장 많이 사용하는 설계인 독립표본 t 검정이고 양측검정의 경우에 사용할 수 있는 Cohen의 표 일부를 <Table 3>에 제시하였다. 예를 들어, 유의수준 .05, 검

정력 .80, 효과크기 .60일 때 <Table 3>을 보면 한 그룹당 필요한 대상자 수는 45명이라는 값을 얻게 된다.

Cohen(1988)은 t 검정으로 자료를 분석할 예정인 실험연구에서 표집 할 적절한 대상자 수를 결정할 때 사용하는 다양한 계산식을 제시하였는데 이는 통계전공이 아닌 연구자가 접근하기는 쉽지 않을 수 있다(Polit & Sherman, 1990). 더군다나 일반적으로 찾아 볼 수 있는 Cohen의 표(1988, pp. 54-55)에는 대응표본 t 검정이나 한 그룹만으로 연구를 진행한 후 시행하는 단독표본 t 검정 상황에 따른 표본 크기가 제시되어 있지 않다. 이러한 특수상황이라면 Kirk(2008, pp. 620-621)가 제시한 표를 참고할 수 있다. Kirk는 t 검정에 적절한 효과크기 .20, .50, 그리고 .80(Cohen, 1988)에 따른 표본 크기를 단독표본, 대응표본, 독립표본, 양측검정, 단측검정 등 다양한 t 검정 상황에 따라 미리 계산하여 제시하였다.

**불충분한 크기의 표본으로 진행된 실험 결과 해석**

이론적으로 대상자 수가 충분하지 않으면 영가설을 기각하기가 어려워진다. 그러므로 대부분의 연구자들은 제 1종 오류를 범할 위험을 감수하고서도 되도록이면 대상자를 많이 표집 하려고 노력하게 마련이다. 위 인용연구 연구자는 대상자 수가 불충분하였지만 자료분석 결과 .05 유의수준에서 여러 개의 영가설을 모두 기각하였다. 인용연구 연구자는 본 연구자가 제시한 그룹 당 대상자 수에 비해 훨씬 적은 수인 12명으로도 ‘통계적으로 유의’한 그룹간 차이를 발견한 것이다. 이 결과는 인용연구 연구자가 사용한 중재방법이 효과가 아주 좋은 중재였기 때문이라고 해석할 수 있으나, 이러한 해석은 연구의 내외적 타당도와 관련된 여러 가지 필요조건이 충족될 때에만 내릴 수 있다(Knapp, 1996). 실험중재자, 측정도구, 측정자, 실험 환경 등과 관련된 외생변수 통제가 적절히 되었음을 전제로 한다.

검정력 분석은 원칙적으로 대상자를 무한하게 큰 규모의 모집단으로부터 무작위로 표집하고, 무작위로 표집한 대상자를 실험군과 대조군으로 무작위로 배정함을 전제로 한다. 또

<Table 3> Part of Cohen's table for estimating sample size in two-tailed t tests when two independent groups and alpha=.05

Power	Effect size (d)										
	.10	.20	.30	.40	.50	.60	.70	.80	1.00	1.20	1.40
.25	332	84	38	22	14	10	8	6	5	4	3
.50	769	193	86	49	32	22	17	13	9	7	5
.60	981	246	110	62	40	28	21	16	11	8	6
.70	1235	310	138	78	50	35	26	20	13	10	7
.80	1571	393	175	99	64	45	33	26	17	12	9

Note. From *Statistical power analysis for the behavioral sciences (2nd ed.)* (p. 55), by J. Cohen, 1988, Lawrence Erlbaum Associates. Copyright 1988 by Lawrence Erlbaum Associates. Adapted and reprinted with permission.

한 모든 변수의 측정값이 정규분포하고 두 그룹 간 분산이 동일하다는 가정을 전제로 한다(Knapp, 1996). 그러나 간호계에서 진행되는 실험연구 중에서 이러한 가정을 모두 만족시키는 연구는 거의 없다. 그러므로 통계적 유의성(statistical significance)이 있는 연구결과일지라도 실제적 유의성(practical significance) 혹은 임상적 유의성(clinical significance)을 생각해보기 위해서는 여러 가지 외생변수를 고려해서 해석해야 한다. 반대로 실험연구에 참여한 대상자 수가 충분하지 않았고, 자료분석 결과 영가설을 기각하지 못한 경우에도 마찬가지이다. 통계적 유의성은 없지만 실제적 유의성을 포함할 가능성도 있기 때문이다.

## 결론 및 제언

너무 크지도 않고 너무 작지도 않은 적절한 크기의 표본을 결정하기 위한 검정력 분석 시 연구자들은 흔히 Cohen(1988)의 방식을 적용한다. 그러나 Cohen이 제시한 것보다 적은 수의 대상자로 연구를 진행하는 경우도 적지 않았다. 그 이유에는 t 검정 상황임에도 불구하고 분산분석 시 참고해야 할 표를 참고하여 t 검정 상황에 적절한 효과크기를 적용하여 표본 크기를 결정하는 것도 포함될 것으로 추측된다. 본 논문은 연구자들이 Cohen의 표를 참고하여 대상자 수 예측 시 오류를 범하지 않았으면 하는 바람에서 논지가 전개되었을 뿐, Cohen의 검정력 분석 방법만이 유일하고도 중요한 대상자 수 결정 방법이라고 주장하려는 의도는 없다. 예전에 Knapp(1996)은 지적하기를, 간호계가 다양한 표본 크기 예측 방법을 무시한 채, 검정력 분석만을 지나치게 강조한다고 하였다. 아직도 이러한 분위기가 지속되고 있지는 않은 지, 연구자들의 주의가 요구된다.

## 참고문헌

참고문헌 중에서 첫 저자의 이름 앞에 별표(\*)가 붙은 것은 분석대상 논문임.

- \* 고자경 (2007). 약물요법, 물리치료, 운동요법이 만성요통환자의 통증, 기능장애 및 우울 정도에 미치는 효과의 비교. *대한간호학회지*, 37(5), 645-654.
- \* 권명순, 황경숙 (2007). 운동 프로그램이 초등학교 비만아동의 신체조성, 심폐기능 및 체력에 미치는 영향. *대한간호학회지*, 37(4), 568-575.
- \* 김명희, 김미영 (2007). 구강인두 감각자극이 비위관 삽입 뇌졸중환자의 연하기능에 미치는 효과. *대한간호학회지*, 37(4), 558-567.
- \* 김애숙, 이은숙, 김성효 (2007). 전화를 이용한 지지간호가 함양화학요법을 받는 여성 암 환자의 자가간호수행과 삶의 질에 미치는 효과. *대한간호학회지*, 37(5), 744-753.
- \* 김현경, 정인숙 (2007). 신생아 돌보기 교육프로그램이 산후 조리원 이용 초산모의 양육자신감과 양육행위 정확도에 미치는 효과. *대한간호학회지*, 37(1), 125-134.
- \* 김희자, 박오장 (2007). 향기요법 마사지가 폐경여성의 복부 비만과 신체상에 미치는 영향. *대한간호학회지*, 37(4), 603-612.
- \* 박명화 (2007). 대학생을 위한 온라인 노화와 건강관리 교육의 개발 및 평가. *대한간호학회지*, 37(4), 540-548.
- \* 석소현, 김귀분 (2007). 여성노인의 만성무릎통증, 일상생활 활동 및 생활만족도에 미치는 근육 전기자극 요법의 적용 효과. *대한간호학회지*, 37(3), 305-312.
- \* 성기월 (2007). 건강유지 프로그램이 시설노인의 신체 건강과 정신 건강에 미치는 효과. *대한간호학회지*, 37(4), 478-489.
- \* 소향숙, 김애영, 김은아 (2007). 가려움증 증재 프로그램이 혈액투석환자의 가려움증과 수면에 미치는 효과. *대한간호학회지*, 37(4), 467-477.
- \* 신아미, 박정숙 (2007). 희망증재가 재가 암환자의 희망과 우울에 미치는 효과. *대한간호학회지*, 37(6), 994-1002.
- \* 신윤희, 안수현, 안주립, 양고운, 오숙경 (2007). 청소년의 운동과 식습관 향상을 위한 멘토링 프로그램 개발 및 효과. *대한간호학회지*, 37(5), 703-714.
- \* 안숙희 (2007). 걷기운동이 폐경기 비만여성의 심혈관계 위험요소와 사회심리적 결과에 미치는 효과. *대한간호학회지*, 37(4), 519-528.
- \* 안혜영, 임숙빈, 홍경자, 허명행 (2007). 학령기 아동을 위한 멀티에이전트 비만관리 프로그램의 효과. *대한간호학회지*, 37(1), 105-113.
- \* 안효자 (2007). 병식-지향 간호프로그램이 정신분열병 환자의 병식, 양성-음성증상, 우울에 미치는 효과. *대한간호학회지*, 37(4), 510-518.
- \* 유은광, 전상희, 양정은 (2007). 집단적 지지 간호증재 프로그램이 뇌졸중 환자 가족 수발자의 부담감에 미치는 효과. *대한간호학회지*, 37(5), 693-702.
- \* 이근자, 이어진, 김차남 (2007). 정상체중여성과 비만여성에게 시행한 운동 프로그램의 효과 비교. *대한간호학회지*, 37(5), 684-692.
- \* 이명숙, 임현자, 김경옥 (2007). 재가 독거노인에 대한 건강 다이어리 프로그램의 효과. *대한간호학회지*, 37(6), 966-975.
- \* 이상복, 신혜숙 (2007). 캥거루식 접촉이 미숙아 어머니의 불안, 모성역할자신감 및 모애애착에 미치는 효과. *대한간호학회지*, 37(1), 105-113.

- 호학회지, 37(6), 949-956.
- 이은옥, 임난영, 박현애 (1998). *간호의료연구와 통계분석 (제 3권)*. 서울: 수문사.
- 이재원, 박미라, 이정복, 이숙자, 박은숙, 박영주 (1998). 간호학 연구에서의 표본크기 결정 방법에 대한 고찰. *여성간호간호학회지*, 4(3), 375-387.
- \* 이준화, 박명화 (2007). 농촌지역 노인의 안전한 약물 사용을 위한 교육의 효과. *대한간호학회지*, 37(3), 295-304.
- \* 이창관 (2007). 입원 호흡재활 프로그램이 만성 폐질환자의 호흡곤란, 운동능력과 건강관련 삶의 질에 미치는 효과. *대한간호학회지*, 37(3), 343-352.
- 이혜경, 양영희, 구미옥, 은영 (2005). *간호연구개론 (제 3권)*. 서울: 현문사.
- \* 최명애, 안경주 (2007). 스테로이드치료 전 운동이 스테로이드치료에 의해 유발된 귀의 위축 Type I, II 뒷다리근육에 미치는 효과. *대한간호학회지*, 37(1), 81-90.
- \* 탁영란, 안지연, 김영아, 우혜영 (2007). 학령기 아동을 위한 신체활동-행동수정 비만관리 프로그램이 대사증후군 위험인자에 미치는 효과 분석. *대한간호학회지*, 37(6), 902-913.
- \* 한상숙, 허정자, 김연정 (2007). 세라밴드를 이용한 근육강화운동이 뇌졸중 환자의 하지기능에 미치는 효과. *대한간호학회지*, 37(6), 844-854.
- \* Chu, S. K., Yoo, J. H., & Lee, C. Y. (2007). The effects of a cognitive behavior program on cognition, depression, and activities of daily living in elderly with cognitive impairment. *J Korean Acad Nurs*, 37(7), 1049-1060.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. US: Lawrence Erlbaum Associates.
- Devane, D., Begley, C. M., & Clarke, M. (2004). How many do I need? Basic principles of sample size estimation. *J Adv Nurs*, 47(3), 297-302.
- \* Hwang, E. H. (2007). Effects of a hand acupuncture therapy on sleep quality in sleep disrupted adults. *J Korean Acad Nurs*, 37(7), 1108-1118.
- \* Jo, K. H. (2007). Effects of a nurse presence program on suicide prevention for elders with a chronic disease. *J Korean Acad Nurs*, 37(7), 1027-1038.
- \* Kim, M. S., Kim, J. S., Jung, I. S., Kim, Y. H., & Kim, H. J. (2007). The effectiveness of the error reporting promoting program on the nursing error incidence rate in Korean operating rooms. *J Korean Acad Nurs*, 37(2), 185-191.
- Kirk, R. E. (2008). *Statistics: An introduction (5th ed.)*. Belmont, CA: Thomson Wadsworth.
- Knapp, T. R. (1996). The overemphasis on power analysis. *Nurs Res*, 45(6), 379-391.
- Kraemer, H. C., & Thiemann, S. (1987). *How many subjects? Statistical power analysis in research*. Newbury Park, CA: Sage Publications, Inc.
- Munro, B. H. (2006). *Statistical methods for health care research* (이은현, 정영혜, 김진선, 송라운, 황규윤 옮김). 서울: 군자출판사. (Original work published in 2005).
- \* Oh, H. E., Lee, Y. S., Shim, M. J., & Kim, J. S. (2007). Effects of a postpartum back pain relief program for Korean Women. *J Korean Acad Nurs*, 37(2), 163-170.
- \* Park, H., & Yoon, H. (2007). A comparison of the effects of intravenous fluid warming and skin surface warming on peri-operative body temperature and acid base balance of elderly patients with abdominal surgery. *J Korean Acad Nurs*, 37(7), 1061-1072.
- \* Park, M. K., Kang, K. S., & Kim, N. Y. (2007). Effects of a smoking cessation program on amount of smoking and nicotine dependence and self-efficacy of smoking cessation for smoking workers. *J Korean Acad Nurs*, 37(7), 1073-1079.
- Polit, D. F., & Sherman, R. E. (1990). Statistical power in nursing research. *Nurs Res*, 39, 365-369.
- \* Roh, Y. S., Cho, H., Oh, J. O., & Yoon, C. J. (2007). Effects of skin rehabilitation massage therapy on pruritus, skin status, and depression in burn survivors. *J Korean Acad Nurs*, 37(2), 221-226.
- \* Seo, N. S., Kim, Y. H., & Kang, H. Y. (2007). The effects of a group smoking cessation program among adult smokers in a rural community. *J Korean Acad Nurs*, 37(7), 1139-1148.
- \* Seo, Y. M., & Suh, S. L. (2007). The effects of a weight control program with competence. *J Korean Acad Nurs*, 37(7), 1177-1183.
- \* Song, R., Lee, E. O., Lam, P., & Bae, S. C. (2007). Effects of a sun-style Tai Chi exercise on arthritic symptoms, motivation and the performance of health behaviors in women with osteoarthritis. *J Korean Acad Nurs*, 37(2), 249-256.
- Thomas, L., & Krebs, C. J. (1997). A review of statistical power analysis software. *Bulletin of the Ecological Society of America*, 78(2), 126-138.

## Power Analysis in Experimental Designs with t test Analysis

Kang, Jeong Hee<sup>1)</sup> · Bang, Kyung Sook<sup>2)</sup> · Ko, Sung Hee<sup>1)</sup>

1) Professor, College of Nursing, Chonbuk National University

2) Professor, College of Nursing, Ajou University

**Purpose:** In the literature power analysis in experimental studies is often executed and reported falsely. This descriptive study was done to promote the correct application of the Cohen(1988)'s power analysis method.

**Method:** Articles of experimental studies from a nursing journal were selected and reviewed to examine the uses of and the reports on the power analysis process. Also, the appropriate power analysis process was discussed with an example of the most common experimental design, an independent two-group design with t test analysis plan.

**Result:** Around half the experimental studies examined reported that they carried out power analysis. Cohen's method was the most frequently utilized but with accuracy in question. **Conclusion:** Power analysis to estimate sample size is the interplay between alpha, power, and effect size, and other factors in the case of t test analysis. Researchers should have a clear understanding of how to apply the Cohen's power analysis method so they do not produce poorly estimated sample sizes.

**Key words :** Experimental design, t test, power analysis, sample size

• Address reprint requests to : Bang, Kyung Sook  
College of Nursing, Ajou University  
San 5, Woncheon-dong, Youngtong-gu, Suwon, Kyunggi-do 443-721, Korea  
Tel: 82-31-219-7016 Fax: 82-31-219-7020 E-mail: ksbang@ajou.ac.kr