

■ 특 강

통계분석결과의 제시방법

박 용 규

가톨릭대학교 의과대학 의학통계학교실

초 록

논문을 구성하는 각 부분 즉, 요약, 서론, 방법, 결과, 고찰 중에서 통계적인 사고와 표현방법이 가장 많이 요구되는 부분은, “방법”과 “결과”라고 할 수 있다. 이 두 부분은 연구의 결론을 이끌어내기 위한 방법이 얼마나 적절하고 합리적인가를 보여주고, 유도된 결과를 얼마나 정확하고 효과적으로 제시했는지를 평가하는 기준이 되어, 논문 전체의 신빙성과 질을 좌우하기 때문이다. 이러한 관점에서 “방법”과 “결과”부분을 기술하는 방법과 주의해야할 점들을 소개하며, 아울러 그 외 연구계획단계 및 논문작성과정 전반에 걸쳐 염두에 두어야 할 통계적인 문제들을 다루고자한다.

논문을 구성하는 각 부분 즉, 요약, 서론, 방법, 결과, 고찰 중에서 통계적인 사고와 표현방법이 가장 많이 요구되는 부분은, “방법”과 “결과”라고 할 수 있다. 논문에서의 이 두 부분은 연구의 결론을 이끌어내기 위한 방법이 얼마나 적절하고 합리적인가를 보여주고, 유도된 결과를 얼마나 정확하고 효과적으로 제시했는지를 평가하는 기준이 되어, 논문 전체의 신빙성과 질을 좌우하기 때문이다. 이러한 관점에서 “방법”과 “결과”부분을 기술하는 방법과 주의해야할 점들을 소개하며, 아울러 그 외 연구계획단계 및 논문작성과정 전반에 걸쳐 염두에 두어야 할 통계적인 문제들을 다루고자한다.¹⁻³⁾

1. 연구계획단계 및 논문작성 전반에 걸쳐 고려해야할 점

가. 모집단과 표본

자신의 주장을 증명하기 위해 어떤 대상을 조사할 것인지를 결정해야 한다. 모집단이란 연구자가 관심을 가진 대상 전체를, 표본은 모집단 중 일부분을 말한다.

1) 전수조사

모집단 자체를 조사하는 것을 전수조사(census)라고 하며, 흔히 인구통계조사와 같은 정부기관에서 실시하는 대규모의 조사를 떠올리게 되지만, 연구자가 소규모 특정집단에만 관심을 가지고 있고, 그 집단 전체를 조사한 경우도 전수조사에 해당한다.

2) 표본조사

모집단의 일부만을 조사하는 것을 표본조사(sample survey)라고 한다.

3) 통계적 추론

연구대상이 모집단 또는 표본에 해당하는지를 구분

해야하는 이유는, 대상에 따라 통계적 추론 (statistical inference), 특히 가설검정을 할 필요가 없을 수 있기 때문이다. 가설검정이란 표본에서 얻은 부분적인 증거를, 모집단에도 확장 적용할 수 있는지를 결정하는 과정을 말한다. 만약 모집단을 조사했다면, 연구자는 자신이 관심을 가졌던 대상에 대해 모두 알았기 때문에 더 이상의 추론이 필요없다. 따라서 가설검정의 결과인 P값을 제시하지 않는다.

4) 확률표본

연구대상이 표본이라면 이 표본은 모집단을 가능한 잘 반영하는 것이라야 한다. 또한 표본의 대표성이 보장되어야만 그로부터 도출된 결과도 객관성을 인정 받을 수 있다. 편향된 (biased) 표본을 얻지 않는 가장 좋은 방법은 확률표본을 뽑는 것 (random sampling) 으로, 흔히 난수표나 컴퓨터의 난수발생 프로그램을 사용한다.

나. 실험 (연구) 방법

연구수행과정에서 부딪치는 현실적인 제약조건들을 고려하면서, 연구자가 원하는 결과를 얻어낼 수 있는 가장 효율적인 연구방법을 찾는 문제로, 통계학에서는 실험계획 (experimental design) 의 분야에 속한다. 많은 경우에 시간, 비용 등이 제약조건이 되며, 이에 따라 연구대상의 수 (표본의 크기) 도 결정된다. 표본의 크기가 이미 정해졌다면, 동일한 표본수로 실험의 효과를 가장 잘 발견해 낼 수 있는 최적 (optimal) 의 실험계획방법이 무엇인지를 찾아야 한다.

1) 표본의 크기

보다 충분한 증거를 확보할 수 있다는 관점에서 보면 대상자의 수는 많을수록 좋을 것이다. 두 집단의 평균을 비교할 때 분산이 일정한 값이라면

가) 두 표본평균값이 차이가 많이 나거나,

나) 평균의 차이가 일정한 경우에는 표본수가 많을수록

서로 차이가 난다고 주장할 수 있는 가능성이 높아진다. 나) 의 경우를 검정력 (power) 이 높다고 한다.

그러나 표본의 수는 현실적인 제약조건들로 인해 무한정 늘릴 수 없을 뿐 아니라, 필요한 수 이상의 대

상자들을 조사할 필요도 없다. 특히 그 실험이 환자들에게 상당한 불편이나 고통을 주는 것이라면, 적정 수 이상의 환자들을 실험대상으로 하는 행위는 비윤리적이라고까지 할 수 있다.

2) 실험계획

사전에 특정 연구를 하려는 계획이 없는 상태에서 얻어진 자료, 즉 의무기록실에 보관된 환자들의 차트나 컴퓨터에 입력되어있는 특정 데이터베이스의 자료를 이용하는 후향적 관찰연구 (retrospective observational studies) 에서는 실험계획을 특별히 생각할 필요가 없지만, 전향적 (prospective) 관찰연구나 실험연구의 경우에는, 자신의 연구목적을 이루기 위한 최선의 실험방법을 찾아야 한다.

가) 짝짓기

연구외적인 요인이 결과에 미치는 영향을 최소화하기 위해 자주 사용되는 방법으로, 예를 들어 두 가지 처리간의 차이를 비교하기 위해서, 같은 대상 또는 결과에 영향을 미치리라 예상되는 요인들의 값이 가장 비슷한 대상들을 짝을 짓고, 두 가지 처리 중 하나를 각각에 적용하는 것을 말한다. 이 방법은 대상들을 두 가지 집단에 랜덤하게 할당하는 완전 랜덤배치법에 비해, 처리효과를 보다 쉽게 밝힐 수 있다는 장점이 있다.

세 가지 이상의 처리를 비교하기 위한 짝짓기 방법을 블록화 (blocking) 라고 하며, 비슷한 특성값을 기준으로 연구대상을 집단적으로 분류하는 방법을 층화 (stratification) 라고 한다.

나) 요인실험법 (Factorial design)

특정 집단에 속한 대상들에게 두 가지 이상의 요인의 처리를 동시에 실시하여, 전체 표본의 수를 줄이고, 각 요인에 의한 차이와 요인들간의 교호작용도 분석할 수 있는 실험방법이다.

다. 통계적 유의성과 임상적 중요성

표본을 대상으로한 연구의 결론은, 검정을 한 후 통계적 유의성 (statistical significance) 을 P값으로 표현한다. 그리고 P값은 표본수와 직접적인 관련성을

갖는다. 만약 표본의 수가 매우 크다면, 연구자는 어떤 것이라도 통계적으로 증명할 수 있고, 반면에 표본 수가 매우 작다면, 연구자는 아무 것도 (통계적으로) 증명할 수 없게 된다.

만약 두 집단간의 작은 차이가 표본이 크기 때문에 통계적으로 유의하게 나타났다면, 그리고 특히 두 값이 모두 임상적으로 정상범위내에 속한다면, 두 집단간의 통계적으로 유의한(의미가 있는) 차이가 임상적인 관점에서는 무의미한 것이 된다. 이와 반대로 두 집단간의 큰 차이가 표본이 작기 때문에 통계적으로 유의하지 않게 나타났다면, 비록 객관적으로 증명되지는 않았지만, 임상적으로는 의미있는 결과로 받아들여질 수도 있다. 통계적으로 무의미(non significance)한 결과는 “차이가 없는 것을 증명한 것”이 아니라, “차이가 있는 것을 증명하지 못한 것”으로 해석해야 한다.

통계적 검정은 의학 연구자들이 자신의 주장을 객관적으로 증명하기 위해 사용하는 수단에 지나지 않는다. 즉, 임상적 중요성(clinical importance)이 통계적 유의성에 우선한다. 따라서 통계분석에 들어가기 전에, 임상적으로 의미를 가질 수 있는 차이 또는 관련성의 크기를 미리 설정할 필요가 있다.

2. “방법”부분에 제시해야 할 내용

연구대상, 연구방법, 통계분석방법으로 구성되는 “방법(Materials and Methods)”부분에는, 연구결과를 얻기 위해 사용된 방법이 적절하고 객관적이라고 인정받을 수 있도록, 또한 다른 연구자들이 쉽게 재현해 볼 수 있을 정도로 명확하고 완벽하게 서술되어야 한다.

가. 연구대상과 방법

1) 연구대상과 관찰(분석)단위

관찰단위란 연구자가 실제로 그 반응을 측정하게 되는 구체적인 대상을 말한다. 연구대상과 관찰단위는 대부분의 경우 서로 동일하지만, 환자의 수와 관찰단위인 안구의 수, 치아의 수처럼 서로 다를 수도 있다. 이처럼 한 연구대상으로부터 여러 개의 관찰값이 얻어진다면, 관련된 자료의 분석방법이 사용되어

야한다.

2) 모집단

연구결과가 적용될 수 있는 대상, 즉 모집단에 대한 인구학적, 진단적, 예후적 요인에 대해 서술을 하여, “서론”부분에서 언급된 연구의 목적에 부합하는 연구대상이 선정되었는지, 즉 대표성을 가진 표본을 사용했는지 판단할 수 있어야 한다.

3) 변수의 척도

연구에 사용될 모든 설명변수와 반응변수의 측정단위들을 구체적으로 제시하고, 추후 분석과정에서 변수변환을 하거나 범주화하였다면, 그 이유와 타당성에 대한 설명이 있어야 한다.

4) 임상적 중요성

임상적으로 중요하다고 간주되는 최소의 차이나 변화를 미리 설정하여, 연구에 필요한 표본의 수를 결정하고, 결과를 서술할 때에는 통계적 유의성보다 임상적 중요성에 초점을 맞추도록 한다.

5) 연구계획(Study designs)

연구가 최소한 다음 중 어디에 속하는지 기술하도록 한다.

- 관찰연구, 실험연구: 관찰연구는 인위적인 개입 또는 중재가 없는 경우를 말한다.

- 후향적 연구, 전향적 연구: 후향적 연구란 연구주체가 결정되기 전에 이미 자료가 수집되어 있는 경우를 말한다.

- 확률화 (또는 대조) 임상실험: 실험군과 대조군에 연구대상들이 랜덤하게 할당(random allocation)되는 전향적 연구의 특별한 형태이다.

참고로 연구계획을 서술할 때에는 다음과 같은 용어들이 주로 사용된다.

관찰연구(Observational studies)

- 사례연구(Case report, Descriptive studies)

- 후향적 사례-대조 연구(Retrospective case-control studies)

- 횡단면적 연구(Cross-sectional studies)

- 전향적 코호트 연구(Prospective cohort studies)

- 후향적 코호트 연구(Retrospective cohort stud-

ies)

- 시계열적 또는 반복측정 연구 (Longitudinal studies)

실험연구 (Experimental studies)

- 대조실험 (Controlled trials)

- 동시 대조집단 (Parallel or concurrent control groups)

- 자기 대조집단 (Sequential controls including self-controls, cross-over trials)

- 과거 대조집단 (Historical controls or external controls)

- 비대조 연구 (Studies without control groups)

6) 짝짓기

만약 짝짓기가 사용되었으면, 짝짓기의 기준과 방법 (사용된 변수와 값의 범위)을 구체적으로 설명하고, 짝짓기를 한 이유에 대해서도 적어야 한다.

7) 혼동변수

주요 설명변수와 반응변수간의 관련성을 왜곡시키는 혼동변수들을 제시하고, 이들의 영향을 보정하기 위해 사용된 방법을 서술한다. 혼동변수의 보정은 실험계획단계에서 짝짓기나 랜덤할당과 같은 방법으로 미리 처리하거나, 자료가 이미 수집된 경우라면 다중회귀분석, 다중 로지스틱분석, 그리고 공분산분석과 같은 통계적인 방법을 통해 할 수 있다.

8) 표본크기의 결정

대부분의 후향적 연구에서는, 정해진 기간동안 특정 질환으로 방문한 환자들의 수처럼, 연구자의 의도와 무관하게 결정되지만, 전향적 연구 특히 임상시험의 경우에는 연구시작 전에 임상적으로 의미를 갖는 일정 수준의 효과를 발견해낼 수 있는 검정력을 미리 정하고, 표본크기를 계산한다.

9) 표본추출방법

표본으로 선정된 연구대상들이 모집단에 대한 대표성을 가지고, 통계검정이 요구하는 가정을 만족하기 위해서는, 랜덤화 방법에 의한 표본추출이 필요하다. 흔히 사용되는 랜덤화 추출방법에는 가) 단순랜덤추출 (simple random sampling), 나) 층화랜덤추출

(stratified random sampling) 등이 있다.

10) 참여기준 및 제외기준

모집단의 성격에 대한 보다 구체적인 설명이 되며, 참여기준 (inclusion criteria)은 보통 특정 질환이나 나이처럼 연구자가 관심을 갖는 설명변수에 의해 결정되고, 제외기준 (exclusion criteria)은 임신, 약물의 복용상태와 같이 연구자가 피하고자하는 잠재적 혼동 변수들에 의해 결정된다.

11) 대상의 할당방법

특정 연구대상자를 실험군 또는 대조군에 할당한 방법을 설명하고, 과거 (historic) 대조군이나 병원 (hospital) 대조군을 사용하였다면 그 이유를 명시한다. 여기서 병원 대조군이란 정상인이 아닌 해당 연구질환과 무관한 질환으로 입원해 있는 환자로 구성된 대조군을 말한다.

12) 다기관 공동연구

다기관 (multi-center) 공동연구의 경우, 각 기관마다 장비, 숙련도, 검사방법, 기준치 등이 달라질 수 있으므로, 분석하기 전에 미리 조정할 필요가 있다. 그 방법을 기술한다.

나. 통계분석방법

1) 가설과 통계검정방법

연구자가 밝히고자 하는 내용과 해당 검정방법을 기술한다.

2) 검정의 반복

동일한 자료에 대해 여러 번의 검정을 반복했을 때에는, P값을 조정하기 위한 방법을 기술한다.

3) 유의수준과 검정력

연구자가 (차이가 있다는) 잘못된 주장을 할 확률을 나타내는 유의수준을 이용하여, 연구자가 내린 결론에 대한 신뢰의 정도를 제시하며, 검정력은 부정적 발견 (negative finding)의 적절성을 지지해 주는 간접적인 척도가 된다.

4) 단측검정, 양측검정

“크다”, “작다”와 같이 차이의 방향을 구체적으로 검정하는 것을 단측검정이라 하고, “다르다”, “차이가 있다”처럼 정해진 방향이 없는 경우를 양측검정이라 한다. 차이의 방향에 대한 충분한 이유와 확신이 없는 경우에는 양측검정을 사용하는 것이 보통이다.

5) 통계패키지

널리 사용되고있는 통계패키지의 경우에는 분석결과의 정확성을 인정받고 있지만, 개별적인 패키지나 프로그램은 잘못된 결과를 제시하는 경우가 가끔 있다.

3. “결과”부분에 제시해야 할 내용

가. 기초값 (baseline values)

여기서의 기초값이란, 특정 처리 전, 후 비교를 할 때, 처리 전의 반응값 뿐 아니라, 각 대상들이 갖는 인구학적 변수들, 즉 성별, 나이, 체중 등을 말한다. 이 값들은 특정 처리에 의한 효과를 판정하는데 혼란을 줄 수 있으므로, 본문의 글 또는 표를 통해 제시하여야 한다. 그러나 이 값들이 비교집단간에 차이가 있는지를 반드시 검정할 필요는 없다.

나. 측정치의 일치도

방사선 사진이나 심전도와 같이 판정자의 주관적 판단에 의해 결정되는 반응값은, 여러 명의 판정자들의 판정결과를 종합하여 사용하되, 판정자들간의 일치도를 피어슨의 상관계수나 카파 (kappa) 등을 이용해 제시해야 한다.

다. 결과 제시

1) 가장 중요도가 높은 결과부터 차례로 제시한다.

2) 변화량이나 변화율을 연구한 경우에는 원래 값과 변화를 나타내는 값을 함께 제시한다.

3) 자료의 변이를 나타내는 표준편차와 추정값의

정확도를 나타내는 표준오차는 구분되어야 한다. 표준편차를 제시하는 것이 일반적이지만, 최근의 각 저널은 95% 신뢰구간을 선호하는 추세이다.

4) 본문에서는 결과가 통계적 유의성 여부에 관계없이 가능한 정확한 P값을 제시한다. 표에서도 지나치게 복잡해지지 않는 경우에는 정확한 P값을 사용한다. 유의수준을 0.05로 정했을 때, 계산된 P값이 0.049 또는 0.051이었다면, 단순히 $P < 0.05$ 또는 $P > 0.05$ 로 제시하는 것에 비해, 독자에게 유의성의 정도에 대한 정보를 줄 수 있다. 또한 통계적으로 유의하지 않은 결과를 얻었을 때에는 “거의 유의한 경향을 보였다”, “경계선상의 유의함을 보였다” 등의 표현을 사용하지 않아야 한다. 이 경우 “유의 (significant) 하다”는 일반적으로 통계 검정의 결과를 나타내는 표현이므로, 굳이 설명하고자 한다면 임상적인 관점에서 결과의 중요성을 강조하는 것이 좋을 것이다.

5) 연구의 주요 결과는 표나 그림을 이용하여 제시한다. 표에 있는 내용을 그림으로 중복표현하지 않아야 하며, 본문을 읽지 않더라도 표나 그림의 내용이 충분히 이해될 수 있을 만큼, 정확하고 자세하게 표현되어야 한다. 또한 본문에서는 표나 그림에 있는 수치를 그대로 반복하지 않고, 중요한 점만 간결하게 서술한다.

6) 결과제시가 지나치게 복잡해지지 않는다면, 추정값, 표준오차, 검정통계량의 값을 구체적으로 제시하여, 추후의 연구나 메타분석에서 연구결과를 이용할 수 있도록 한다.

7) 처리와 관련된 부작용이 있을 경우에는, 비록 그것이 전체 연구결과에 불리하게 작용한다하더라도 반드시 서술해야 한다.

8) 이상값 (outliers)이 측정 또는 자료입력의 잘못으로 인해 발생했다면, 그 값들을 제외하고 분석해도 별다른 문제가 없지만, 측정 가능한 값일 때 무조건 제외시키면 왜곡된 결과를 얻게 된다. 후자의 경우 이상값을 포함한 분석결과와 제외시킨 결과를 함께 제시하는 것도 좋은 방법이 된다.

9) 결측치(missing values)는 특정 처리의 부작용으로 인해 나타날 수도 있으므로, 분석에서 제외된 대상자들의 특성을 상세히 서술하여, 연구결과가 편향된 대상으로부터 얻어진 것이 아니라는 점을 강조해야 한다.

4. 통계분석결과와 제시방법

가. 자료의 요약

1) 유효숫자

정확성을 강조하기 위해 필요이상으로 자세한 숫자를 제시하지 말아야 한다. 숫자의 유효 자리수는 2자리로 충분하다.

2) 백분율

백분율만 제시하지 말고, 해당 빈도수까지 제시한다. 이 때 표본수가 100이 넘으면 소수점아래 한자리까지 백분율을 제시하고, 20보다 작으면 백분율 대신 빈도수를 제시한다.

3) 범주형 자료

백분율을 제시하되 대상수도 함께 나타낸다. 만약 연속형 자료가 절단점(cut-off points)에 의해 나누어졌다면, 절단점을 제시하고 그 값이 절단점으로 선택된 이유를 밝힌다.

4) 연속형 자료

적절한 대표값과 변이를 사용해야한다. 정규분포를 하는 자료는 평균과 표준편차, 비정규형 자료는 중앙값과 범위 (또는 사분위수 범위)가 적절하다. 또한 표준오차(SE 또는 SEM)는 추정값의 정확성을 나타내는 척도이므로 자료의 요약으로는 적절하지 않다. 이 때 평균은 원 자료보다 소수점아래 자리수를 하나 더, 표준편차는 두 자리를 더 나타내는 것이 일반적이다. 정규분포를 하는 여러 집단자료의 변이를 비교할 때에는, 측정단위에 영향을 받는 분산 대신 변동계수를 사용한다.

5) 짝지워진 자료

짝짓기에 의해 얻어진 자료는 각각의 값과 변화량

을 차이 또는 비율로 함께 제시한다.

6) 변환된 자료

변수변환을 한 후 분석을 했다면, 변환의 이유와 적절성을 설명하고, 결과해석을 할 때에는 원래의 척도로 환원해서 설명한다.

7) 작은 표본

표본수가 작을 때에는 평균이나 비율 등을 통해 자료의 요약을 하는 대신 원 자료를 그대로 제시한다.

나. 통계분석결과와 제시 예

1) 독립인 두 집단의 t 검정

가) 본문 서술

“실험군의 비만지수는 평균 23.4 ± 3.37 kg/m²(표준편차)으로, 대조군의 평균 22.0 ± 3.39 kg/m²과 유의한 차이가 없었다. (t=1.68, df=73, P=0.09)”

- 정규분포를 하는 반응값이므로, 평균±표준편차를 이용하여 대표값과 변이를 제시하였다.

- 간혹 평균±표준오차를 제시하는 경우가 있기 때문에, 괄호 안에 ±기호 다음의 값이 표준편차임을 나타내었다.

- 통계적 유의성에 관해 서술하였다.

- 검정통계량과 자유도, 그리고 정확한 P값을 제시하였다.

“실험군의 비만지수는 평균 23.4 kg/m²(95% 신뢰구간, 22.4 - 24.4 kg/m²)으로, 대조군의 평균 22.0 kg/m²(95% 신뢰구간, 20.7 - 23.3 kg/m²)과 유의한 차이가 없었다. (t=1.68, df=73, P=0.09)”

- ±표준편차 대신 95% 신뢰구간을 사용하였다. 최근에는 단순히 표준편차를 제시하는 것보다 95% 신뢰구간을 선호하는 경향이 있다.

나) 표 작성

- 표의 제목은 연구내용을 적절히 반영하되, 지나치게 통계적인 표현은 피한다.

- 측정의 단위 (kg/m²)와 제시된 숫자 (Mean±SD)가 무엇을 나타내는지를 표의 제목에 나타내었다.

- 각 집단의 표본수 (N)를 제시하였다.

Table 1. Difference of BMI(kg/m²) between two groups (Mean±SD)

Group	N	BMI	t (df)	P - value
Control	30	23.4±3.37	1.68 (73)	0.09
Experimental	45	22.0±3.39		

- 분석에 사용된 검정방법과 분석결과를, t(df), P 값으로 나타내었다. 중앙값 14 mm(사분위수 범위, 15 - 15 mm)였으며, 대조군은 중앙값 25 mm(사분위수 범위, 23 - 27

Table 2. General characteristics of each group

Characteristic	Control (N=30)	Experimental (N=45)
Age (years) *	45.6±5.45	54.7±4.96
BMI (kg/m ²)	23.4±3.37	22.0±3.39
...

* P<0.05, by two-sample t-test

- 만약 여러 가지 비교결과를 하나의 표에 실어야 한다면, 표 안에는 평균±표준편차를 제시하고, 주석 기호와 설명을 통해 검정방법과 결과를 제시할 수 있다.

mm)로 유의한 차이가 있었다. (S=16.5, P=0.005)”

- 이 예에서는 변이는 범위 대신 사분위수 (interquartile) 범위를, 검정통계량은 U 대신 S를 사용하였다.

2) 윌콕슨의 순위합 검정

나) 표 제시⁴⁾

Table 3. Distance between pole calyx and spine in each group (Median, minimum-maximum)

	Disease (N=5)	Control (N=9)	U	P - value*
Distance (mm)	14 (10 - 22)	25 (18 - 34)	1.5	0.0005

* by Wilcoxon's rank sum test

가) 본문 서술

“X선 필름상에서 측정된 우측상부 pole calyx와 척추사이의 거리는, 신장질환이 있는 어린이의 경우에는 중앙값 14 mm, 범위 10 - 22 mm였으며, 대조군은 중앙값 25 mm, 범위 18 - 34 mm로 유의한 차이가 있었다. (U=1.5, P=0.005)”

- 표안에 제시된 값, 중앙값(최소 - 최대값)을 제 목부분에서 설명하였다.

- 표 안에 표본수 (N), 측정단위 (mm), 검정결과 (통계량, P값)를 제시하였다.

- 주석기호를 이용하여 검정의 이름을 설명하였다.

- 반응값이 정규분포를 하지 않기 때문에 중앙값과 범위, 그리고 순위합 검정 (또는 Mann-Whitney U검정)의 결과를 제시하였다.

3) Paired t-test

“X선 필름상에서 측정된 우측상부 pole calyx와 척추사이의 거리는 신장질환이 있는 어린이의 경우에는

가) 본문 서술

“자극을 주기 전 수축기혈압은 평균 128.7 mmHg (±6.93 mmHg, 표준편차)였으나, 자극 후에는 131.1 mmHg (±6.72 mmHg)로, 평균 2.42 mmHg만큼 유의

박용규: 통계분석결과와 제시방법

하게 증가하였다. (t=2.46, df=11, P=0.314)”
 - 자극 전, 후의 평균과 표준편차, 그리고 차이값의 평균을 제시하였다.

나) 표 제시⁴⁾

5) 분산분석

가) 본문 서술

“제 2 피로군과 제 3 피로군의 스트레스 강도는 각각 10.4±2.3, 10.5±2.3 (평균±표준편차)으로, 정상

Table 4. Systolic blood pressure(mmHg) before and after stimulus (N=12, Mean±SD)

	Before	After	Difference	t (df)	P - value*
SBP	128.7±6.93	131.1±6.72	2.4±3.40	2.46 (11)	0.0314

* by paired t-test

- 제목부분에서 측정단위 (mmHg), 표본수 (N), 표 안에 제시된 수치 (Mean±SD)를 설명하였다.
 - 주석기호를 이용하여 검정방법을 설명하였다.

4) 윌콕슨 부호순위 검정

가) 본문 서술

“진통제 투여 전 토끼의 혈당수준의 중앙값은 125.5 mg/kg (120-134 mg/kg, 사분위수 범위)였으나, 투여 후에는 136.5 mg/kg (131-171 mg/kg)으로 유의하게 증가하였다 (S=85.5, P=0.0007)”

- 대표값과 변이로 중앙값과 사분위수 범위를 사용하였다.

나) 표 작성

군 (8.3±1.8), 제 1 피로군 (8.7±1.7)에 비해 유의하게 높았다. (Scheffe의 다중비교, P<0.05)”

나) 표 작성⁵⁾

- 제목부분에서 표 안에 제시된 수치 (Mean±SD)를 설명하였다.

- 표 안에 표본수 (N)를 제시하였고, 주석기호를 이용하여 분산분석과 다중비교의 결과를 설명하였다.

6) 상관분석

가) 본문 작성

“나이가 많을수록 요추 (r=-0.405)와 Ward 삼각부 (r=-0.479)의 골밀도는 낮아졌고, 체중이 증가할수록 요추 (r=0.505), 대퇴골부 (r=0.452), 대퇴전자부 (r

Table 5. Blood sugar levels(mg/kg) before and after treatment (N=18, Median(Interquartile range))

	Before	After	% of change	P - value*
Blood sugar	125.5 (120 - 134)	136.5 (131 - 171)	8.76	0.0007

* by Wilcoxon's signed rank test

- 제목부분에서 측정단위 (mm/kg), 표본수 (N), 표 안에 제시된 수치 (Median, Interquartile range)를 설명하였다.

- 전후 변화율 (% of change)을 추가로 제시하였다.

- 주석기호를 이용하여 검정방법을 설명하였다.

=0.538), 대퇴체부 (r=0.453)의 골밀도는 높아졌으며, 체질량지수가 증가하면 모든 부위의 골밀도가 유의하게 높아졌다. (P<0.05) 칼슘과 알콜 섭취량과 골밀도와는 유의한 관련성이 없었다”

- 두 변수들간의 관련성을 구체적으로 설명하였다. (~할수록 ~하다)

Table 6. Stress levels of fatigue groups

(Mean ± SD)

	Normal (N=90)	Fatigue			P-value*
		Group 1 (N=44)	Group 2 (N=50)	Group 3 (N=34)	
Stress	8.3 ± 1.8	8.7 ± 1.7	10.4 ± 2.5 ^{†‡}	10.5 ± 2.3 ^{†‡}	0.01

* P-value for one-way ANOVA

† P < 0.05, compared with normal by Scheffe's test

‡ P < 0.05, compared with Fatigue group 1 by Scheffe's test

- Pearson의 상관계수가 너무 많은 관계로 일부만 제시하였다.

요인을 이용하여 전체 산소 소비량의 변동을 81% 정도 설명할 수 있었다”

- 회귀식을 제시하고, 각 회귀계수가 뜻하는 내용을 설명하였다.

나) 표 작성⁶⁾

Table 7. Relation between regional BMD(g/cm²) and physical and diet characteristics (Pearson's correlation coefficient, N=28)

	LS	Bone FN	Mineral FW	Density FT	FS
Age (years)	-0.405*	-0.311	-0.479**	-0.266	-0.215
Body weight (kg)	0.505**	0.452*	0.269	0.538**	0.453*
Body mass index (kg/m ²)	0.417*	0.406**	0.389*	0.647**	0.589**
Calcium intake (mg/day)	0.127	0.164	0.089	0.180	0.019
Alcohol intake (g/day)	0.288	0.171	0.224	0.341	0.338

LS: lumbar spine, FN: femoral neck, FW: Ward's triangle, FT: trochanter, FS: femoral shaft.

* P < 0.05, ** P < 0.01, by Pearson's correlation analysis

- 제목부분에서 BMD의 측정단위, 표 안에 제시된 수치 (correlation coefficient) 와 표본수 (N) 를 설명하였다.

- 회귀모형의 적절성을 설명계수 (R²=0.81) 를 이용하여 설명하였다.

- 표 안에 각 변수들의 측정단위를 제시하였다.

나) 표 작성⁷⁾

- 주석기호를 이용하여, 두 가지 종류의 유의수준을 구분하였다.

- 표 제목에 통계적인 용어의 사용을 피하고자 하였다.

7) 회귀분석

- 각 변수의 측정단위를 제시하였다.

- 회귀계수, 표준오차, P값, 부분결정계수를 제시하였다.

가) 본문 서술

- 주석을 통해 분석방법과 전체 모형의 설명력을 제시하였다.

“달리는 시간이 오래 걸릴수록, 나이가 많을수록, 그리고 달리는 동안의 심박동수가 높을수록, 운동 중 산소 소비량이 유의하게 적어지는 것으로 나타났으며, (산소 소비량 = 111.72 - 2.88 × 달리는 시간 - 0.26 × 나이 - 0.13 × 심박동수, P < 0.001) 이 세 가지

8) 로지스틱 회귀분석

가) 본문 서술

Table 8. Factors associated with oxygen consumption

(ml/kg/minute)*

Step	Factor	Coefficient	SE	P	Partial R ²
0	Intercept	111.72	10.24		
1	Run time (minutes)	-2.83	0.34	0.0001	0.74
2	Age (years)	-0.26	0.10	0.0129	0.02
3	Run pulse (per minute)	-0.13	0.05	0.0154	0.05

Run time: time to run 1.5 miles

Run pulse: heart rate while running

* by stepwise multiple regression analysis, model R² = 0.81

Table 9. Relationship between weight and high(>220) or low(≤220) serum levels (mg/dL)

Variable	Coefficient	SE	P-value	OR	95% CI
Intercept	-1.89	0.48			
Weight (kg)	0.44	0.11	0.001	1.55	1.25 - 1.93

“체중이 1 kg 증가할수록 혈청농도가 높은 집단 (>220 mg/dL)에 속할 odds ratio가 1.55배 (95% 신뢰구간, 1.25 - 1.93) 더 높았다”

- 분석결과를 odds ratio와 신뢰구간을 이용하여 설명하였다.

나) 표 작성

- 반응변수의 측정단위 (mg/dL)와 이분화한 기준값 (220)을 제시하였다.

- 설명변수의 측정단위 (kg)와 회귀계수, 표준오차, P 값, odds ratio, 95% 신뢰구간을 제시하였다.

참 고 문 헌

1. Lang TA, Secic M. How to report statistics in medicine - Annotated guidelines for authors,

editors, and reviewers. Philadelphia (PA) : American College of Physicians; 1997.

2. 논문작성법교재편찬위원회. 논문작성법 (자연과학 편). 서울: 건국대학교 출판부; 1984.

3. 박용근, 민상기, 남용택. 의학 논문 작성법. 서울: 아카데미아; 2000.

4. 송혜향, 김동재. 통계학. 서울: 청문각, 1998.

5. 김윤진, 이상엽. 중년 직장 남성의 피로 관련 인자. 가정의학회지 2000;21(10):1277-86.

6. 옥선명, 최환석, 김철민. 태극권을 수련하는 폐경 여성에서의 척추 및 대퇴부 골밀도와 체력 요인. 가정의학회지 2000;21(1):46-55.

7. SAS Institute Inc. SAS guide for personal computers, Version 6.03. Cary (NC) : SAS Institute Inc.; 1988.