

## Power analysis

가천대길병원 G-ABC 방태모

### 1. Revision 요구 답변

Reviewer #2

Q5.

22 VJT is a high number for this procedure, however, there is no power analysis to support the conclusion that the lack of difference between range of motion and joint inclination is truly significant rather than underpowered.

A.

22건의 VJT는 꽤 많은 표본이긴 하지만, VJT 이후 ROM과 toe PIPJ의 articular surface shape간의 상관관계가 유의하지 않다는 해당 연구의 결론을 지지할만한 검정력 분석(power analysis, 이하 검정력 분석)이 결과에 포함되지 않은 점을 지적하는 것 같습니다.

검정력 분석에는 여러 가지 목적이 존재하고, 일반적으로 가장 많이 쓰이는 경우는 연구에 필요한 최소 표본 수(sample size)를 계산할 때입니다. 이러한 경우에는 검정력 분석은 연구 수행 이전에 수행될 것이고, 적절한 표본 수는 실험 결과를 뚜렷하게 확인할 수 있게 해줄 것입니다. 검정력 분석에는 다음의 4가지 값 중 3가지가 필요로 됩니다.

- (i) Effect size: 검정통계량과 귀무가설하의 검정통계량 값의 차이  
e.g. 다음과 같은 가설을 갖는 피어슨 상관관계수 검정으로 가정.

$$H_0 : \rho = 0 \text{ vs } H_1 : \rho \neq 0$$

여기서 주어진 자료로부터 계산된 상관관계수 추정량  $\hat{\rho}$ 의 값이 0.5라면:

$$\text{effect size} = 0.5 - 0.0 = 0.5$$

즉, 통상적으로 상관관계수 검정은 해당 상관관계수가 0인지에 관해 검정하므로, 상관관계수 검정의 effect size는 곧 표본으로부터 계산된 상관관계수 추정량의 값이 됩니다.

- (ii) Sample size: 표본 수

- (iii) Significance: 통계적 검정에 사용된 유의 수준. e.g. 일반적으로 0.05(5%)

- (iv) Statistical Power: 검정력. 즉, 귀무가설을 기각시킬 가능성을 의미.

본 연구에서는 상관관계수 검정이 수행되었고 각 검정을 통해 계산된 상관관계수 값(effect size)과 표본 수(n=22)가 주어져 있으므로, 본 연구의 유의 수준(significance)을 5%로 가정하면 연구에서 수행된 각 상관관계수 검정의 검정력을 계산할 수 있습니다. R 4.0.4의 pwr 패키지

(1.3.0)의 내부함수 `pwr.r.test()`를 통해 수행할 수 있습니다. 해당 연구 논문에는 수행된 4가지 각 검정의 p-value만 주어져 있어서, 다음과 같이 R 코드 러닝하시면 됩니다:

```
> pwr.r.test(n = 22, r = (각 검정의 상관계수 값 - 0), sig.level = 0.05)
```

## 2. 참고 사항

먼저, 검정력(power)의 의미를 간단하게 말하면 "귀무가설을 기각시킬 확률"입니다. 결론부터 말씀드리면 제가 보내드린 표에 계산된 각 검정의 검정력을 보시면 대부분이 낮은 값( $<0.5$ )을 취하고 있어, 본 논문의 주요 결과들인 "둘 간의 상관이 유의하게 존재하지 않는다."라는 주장에 대한 신뢰를 확보하기 힘들 것 같습니다.

일반적으로 검정력 분석은 적절한 표본 크기 결정을 위해 사전에 수행되는데(priori power analysis, 이하 사전 검정력 분석), 이때 통상적으로 검정력은 0.8, 유의 수준은 0.05가 되도록 기준을 삼고 effect size는 이전 연구에서 알려진 값을 참고하거나 연구자가 적절한 차이를 가늠하여 설정하고 실험 수행 이전 적절한 표본 크기를 계산하게 됩니다.

해당 연구에서 수행된 연구는 사전 검정력 분석 없이 수행되었고 리뷰어가 본 연구에서 수행된 대부분의 검정에서 귀무가설을 기각시키지 못하는 점을 보고, 혹시 검정이 underpowered여서 귀무가설을 기각시키지 못한 것은 아닌지를 의심하여 검정력 값에 대한 명시적 요구를 한 것으로 미루어 봤을 때, 논문 accept가 되기 힘들 수도 있지 않을까 하는 생각이 듭니다.

underpowered된 검정에 대한 근본적인 해결책은 사실 표본 크기를 더 확보하는 것인데 해당 연구는 표본 크기가 제한되어 있다고 하셔서 그마저도 힘들 거라고 생각이 됩니다. 마지막으로 제가 생각해본 한 가지 해결책은 후향적 연구에 해당하는 본 연구의 특성상 표본 크기를 연구자가 원하는 만큼 충분히 확보하기는 힘들다는 점을 이유로 들어, 검정력이 낮은 (underpowered) 검정들을 언급하며 해당 부분을 연구 한계점에서 언급해보는 방법도 있을 것 같습니다.