

연구설계

경북의대 예방의학교실 천 병 렬

I. 연구설계가 중요한 이유

과학적인 연구방법은 객관적인 접근법에 의해 현상에 내재하고 있는 진리를 규명하는데 중요하다. 즉, 현상에 대한 서술, 이해, 그리고 예측이 객관적이고 실증적인 연구방법에 의해 이루어질 때 그 연구의 결과는 믿을 수 있을 것이다. 의학에서 연구설계가 중요하게 인정되기 시작한 것은 1854년 런던에서 발생한 콜레라의 원인을 밝히고 그 대책을 강구하여 콜레라의 전파를 차단한 유명한 John Snow의 빛나는 업적 이후이다. 이는 Robert Koch가 세균을 발견하기 이전에 이룩한 연구로서 그 과학적인 연구수행과정과 연구결과의 정리는 연구방법론의 기초가 되었으며 이후에도 연구방법론은 꾸준히 발전하여 왔다. 미국에서 현재에도 계속되고 있는 Framingham heart study를 비롯하여 세계보건기구의 협력 하에 유럽을 중심으로 전 세계적인 network을 형성하고 있는 MONICA Project는 기술역학, 분석역학 그리고 임상시험까지를 포함한 연구이다. 우리 나라에서도 최근에는 이런 종류의 개입연구가 시작되고 있어서 타당한 연구결과를 얻기 위한 연구설계의 중요성에 대한 인식이 확대되고 있다.

산업의학의 영역에서도 직업병의 진단에 분자생물학이 개입하고 동시에 미량의 독성물질이 인체에 미치는 영향을 예측할 수 있는 모델의 개발이 요구됨을 볼 때, 산업장에 근무하는 근로자와 공해에 시달리는 인구집단에 관한 건강평가는 과학적이고 체계적인 연구설계에 의해 이루어져야 될 것이다. 또한 임상의학에서 신약개발에 대한 정부의 기준과 감시가 강화되어 과거와 같이 적당히 자료를 만들어 결과를 제시하는 임상시험은 인정되지 못할 것이다. 컴퓨터를 이용한 의료정보체계의 구축은 지구를 하나의 지역사회로 생각할 수 있는 길을 열어 놓았다. 그러나 중요한 것은 정보체계구축과 함께 양질의 정보를 제공하는 것이다. 특히 의학의 영역에서 잘못된 정보의 역기능적인 파급효과는 심각할 것이다. 연구설계는 이러한 양질의 정보의 생산에 가장 중요한 역할을 할 것이다. 연구방법론은 타당성 있는 연구자료의 수집에 더 중점을 두기 때문에 일반성보다는 타당성을 더 중요시한다(Good statistical analysis do not salvage poor data). 그러나 어떠한 연구방법이라도 이상적인 것은 없다. 따라서 여러 가지 연구방법의 장단점을 고려하면서 불완전성을 해결하는 것이 최선의 방법이다.

II. 연구설계의 목적

잘 고안된 연구설계에 의해 타당성 있는 자료를 얻을 수 있다. 그러므로 사전에 충분히 심사숙고하여 연구설계를 해야 한다. 연구설계의 궁극적인 목적은 타당한 연구 결과를 얻으려는 것이다.

연구설계의 구체적인 목적은 다음과 같다.

1. 연구 문제에 대한 해답을 경제적인 비용으로 얻기 위해서
2. 변량의 통제를 위해: 원인과 결과의 관계를 명확히 할 수 있다.
3. 외적타당성과 내적 타당성이 있는 결과를 얻기 위해서

III. 연구설계의 형태

연구는 크게 관찰연구(observational study)와 실험연구(experimental study)의 두 범주로 나눈다. 관찰연구는 기술연구(descriptive study)와 분석연구(analytic study)로 나눌 수 있고 분석연구는 다시 환자-대조군연구(case-control study)와 추적연구(cohort study)로 세분된다. 그리고 실험연구는 지역사회에서 실시되는 개입연구(intervention study)와 병원에서 주로 이루어지는 임상시험(clinical trial)으로 나누어진다. 관찰연구와 실험연구의 차이점을 이해하기 위해서는 두 가지 중요한 요소들을 알아야 한다(표 1.1). 이들은 연구요인에 관한 개입여부(intervention of study factor)와 무작위할당여부이다.

표 1.1 연구설계의 형태

연구 형태(type)	연구 소형태(subtype)
연구요인에 관한 인위적인 개입과 무작위할당이 있음	
실험연구(Experiment)	실험실(laboratory) 임상시험(clinical trial) 지역사회개입연구(community intervention)
연구요인에 관한 인위적인 개입은 있지만 무작위할당이 없음	
준실험연구 (Quasi-experiment)	실험실(laboratory) 임상실험(clinical trial) 정책, 계획(program/policy)
연구요인에 관한 인위적인 개입도 없고 무작위할당도 없음	
관찰연구(Observational)	기술연구(descriptive) 단면연구(사례연구) 분석연구(analytic) 환자대조군 연구 코호트 연구(전향성, 후향성)

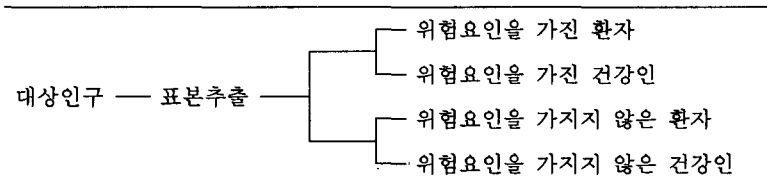
1.1 단면연구(Cross-Sectional Study) = 사례연구(The Case Study)

1.1.1 정의 및 연구설계 모형

관찰적인 연구방법의 한가지로 가장 기본적인 기술연구이다. 어느 한 시점에 한번의 조사를 하는 연구형태이다. 연구대상집단의 유병률 또는 위험요인의 분포를 조사하는 것이 주된 목적이다. 따라서 이를 유병조사(prevalence study)라고도 한다. 일반적으로 이와 같은 단면연구는 인구 집단에서 질병의 발생정도와 질병의 원인이 밝혀져 있지 않을 때 질병의 유병률을 알아보고 질병의 위험요인에 관한 가설을 얻기 위하여 실시한다. 이 연구는 보건의료문제 파악을 위한 기초조사이다.

단면조사는 어떤 연구집단(모집단)으로부터 표본 추출된 연구대상에 대하여 질병의 유병률을 조사하거나 질병과 관련된 위험요인을 알아내기 위한 연구이므로 그림 1.1의 연구설계를 참조하면 된다. 즉, 모집단으로부터 이를 대표할 수 있는 표본이 추출되면 개인들을 대상으로 관심질병의 유병상태와 이와 관련되었다고 생각되는 위험요인들에 관한 정보를 수집한다. 예를 들면, 혈청 총콜레스테롤치의 분포조사 또는 고혈압 유병률조사 등이다.

그림 1.1 단면연구의 설계



1.1.2 연구설계 시 고려할 사항

단면조사는 한 시점을 기준으로 하여 단 한번의 조사로 연구를 끝내는 경우가 대부분이므로 표본의 추출방법과 표본의 크기를 사전에 결정하여야 한다. 따라서 모집단을 대표할 수 있는 표본추출이 가장 중요하다. 단면연구는 연구설계의 특성상 인과관계에 관한 정확한 정보를 제공하기 어렵다. 그 이유는 위험요인이 질병의 발생보다 시간적으로 먼저 존재했는지를 확인하기 어렵기 때문이다. 그리고 일정한 시점에서 연구대상자의 특성을 관찰함으로써 누적유병률이 낮은 질병에 관한 연구로는 부적절하다. 즉, 단순한 유병률조사는 가능하지만 위험요인들과의 관련성을 조사하는 연구에서는 적당한 표본 수가 확보되어야하는 문제점이 있다. 그러나 단면연구는 비교적 짧은 시간 내에 연구결과를 얻을 수 있고 비교적 비용이 적게 들어 연구를 수행하기 쉽다. 그리고 새로운 가설을 제시하기가 용이하다.

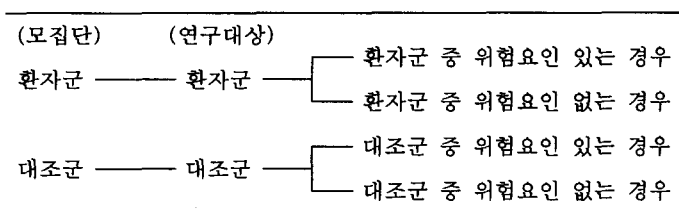
1.2 환자-대조군 연구(Case-Control Study)

1.2.1 정의 및 연구설계 모형

기술연구의 결과 어떤 요인이 연구의 결과인 질병발생과 관련이 있을 가능성이 발견되면 이것을 가설로 정립하고 이 가설을 검증하기 위한 분석연구가 필요하게 된다. 이 때 비용이 적게 들고 단시간 내에 결과를 알 수 있는 환자 대조군 연구가 있다. 환자-대조군 연구는 현재 조사하려는 질병에 이환된 환자(case)와 그 질병에 이환 되어 있지 않는 건강인 또는 다른 질병에 이환된 환자 중에서 선정된 대조군(control)을 대상으로 하여 연구대상 질병과 관련되었다고 생각되는 위험요인의 분포를 비교하는 연구이다. 그림 1.2의 연구설계를 보면, 환자군의 모집단에서 추출된 환자군과 대조군의 모집단에서 선정된 대조군의 위험요인의 분포를 비교 분석하는 형태의 연구이다.

환자 대조군 연구는 질병의 발생과 관련된 것으로 여겨지는 연구요인과 발생된 질병과의 관련성을 조사하는 것이므로 연구시점보다 연구요인의 발생이 선행되어 있고 질병의 발생이 이미 일어난 환자를 대상으로 하는 경우가 대부분이다. 따라서 시간의 경과를 중심으로 볼 때 후향성 연구(retrospective study)라고도 한다. 이 연구에서 유의할 것은 대조군이 반드시 건강인만을 의미하는 것이 아니고 다른 질병을 가진 환자도 대조군이 될 수 있음을 유의해야 할 것이다. 그리고 환자군이 반드시 환자군의 모집단을 대표할 필요는 없다는 견해도 있다.

그림 1.2 환자-대조군 연구의 설계



1.2.2 연구설계 시 고려할 사항

1. 환자군의 선정

환자-대조군 연구는 환자군의 선정이 가장 먼저 수행되어야 한다. 이 때 환자의 정의가 무엇이나를 결정하고 그 진단기준을 설정해야 한다. 이 진단기준의 설정은 객관적인 정보에 근거를 둔 명확한 것이라야 한다. 가능하면 정확한 검사방법에 의한 기준이 요구되지만 그렇지 않은 경우도 있을 수 있다. 그리고 진단기준의 설정시 반드시 선행연구를 참고로 하여 연구결과의 비교시 진단기준의 차이에 따른 문제점을 사전에 고려하는 것이 좋다. 진단기준을 결정할 때 또 한 가지 중요한 것은 비용의 문제이다. 즉, 진단기준을 너무 구체화시킴으로써 진단을 위한 비용이 연구를 수행하기 힘들 정도가 되면 안된다.

환자군의 선정시 두 가지 방법에 의해 연구대상자를 선정한다. 하나는 지역사회에서 발생한 환자를 환자군으로 하는 방법과 다른 하나는 병원에 입원한 환자를 연구대상으로 선정하는 방법이다. 두 가지 방법 모두 장단점이 있지만 연구목적상 전자가 후자보다 더 타당성 있는 연구 결과를 얻을 수 있다. 그러나 연구수행상의 문제 때문에 비용과 시간이 많이 든다. 따라서 병원 환자를 대상으로 한 연구가 실질적으로 많이 이용되고 있다. 이 연구는 환자군 모집단의 대표성이 연구결과의 타당성에 영향을 미칠 수 있다.

2. 대조군의 선정

질병진단의 기준이 설정되고 환자군의 정의가 명확해진 후 대조군에 관한 정의를 확정 지워야 한다. 대조군은 환자군의 선정과정과 동일한 방법에 의해 환자군이 선정된 동일 집단에서 선정되는 것이 좋다. 이는 동일 모집단의 개념이다. 이와 더불어 현실적인 가능성과 경제적인 측면을 고려해야 한다. 대조군의 선정이 환자-대조군 연구의 결과를 좌우하는 경우가 많다. 따라서 정확한 대조군 선정은 매우 어렵다. 환자-대조군 연구를 쉬운 연구로 보는 것은 환자-대조군 연구의 결과를 해석할 때 부딪히는 어려움을 외면한 결과이다. 대조군 선정에서 가장 중요한 것은 환자군과의 비교성(comparability)이 높아야 되는 점이다. 대조군이 환자군과 연구하려는 연구요인(study factor)외에는 다른 차이가 없을수록 연구의 비교성은 높아진다. 이 비교성을 높이기 위해서 환자-대조군 연구에서는 다음과 같은 방법을 이용하여 연구대상자를 선정한다.

1) 표본추출방법

표본추출에 의한 대조군의 선정방법은 연구자의 편견을 줄일 수 있지만 혼란변수를 고려할 수 없는 경우가 많다. 그리고 실제 환자-대조군 연구 수행 상에 있어서 전체 인구를 대상으로 표본추출이 쉽지 않다.

2) 제한적인 연구대상자 선정방법

혼란변수를 통제하기 위한 방법으로 쓰여지는 제한(restriction)은 어떤 경우에는 연구대상자 선정의 비용과 실행성을 높여주는 효율적인 방법이다. 제한은 완전제한과 부분제한으로 나눌 수 있는데 환자-대조군 연구에서는 부분제한(짝짓기; matching)을 주로 이용한다. 그러나 연구방법론적으로 실행성을 높여주는 것은 확실하더라도 연구결과의 타당성을 높이는 관점에서 볼 때 환자-대조군 연구에서 짝짓기는 큰 도움이 안된다. 일반적으로 짝짓기를 하면 연구결과의 타당성이 증가되는 것으로 잘못 알고 있으므로 주의해야 할 것이다. 그 이유는 짝짓기나 제한을 통해 비교성이 높아지는 경향이 있는 데 이를 그대로 연구의 타당성과 연결시켜 생각하기 때문으로 보여진다. 그리고 짝짓기를 할 때 어떤 변수를 짝짓는 변수로 사용할 것인지에 대해 충분한 고려 없이 짝짓기를 하면 안되며 일단 짝짓기를 하면 짝짓기에 사용된 변수들과 질병과의 관련성은 볼 수 없다는 제한점이 있다. 뿐만 아니라 통계학적으로도 짝짓기가 반드시 더 낫다는 보장도 없다.

3. 위험요인에 대한 정보 수집상의 문제

위험요인은 연구결과가 일어나기 이전의 상태를 조사하는 것이므로 환자-대조군 연구의 특성

상 과거 수십년전 또는 수년전의 기억이나 자료에 의존해야한다. 이런 경우 환자군과 대조군사이의 차이가 기억에 의한 차이가 되지 않도록 주의해야한다. 즉 회상편견(recall bias)이 일어나지 않도록 세심한 주의를 기울여야 한다. 특히 대부분 위험요인에 관한 조사는 설문지를 통한 면접조사이므로 면접자는 연구대상자로 하여금 과거의 상태를 회상하면서 대답할 것을 주지시켜 현재의 상태를 대답하지 않도록 주의해야 한다.

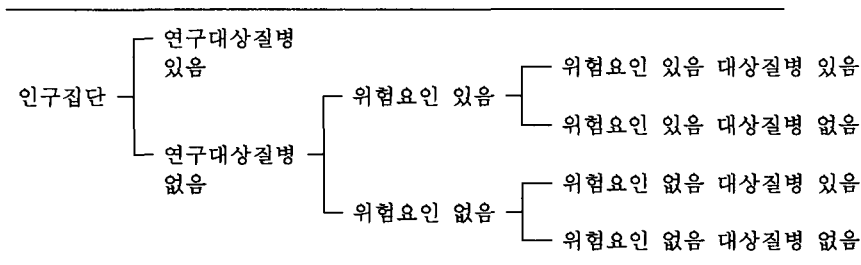
환자-대조군 연구의 장점은 비용이 적게 들고 연구계획과 연구수행이 비교적 용이하고 연구결과도 신속하게 얻을 수 있다는 것이다. 특히 발생률이 낮고 질병의 발생에 요구되는 시간이 오래 걸리는 비전염성질환의 연구에 많이 이용된다. 즉, 암과 같은 질환의 위험요인 연구에 용이하다. 발생률이 상당히 낮은 질병에 관한 연구는 환자-대조군 연구이외의 방법으로 하기는 힘들다. 환자-대조군 연구는 표본 수가 적어도 가능하고 기존자료를 이용할 수 있다. 그러나 단점으로는 여러 종류의 편견이 개재되기 쉬워 정확한 연구결과를 얻기가 힘들다. 과거의 기억에 의해 정보를 얻는 경우가 대부분이므로 설문지나 면접조사시 연구자의 노력에도 불구하고 연구대상자의 편견이 부지불식간에 일어날 수 있다. 따라서 연구결과 해석 시 많은 주의를 요한다.

1.3 추적 연구(Follow-Up Study) = 코호트 연구(Cohort Study)

1.3.1 정의 및 연구설계 모형

추적연구란 질병발생요인과 관련되었다고 의심되는 어떤 연구요인에 노출된 사람과 노출되지 않은 사람들을 연구대상자로 선정하여 이들을 일정기간 동안 추적 관찰하여 연구대상질환의 발생률이 어느 집단이 높은지를 비교함으로써 그 요인과 질병발생간의 관련성을 규명하는 것이다. 그림 1.3은 추적연구의 설계이다.

그림 1.3 추적연구의 설계



1.3.2 추적연구의 종류

추적연구는 연구시작시점과 연구결과가 일어난 시점이 시간상 명백한 선 후 관계가 확립되어 있지만 연구수행시점이 이미 연구결과가 확인된 경우일 때와 연구수행시점이 요인에 노출된 시점 일 때에 따라 두 가지로 구분된다. 즉 전향성추적연구와 후향성추적연구이다.

1. 전향성 추적연구(Prospective follow-up study)

연구시작시점에서 질병발생원인에 노출된 집단과 노출되지 않은 집단을 구분하고 그 때부터 일정기간동안 추적 관찰하여 대상질환의 발생률을 비교 관찰하는 방법이다.

2. 후향성 추적연구(Retrospective follow-up study)

연구시작시점에 이미 대상질환의 발생이 일어난 경우로 연구시점이전의 기록 또는 조사에 의해서 질병발생원인에 노출된 집단과 노출되지 않은 집단을 구분하고 그 때부터 현재 연구시작시점동안 추적된 대상질환의 발생률을 비교 관찰하는 방법이다. 이 연구는 환자-대조군연구의 장점과 추적연구의 장점이 결합되면 타당한 연구결과를 단시간 내에 얻을 수 있다. 이는 정보화사회에서 매우 유용하게 쓰일 수 있는 연구방법이다.

1.3.3 연구설계 시 고려할 사항

추적연구에서 중요한 것은 연구대상질환에 이환 되지 않은 사람들을 선정하는 것이다. 그러므로 질병에 이환 되지 않았다는 것을 증명하는 방법을 정의하고 이를 먼저 확인하여야한다. 두 번째로 중요한 것은 질병발생이전에 연구요인에 노출된 지를 확인하고 이들의 노출여부를 판단하여 연구대상자를 분류하는 일이다. 특히 추적연구는 장기간 관찰하여야하므로 추적가능한 사람들을 선정하여야 한다. 추적연구에서는 장기간의 연구기간이 요구되므로 진단기준의 확실성이 요구된다. 즉 진단기준의 차이에 의한 분류오차를 최소화하기 위해서 항상 연구시작이전에 진단기준을 명백히 정해야 한다. 추적관찰을 얼마나 정확하게 많이 하느냐가 이 연구의 가장 중요한 요소이다. 따라서 추적불능인 사람들을 수를 최소한 줄여야하며 동시에 추적이 안된 집단이 선택적인 특성을 가지지않도록 반드시 확인해야 한다. 그러나 아무리 추적율이 높더라도 추적이 안된 사람들과 추적된 사람들 간의 특성이 차이가 심해 연구결과 해석에 영향을 준다면 곤란하다. 즉 양적인 추적과 질적인 추적의 상호보완이 필요하다. 일반적으로 연구대상자의 80% 이상이 추적되면 양적으로는 추적율이 양호하다고 한다. 추적연구의 한계성은 항상 추적불능에 있다. 추적조사의 경우 연구의 타당성은 추적율에 좌우된다. 따라서 추적율을 높이도록 연구자는 최선을 다해야 한다.

추적연구는 질병의 원인규명에 더 결정적인 정보를 제공하여주는데 그 이유는 원인이 결과보다 항상 먼저 일어난다는 점이 방법론상 확실하기 때문이다. 추적연구는 환자-대조군 연구와 달리 발생률을 알 수 있으므로 비교위험도와 귀속위험도를 직접 구할 수 있다. 연구결과를 알기전에 연구요인에 노출된 지를 알 수 있으므로 인과관계를 확실하게 평가할 수 있고 정보편견이 개입할 가능성이 훨씬 적다. 한가지 위험요인과 관련된 여러 가지 질병의 연구가 용이하다. 그러나 연구기간이 길고 연구대상자 수가 많아져 비용과 시간이 많이 든다. 추적기간 동안에 연구대상자의 연구요인에 대한 노출상태나 정도가 변화할 가능성이 있으며 진단의 기준이 달라질 수 있다. 추적도중 탈락될 가능성이 높다. 그리고 연구대상질환의 발생률이 너무 낮을 때는 연구를 수행하기 힘들다.

1.4 실험 연구

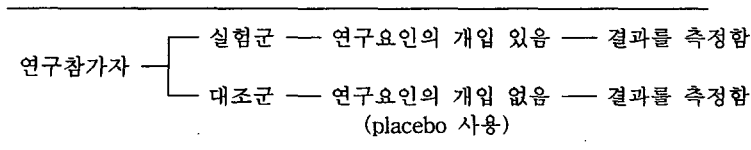
1.4.1 정의 및 연구설계 모형

실험연구는 관찰적인 연구방법과 달리 연구요인을 통제할 수 있기 때문에 인과관계를 밝히는 데 가장 유용한 방법이다. 실험은 연구자가 실험대상자에 대해 인위적으로 조작을 가할 수 있는 이점이 있지만 윤리적인 문제를 고려해야 하는 제한점도 있다. 따라서 동물실험을 실시하여 이와 같은 제한점을 극복하려하지만 동물실험의 결과를 인간에게 직접 적용하지 못한다는 단점이 있다.

이 연구는 추적조사의 속성을 띠고 있다. 연구자가 연구대상자를 실험군이나 대조군에 무작위할당시키고 연구요인을 조작하는 기법이다. 기본적인 연구의 설계는 그림 3.4와 같다. 즉, 연구에 참여하기를 원하는 사람들을 대상으로 무작위할당(random allocation)을 실시하여 실험군과 대조군으로 나눈다. 이 때 윤리적인 문제를 고려하고 연구대상자로부터 연구참여동의서를 받아야 된다. 무작위할당의 방법은 여러 가지가 있다. 무작위할당을 실시하는 이유는 인간을 연구 대상으로 하고 연구자가 실험의 결과판정에 깊이 관여할 수 있고 연구대상자도 어느 정도 연구의 목적을 알고 있기 때문에 이들에 의한 편견의 개재가능성을 배제하기 위함이다. 실험연구는 이와 같은 편견의 개입을 사전에 효과적으로 방지할 수 있는나에 의해 연구결과의 타당성이 좌우된다. 따라서 실험연구에서는 이와 같은 무작위할당이 제대로 되었는지가 가장 중요하다.

그리고 무작위할당이 된 다음에는 결과의 판정이 공정하게 이루어지도록 노력해야 한다. 이를 위해서는 맹검법(blind technique)이 이용된다. 가장 대표적인 것으로는 이중맹검법(double blind technique)이 있다.

그림 1.4 실험연구의 설계



1.4.2 실험연구의 종류

실험연구는 크게 3가지로 분류한다.

- 1) 실험실 실험연구(Laboratory experiment study)
- 2) 임상시험 (Clinical trial)
- 3) 지역사회 개입연구 (Community intervention study)

1.4.3 연구수행 시 고려할 사항

사람을 대상으로 하는 연구이므로 근본적으로 추적조사가 가지는 제한점을 다 가지고 있다. 뿐만 아니라 윤리적인 문제를 고려해야하므로 실험연구 자체를 수행하는 것조차도 불가능할 수 있다. 따라서 항상 자원자를 대상으로 연구를 수행해야 함으로 선택편견의 개입가능성이 높다. 연구대상자들이 연구자의 의도대로 따라주지 않을 때는 통제하기 불가능하다. 따라서 연구시작 이전에 윤리적인 문제를 고려한 후 실험연구는 시작되어야한다.

이 연구는 연구요인이외의 혼란변수들에 관해 사전에 조정이 가능하고 연구요인의 인위적인 조작이 가능하므로 인과관계를 밝히는 데 가장 유용한 연구방법이다. 특히 임상에서 치료효과와 판정에 매우 유용하다. 일단 제대로 무작위할당과 맹검법이 시행된 후에는 연구결과와 판정이 명확하고 타당성 있게 이루어질 수 있으므로 결론을 내릴 수 있다. 그러나 앞서 지적한 바와 같이 추적연구의 단점은 모두 가지고 있으며 윤리적인 문제를 사전에 해결해야 하는 어려움이 있다. 대개의 실험연구는 연구요인 개입전 후의 차이를 실험군과 대조군을 대상으로 비교해 보는 것이다. 따라서 여러 가지 지표가 이용된다. 추적연구의 지표와 같은 발생률, 비교위험도 등을 측정할 수도 있지만 어떤 지표의 변화정도도 측정될 수 있다.

IV. 연구설계 시 고려할 사항들

1. 연구의 구체적 목적과 가설
2. 독립변수(연구요인)의 조작 가능성
3. 외적변수(extraneous variables)의 통제 정도
4. 연구상황의 자연성
5. 연구결과와 일반화 정도
6. 연구수행의 용이성(시간, 비용)

참고문헌

1. Baker TL: Doing social research, McGraw-Hill, Inc. Singapore, 1988
2. Bouma GD: The research process, Oxford University Press, Revised edition, Australia, 1993
3. Beaglehole R, Bonita R, Kjellstrom T: Basic epidemiology, World Health Organization, Geneva, 1993
4. Fletcher RH, Fletcher SW, Wagner EH: Clinical epidemiology, the essentials, second edition, Williams and Wilkins, 1988.
5. Gordis L: Epidemiology, W.B. Saunders Company, 1996.