

일차의료 중재에 관한 리얼월드 평가 안내서

(A guide to real-world evaluations of
primary care interventions)

**일차의료 중재에 관한
리얼월드 평가 안내서**

본 책자는 아래 원저자로부터 번역 및 활용에 대한 허가를 받아 작성되었습니다.

This publication was translated and adapted in part, with permission, from:

A guide to real-world evaluations of primary care interventions: Some practical advice. U.S. Agency for Healthcare Research and Quality; Rockville, MD (USA). October 2014.

일차의료 중재에 관한 리얼월드 평가 안내서

(A guide to real-world evaluations of primary care interventions)

발행인: 이영성

발행처: 한국보건 의료연구원

인쇄: 2018년 11월 1일 초판 1쇄 발행

서울시 중구 퇴계로 173 남산스퀘어빌딩 7층

Tel. 02-2174-2700

Homepage: www.neca.re.kr

ISBN: 978-89-6834-486-2

집필진

본 안내서는 다음과 같이 인용될 수 있습니다.

Suggested citation:

Peikes, D., Taylor, E., Genevro, J., & Meyers, D. A guide to real-world evaluations of primary care interventions: some practical advice. HS Kim, YS Suh, JH Park, BN Yoo, JH Kim, JY Park, translator. U.S. Agency for Healthcare Research and Quality; Rockville, MD (USA). 2014.

저자 Deborah Peikes, Ph.D., M.P.A., Mathematica Policy Research
Erin Fries Taylor, Ph.D, M.P.P, Mathematica Policy Research
Janice Genevro, Ph.D., Agency for Healthcare Research and Quality
David Meyers, M.D., Agency for Healthcare Research and Quality

역자 김희선 한국보건의료연구원 부연구위원
서유신 한국보건의료연구원 과제연구원
박진향 Eötvös Loránd University(ELTE) 대학원
Health Policy, Planning and Financing 석사
유빛나 한국보건의료연구원 연구원
김진희 한국보건의료연구원 과제연구원
박종연 한국보건의료연구원 선임연구위원

발간사

한국보건 의료 연구원은 공공기관 자료 연계의 법적 권한을 가지고 보건 의료 근거를 생성하고 의료 기술을 평가하는 기관으로, 보건 의료 정책에서의 공익적 근거 창출에 기여하는 공공기관입니다. 이러한 역할의 일환으로 보건복지부는 HTA 연구 방법론을 활용한 보건 의료 정책 평가 연구를 우리 연구원에 위탁하였고, 이를 통해 일차 의료 정책에 대한 실질적인 문제점과 효과를 과학적으로 도출하기를 기대하였습니다. 이에 우리 연구원은 약 3년간 국가의 보건 의료 정책 성과 평가를 위한 연구팀을 만들어 운영하고 연구 외에도 '일차 의료 발전 정책 근거 지원 TF'를 구성하여 정책이 실제 개발되고 적용되는 전반적인 의사결정을 지원 하였습니다.

최근 유럽을 중심으로 공중 보건 증재에 대해 의료 기술 평가 방법론을 적용하기 위한 노력이 활발하게 진행되고 있는데, 이미 우리 연구원은 한국적 상황에서 이를 고민하고 있었다고 해도 과언이 아닐 것 같습니다. 물론, 우리 연구원에서 수행한 공중 보건 증재에 관한 연구(보건 의료 정책 성과 평가 연구들)들은 시작 단계에 불과하나, 지난 몇 년간 수행한



연구 사례를 통해 리얼월드 데이터를 보다 체계적으로 수집하고 과학적 방법론으로 분석하여 얻은 현장 중심 근거가 정책에 어떻게 선순환적으로 활용될 수 있는지 경험하였습니다.

그러나 아직 연구를 통해 고민해야 할 것이 많고, 새롭게 적용해야 할 방법론을 체계적으로 정리하고 확산해야 할 과제가 많이 남아있습니다. 이러한 고민을 앞서 경험한 미국의 사례를 살펴보고 우리나라의 경험을 반추하면서 보다 발전적인 공중보건학적 중재 평가로서의 보건의료 정책 성과 평가 연구가 활성화되길 기대합니다.

2018년 10월

한국보건의료연구원 

역사의말

한국보건의료연구원은 공공기관 자료 연계의 법적 권한을 통해 보건 의료 근거를 생성하고 의료기술을 평가하고 있는 공공기관으로, 지난 3년간(2015~2017) 보건복지부로부터 ‘지역사회 일차의료 시범 사업 평가 연구’를 의뢰받아 수행한 바 있습니다. 이러한 경험을 바탕으로 2018년 현재 ‘근거기반 보건의료정책 성과평가 방안 연구’(연구 책임자 : 김희선)를 진행하던 중, 최근 미국 AHRQ(The Agency for Healthcare Research and Quality)에서 발표한 일차의료 세팅에서의 리얼월드 데이터(Real-world data, RWD)를 활용한 평가 가이드라인을 접하게 되었습니다.

미국의 평가 경험은 저희 연구진의 지난 3년간의 연구에서 경험한 바와 유사한 것들이 많았습니다. AHRQ는 이러한 경험을 살려 보다 많은 연구자들이 현장의 데이터를 활용할 수 있도록 가이드라인을 만들었습니다. 과학적이고 체계적인 평가방법으로 검증함으로써, 근거기반의 보건의료정책 실현을 도모하고자 하는 미국의 노력이 공감되었습니다. 저희 연구진과 같이 리얼월드 데이터를 활용한 보건의료정책 평가 연구, 특히, 일차의료 현장에서 수집되고 축적된 데이터를 활용한 보건

의료 정책의 근거 생성에 관심을 가지고 계실 현장의 의료 관계자, 연구자, 정책가분들께서 미국의 경험을 보다 손쉽게 접할 수 있도록 공유하고 싶었습니다. 이에 AHRQ의 허가를 받아 작은 책자로 발간할 용기를 가지게 되었습니다.

단순 번역 자료이기 때문에 국내 평가 경험을 덧붙여 낼 수 없어 아쉽지만, 우리나라 보건의료체계의 개선을 위한 주요과제인 일차의료 발전에 기여할 자료로써 활용될 수 있기를 기대합니다. 저를 비롯한 연구진은 앞으로도 한국보건의료연구원이라는 공공기관의 연구원으로서 리얼월드 데이터를 공익적 용도로 사용하고 이를 잘 활용하여 보건정책의 의사결정에 활용하는 것을 중요한 가치로 여기고 최선을 다해 연구로서 지원하고자 합니다.

부족한 자료겠지만 지원과 응원의 말씀을 부탁드립니다. 지금까지 같이 애써주신 여러 연구진 선생님들께 감사의 말씀을 드립니다.

2018년 10월
대표역자 김희선

일러두기

1. 「일차의료 중재에 대한 리얼월드 평가 안내서」는 한국보건의료연구원에서 AHRQ의 허가를 받아 일차의료 중재 평가 단계 및 예시, 방법론 리소스를 정리한 문헌을 번역한 한국보건의료연구원의 공식 출간물입니다.
2. 본 안내서는 한국보건의료연구원 외의 다른 기관이나 단체에서 일차의료 중재 평가를 수행할 때 활용될 수 있습니다. 그런 경우에는 출처를 밝히도록 합니다.
3. 별첨은 안내서와 별도로 한국보건의료연구원에서 수행한 일차의료 시범사업 추진 관련 연구를 요약 소개한 자료입니다.
4. 본 안내서는 한국보건의료연구원 홈페이지(<http://www.neca.re.kr>)에서 열람 가능합니다.

차례

평가 안내서에 대한 간략한 소개	1
Ⅰ. 평가가 필요한가?	4
Ⅱ. 평가를 위해 무엇이 필요한가?	6
Ⅲ. 평가 계획을 어떻게 세울 것인가?	8
Ⅳ. 평가를 어떻게 수행하며 이를 통해 우리는 무엇을 알 수 있는가?	15
Ⅴ. 결과를 어떻게 활용할 수 있는가?	22
Ⅵ. 일차 의료 모델 평가를 위한 리소스 모음집(resource collection)	23
부록 A : 유의미한 효과를 감지할 수 있는 가능성을 평가하기 위한 클러스터링을 통한 통계능력 계산하기	27
부록 B : 중재가 없었다면 일어났을 일을 설명하기 위해 비교 그룹 사용하기	29
부록 C : 프로그램 영향평가를 위한 분석 방법의 활용	31
참고문헌	32
별첨 : 시범사업 추진 관련 NECA 수행 연구 요약	34

평가 안내서에 대한 간략한 소개

목표. 효과적인 일차의료는 건강 및 비용, 환자, 임상과의 의료관계자의 경험을 개선하며, 평가를 통해 이러한 목표를 달성하기 위한 최선의 일차의료 개선 방법을 모색해 볼 수 있다. 본 평가 안내서는 환자중심메디컬홈(PCMH)^a과 다른 서비스 모델과 같은, 일차의료 제공 체계를 개선하기 위한 중재들의 현실 평가방법을 설계하는데 도움이 될 수 있는 실질적인 조언을 제공한다.

목표대상. 이 안내서는 의료전달체계 관련 분야에 종사하는 평가자, 고용주, 진료기반 연구 네트워크, 지역 또는 광역 보험사, 그리고 상대적으로 적은 수의 일차의료 사례를 바탕으로 새로운 중재 방법에 대한 평가를 진행 및 그 중재를 평가하기 위해서 제한적인 자원을 가진 이들을 위해 고안되었다.^b

개요. 이 안내서는 소수의 사례를 바탕으로 한 일차의료 중재를 평가하는 데 필요한 몇 가지 실질적인 단계를 제시한다. 이를 통해 새로운 의료모델의 수행 사례를 평가하고, 진료 및 결과를 개선시키는 데 필요한 세부 지침으로 사용할 수 있는 정보를 제공한다. 이 안내서는 소규모 프로젝트 평가자를 위한 문제해결 방안 및 대규모 프로젝트 평가자를 위한 통찰력을 제공한다. I부터 V장까지는 아래의 질문에 대한 답변으로 구성되어 있다. VI장의 리소스모음집(resource collection)에는 미국의 보건부 산하 AHRQ^c에서 제공한 자료를 비롯하여 평가를 기획하고 수행하는 데 도움이 되는 도구와 자료가 포함되어 있다. 몇 가지 부록에는 정량적 효과 평가에 도움이 되는 추가적인 기술적 세부사항이 포함되어 있다.

a. 역자주) Patient Centered Medical Home. 단순히 의료를 제공하는 장소가 아닌 다음의 핵심 기능을 수행하는 일차의료 서비스 조직 모델: 포괄적 의료(comprehensive care), 환자 중심성(patient centered), 조직적 진료(coordinated care), 접근가능한 서비스(accessible services), 품질 및 안전(quality and safety) (출처: PCMH 홈페이지. <https://pcmh.ahrq.gov>)

b. 역자주) 미국의 의료보험제도는 고용 관계를 기초로 한 민간보험이 주를 이루기 때문에 고용주, 보험사 등이 보건의료 전달과 밀접하게 관계되어 있음.

c. 역자주) Agency for Healthcare Research and Quality. 미국의 보건의료 시스템의 안전성과 품질 개선에 필요하거나 관련 정책 입안에 필요한 지식, 도구 및 데이터를 개발하는 연방 기관(출처: AHRQ 홈페이지. <https://www.ahrq.gov>)

I. 평가가 필요한가? 모든 중재를 평가할 필요는 없다. 중재 사안이 상대적으로 사소하거나 저비용으로 행해지는 경우, 또는 확고한 근거가 존재하거나 품질 개선 노력의 일환으로 행해지는 경우에는 평가가 필요 없을 수도 있다. 그렇지만 다수의 중재는 연구를 통해 혜택을 받을 것이다. 평가 수행 여부를 판단하기 위해서는 평가의 수행 비용을 고려하여 특별한 결정 사항을 파악하는 것이 중요하다. 평가는 핵심이 명확하고 비용이 많이 들며, 확실한 근거가 존재하지 않는 중재 사례에 유용하다. 평가를 통해 특정 중재가 해당 사례에서 의료를 제공하는 방법에 영향을 미쳤는지, 의료 제공의 변화가 결과에 어떻게 영향을 미쳤는지 등과 같은 핵심적인 문제에 대해 해답을 얻을 수 있다. 중재모델에 대한 피드백과 초기의 성공 지표는 중재를 개선하는 데 도움이 될 수 있다. 평가 결과를 바탕으로 다른 사례를 유도해 내는 것도 가능하다. 고려해야 할 핵심 질문은 “수행할 자원이 있는지 평가하여 신뢰할 수 있는 유효한 결과 생성이 가능한가?”이며, 프로그램의 영향에 대한 편향된 추정치는 이해당사자를 오도하기 때문에 아무런 결과를 얻지 못하는 것보다 오히려 더 나쁠 수 있다. 본 안내서는 평가 수행 여부와 평가가 자원과 상황에 적합한 선택인지를 판단할 수 있게 돕는다.

II. 평가를 위해 무엇이 필요한가? 중재를 실시하고 평가를 수행하기 위해서는 이에 필요한 자원을 파악하는 것이 중요하다. 평가에 필요한 일부 자원으로는 (1) 리더십 바이 인(buy-in)과 지원, (2) 데이터, (3) 평가 기술, (4) 평가자, 임상 의료진 및 관계자들이 역할을 수행할 수 있는 데이터를 수집하기 위한 시간 등을 들 수 있다. 이 단계에서는 올바르게 설계된 평가를 수행하는데 드는 비용을 정확하게 파악하는 것과 이 비용을 중재의 성격, 범위 및 비용과 관련시켜 고려하는 것이 중요하다.

III. 평가 계획을 어떻게 세울 것인가? 가능한 한 질 높은 정보를 제공하기 위해서는 중재가 시작되기 전에 평가를 기획하는 것이 가장 좋다. 적절한 연구 질문을 선정하고 그에 맞게 평가를 기획할 수 있도록 목표와 대상자를 먼저 설정해야 한다. 그 다음엔 평가를 위해 사용할 수 있는 자원 목록을 작성하고 이를 이용했을 때 어떤 문제에 해답을 제공해 줄 수 있을지에 대한 기대치를 조정한다. 그런 다음, 중재를 위한 기본 논리 또는 변화 이론을 설명한

다. 이는 중재가 결과를 향상시킬 것이라고 기대하는 이유와 결과가 향상될 것이라고 기대하기 전에 발생할 수 있는 몇 가지 단계를 설명해야 한다. 그러나 논리모형을 통해 정해야 할 사항과 시기를 설명할 수는 있지만, 예상치 못한 정보와 의도하지 않은 결과에 대해서도 수용적인 자세를 취해야 한다. 이러한 논리 모델은 평가의 범위와 기획을 이용 가능한 자원에 맞게 조정하는 데에도 도움을 줄 것이다.

IV. 평가를 어떻게 수행하며, 이를 통해 우리는 무엇을 알 수 있는가? 다음 단계는 중재의 수행방안에 대한 연구를 설계하는 것이다. 또한 통계적으로 유의미한 변화를 감지할 수 있는 충분한 사례가 포함되어 있으면, 그 영향에 대한 연구를 설계하는 것도 이 단계에 포함된다. 이해당사자들이 연구결과에 관심을 가지고 있을지라도, 테스트한 사례의 수가 적은 경우에는 일반적으로 비용과 품질에 미치는 영향에 대한 신뢰할 만한 추정치를 산출하기가 어렵다. 이러한 경우, 해당 모델을 이행하는데 어떤 것이 장애물이 되거나 촉진제로 작용하는지를 이해하기 위해 정성적 분석(qualitative analysis) 방법을 사용할 수도 있고, 진료 절차의 변화나 환자의 경험과 같은 중간 결과를 측정하기 위해 정량적 분석(quantitative analysis)을 사용할 수도 있다. 이러한 분석과정은 좀 더 세심하게 살펴보아야 할 영역과 결과를 향상시킬 수 있는 잠재력을 알아내는데 도움이 된다.

V. 결과를 어떻게 활용할 수 있는가? 실행평가(implementation evaluations) 결과는 중재를 시행할 수 있을 지 여부와 중재를 개선할 수 있는 방법을 알려줄 수 있다. 진료 수행 및 영향평가(impact evaluation)의 결과를 종합하면(영향평가를 수행할 수 있는 경우) (1) 테스트 중인 모델의 효과에 대해 보다 깊은 이해를 제공할 수 있고, (2) 가장 효과적인 대상 환자, 사례 및 환경 유형을 식별할 수 있으며, (3) 개선과 확산에 대해 현명한 결정을 내릴 수 있도록 한다.

VI. 활용할 수 있는 리소스에는 무엇이 있는가? 본 안내서의 리소스 모음에는 논리 모델 개발, 진료 시행 및 결과 측정 방법 선택, 분석 방법 설계 및 수행, 실행 및 영향평가의 결과 종합에 사용할 수 있는 리소스와 도구가 포함되어 있다.

1. 평가가 필요한가?

당신은 중요한 결과를 도출하기 위한 희망 속에 일차의료를 제공하는 방식을 변경해보기로 결정했다. 첫 번째 질문은 중재 평가 여부에 관한 것이다.

모든 일차의료 중재가 평가 대상이 되는 것은 아니다. 변화의 필요성이 명백할 때에는 간단히 중재를 수행할 수 있다. 예를 들어, 환자가 저녁 진료 시간이 부족하다는 의견을 주거나, 긴급 치료 센터가 사업 경쟁력을 잃게 되면 일차의료 기관은 평가를 거치지 않고도 저녁 진료 시간을 추가할 수 있다. 의료 이용률 및 접근성에 관한 환자의 피드백을 확인하고 싶겠지만, 중재에 대한 완전한 평가가 보장되지 않을 수도 있다. 또한 일부 운영 및 질 향상 변화는 지역에서 관리하는 계획-실행-평가-개선 주기(plan-do-study-act cycles)를 통해 평가될 수 있다. 이러한 변경의 예로는 예약 기간의 변경과 환자를 위한 교육 자료의 개선 등이 있다. 마지막으로, 이전에 발표된 연구가 유사한 인구 집단과 환경에서 명백한 증거를 제공했을 때에는 그러한 중재를 다시 테스트할 필요가 없다.

일차의료 중재를 채택하는 데 비용이 많이 들고, 중재를 널리 보급할지 여부를 고려하고 있다면 더욱 철저한 평가가 필요할 수 있다. 평가는 중재를 가장 효과적으로 수행하는 방법과 그것이 결과에 어떤 영향을 미칠 수 있는지에 대해 파악할 수 있도록 돕는다. 진료를 통해 추구하는 변화를 도출할 수 있는지, 개선된 중재를 보다 원활하게 실행할 수 있는지, 그러한 노력이 타당한 것이었는지를 입증할 만큼 결과가 충분히 개선되었는지 검토할 수 있다. 또한 의원, 시장 및 환자의 특성에 따라 결과가 어떻게 달라졌는지 확인할 수 있다. 중재의 결과에는 일반적으로 진료 비용 및 품질, 환자, 임상가와 의료진의 경험이 포함된다. 중재의 효과가 투입된 비용에 비해 좋을 경우, 실행 및 영향분석으로부터 얻은 결과는 중재를 개선하고, 중재에 자금을 계속적으로 조달하고, 또는 그것을 더 많은 기관으로 확산시키는 데 도움이 될 수 있다.

[그림 1]은 일차의료 중재의 평가를 계획하고 시행하는 단계를 요약한 것이다. 우측 두 개의 박스는 평가의 이점을 보여준다.

평가가 필요한가? 그렇다면 어떤 리소스가 필요한가?
(12장 참고)

중재를 평가할 가치가 있는가?

강력한 중재를 위한 리소스:

- 리더십 바이 인 재정적 및 기술적 지원 도구, 시간

강력한 평가를 위한 리소스:

- 리더십 바이 인 재정적 리소스, 연구 기술과 전문지식, 데이터, 시간

평가를 어떻게 계획할 것인가?
(3장 참고)

평가의 목적과 대상자를 고려하고 중재를 계획하는 것과 동시에 평가를 계획할 것. 어떤 질문에 답해야 하는가?

- 주요 평가 과제를 이해할 것

현실적인 기대를 할 것. 접근 방식을 리소스 및 데이터와 일치시킬 것.

중재의 모든 구성 요소를 아우르는 논리를 결정할 것.

- 중재는 어떻게 실행되고 있는가?
- 어떻게 A가 B로 B가 C로 연결되는가?
- 비용, 질, 환자, 임상 및 의료진의 경험에 대한 중재의 예상 효과는 무엇인가? 언제 이것이 발생하기를 기대하는가?
- 어느 상황별 요인, 과정 지표 및 결과를 추적해야 하며, 그것을 언제 실행해야 하는가?
- 의도하지 않은 결과를 예상할 수 있는가?

평가를 어떻게 수행하고 평가가 어떤 질문에 답할 것인가?
(4장 참고)

각 데이터 소스의 부담과 비용을 고려한 실행 연구를 설계하고 수행한다:

- 중재는 어떻게, 그리고 얼마나 잘 실행되고 있는가?
- 실행 장벽과 가능한 제거 방법을 확인할 것.
- 실행에 사용된 계획의 변형과 그 이유를 확인할 것.
- 의도되지 않은 결과를 확인할 것.
- 필요 시 시간의 경과에 따라 중재를 조정할 것.

충분한 통계력이 있다면 영향에 관한 연구를 계획하고 실시할 것:

- 서로 다른 데이터 소스에 대한 비교 그룹 설계, 추정 방법 및 샘플을 고려할 것.

결과를 종합할 것.

- 중재 활동이 단기 또는 중간 결과와 관련 있는 것으로 보이는가?
- 결과가 명확하지 않을 수도 있지만, 이러한 측정이 적절하게 수행되었는가?

중재가 비용, 품질, 환자, 임상 및 의료진의 경험을 변화시키는 결과로 보이는가? (평가의 길이와 포함범위에 따라 다를 수 있음)

결과를 어떻게 활용할 수 있는가?
(5장 참고)

중재가 무엇을 달성했고 무엇을 하지 못했는지에 대한 증거를 확보할 것:

- 중재가 누구를 위한 것이었나?
- 전달을 변화시켰는가?
- 모범 사례들
- 팀 구성원을 위한 최상의 인력과 역할
- (영향 분석이 가능한 경우) 세팅 및 환자의 하위 그룹별로 진료 실행 및 영향이 어떻게 다른가?

결과가 유사한 경우, 결과를 통해 이 중재의 상대적 비용과 이익을 다른 중재의 비용과 비교할 수 있다.

결과는 다음과 같은 경우를 만들는데 도움이 될 수 있다:

- 중재에 지속적으로 자금을 지원하여 중재를 개선시킬 것.
- 다른 세팅에도 중재를 확산시킬 것.

일차의료 중재 평가 시 공통적인 과제

- 주어진 시간이 너무 짧거나 중재가 너무 사소하여 의료전달 및 결과의 변화를 관찰할 수 없다.
- 사례가 적은 경우, 클러스터링으로 인해 영향을 통계적으로 감지하기 어렵다.
- 데이터가 제한되거나 품질이 낮거나, 처리 시간이 너무 오래 걸린다.
- 중재 진료의 특성으로 인해 결과를 일반화할 수 없다. (예: 참가자가 열리 어렵터일 수 있음)
- 결과는 중재와는 무관한 이유로 인해 개선되거나 쇠퇴할 수 있으며, 비교 그룹 또는 평가 설계가 이를 적절하게 설명하지 못할 수 있다.
- 중재가 실행 되기 전, 중재 진료와 비교 진료 간에는 기본적으로 차이가 있다.
- 사례 비교를 통해 중재의 형태 또는 수준을 정할 수 있다.

II. 평가를 위해 무엇이 필요한가?

평가를 성공적으로 계획하고 수행하는 데 중요한 단계는 필요한 자원을 이해하고 확보하는 것이다. 중재와 평가를 수행하는 데 필요한 자원은 [표 1] 및 [그림 1]에 제시되어 있다. 평가를 위해 초기 계획 단계에서 이러한 목록을 준비하는 것이 좋다. 기관의 고위 경영진 등의 사람들은 필요한 리소스를 확인하고 지원하는 것이 필요할 수도 있다.

중재에 사용할 수 있는 자원은 평가와 관계가 깊다. (1) 의료기관이 케어를 변화시킬 수 있는 범위와 (2) 기대 효과의 크기에 영향을 끼치기 때문이다. 이를 파악하기 위해 다음과 같은 질문을 할 수 있다. 얼마나 많은 진료가 변화될 수 있는가? 변화를 실현하기 위해 의료진에게 얼마나 많은 시간이 필요한가? 변화를 이끌어내는 재정적, 기술적 지원(공동의사결정도구 또는 환자 등록부 개발 지원 등)과 도구가 필요할 것인가? 신규 또는 기존 파트너십을 통해 추가로 자원을 사용할 수 있는가? 이 중재 패키지는 결과의 변화를 기대하게 할 만큼 충분한가? 마지막으로, 진료가 의료 제공을 변화시키고 이 변화가 결과를 개선하는데 시간이 얼마나 걸릴 것인가?

마찬가지로, **평가에 사용할 수 있는 자원**은 평가의 엄격성과 깊이를 결정하는데 도움이 된다. 평가를 진행하려면 데이터, 연구 기술, 전문 지식 및 재정 자원이 필요하다. 기관은 내부적으로 활용할 수 있는 기술과 전문성에 따라 내부 직원이 평가를 진행하거나, 외부의 평가자를 고용하여 평가 또는 공동 작업을 진행하고 설계와 분석에 대한 지침을 제공할 수 있다. 외부의 평가자는 종종 평가의 전문성과 객관성을 보장할 수 있다. 내부 및 외부의 전문가 또는 둘 다에 의해 평가를 진행하는지 여부에 관계없이 내부 직원의 평가(예를 들어 청구한 데이터를 얻고 인터뷰와 설문 조사에 참여하는 등)에 대한 지속적인 지원이 중요하다. 평가에 사용할 수 있는 시간은 데이터 수집에 필요한 시간뿐만 아니라 결과가 변경되는 데 필요한 시간 때문에 측정할 수 있는 결과에 영향을 미친다.

• 평가에 전념할 수 있도록 재정 및 연구 자원과 데이터 리소스를 목록화하고 이에 따라 평가를 조정할 것

[표 1] 일차의료 중재를 테스트하는 데 필요한 자원 목록

자원 유형	예시
중재를 위한 자원	
리더십 바이인	중재 시도에 대한 동기 부여와 지원
재정적 자원	중재를 위해 사용할 수 있는 자금(테스트할 수 있는 진료 횟수 포함)
기술적 도움	데이터 피드백, 진료 촉진·코칭, 전문가 자문, 공동학습 및 정보 기술(IT) 전문 지식과 같은 진료를 변화시키는데 도움을 주는 자원
도구	환자 등록부(registries), 의료정보기술 및 공동의사결정도구(shared decision making tools)와 같은 진료를 위한 도구
시간	중재 수행에 할당된 시간, 진료가 변화되고 결과가 바뀌는 데 걸리는 시간
평가를 위한 자원	
리더십 바이인	중재 평가에 대한 동기 부여와 지원
재정적 자원	필요한 경우 외부 평가자를 고용할 수 있는 자금 및 평가를 위한 자금 제공
연구 기술, 전문 지식 및 책임	평가 설계, 데이터 활용, 이행 및 영향 분석 수행 및 결과에서 결론을 도출하는 기술과 전문 지식 데이터를 제공할 임상의 및 직원과 같은 다른 관련 이해당사자와 평가자의 동기 및 바이인 평가 접근법과 분석계획의 설계, 환자와 청구 자료가 포함된 파일 작성과 분석 진행에 대한 전문 지식
데이터	연구 질문에 따라, 데이터에는 청구자료, 전자의무기록, 종이 진료기록부, 환자 섭식 기록, 진료계획, 환자 설문조사, 임상익의 의료진 설문조사, 환자 등록부, 관리 추적 데이터, 현장에서 관찰/인터뷰한 정성적 데이터와 중재 비용 등 데이터의 질은 적정 수준을 유지
시간	데이터를 얻고 분석하며 결과가 바뀌는데 필요한 시간

III. 평가 계획을 어떻게 세울 것인가?

시험조사(pilot test)를 시작하기 전에 평가 방법을 개발할 것

가능한 먼저 평가 계획을 세워야 한다. 면밀하고

시의 적절한 계획은 여러 측면에서 유용한 결과를 도출하는 데 도움을 준다. 첫째, 중재 이전의 데이터를 수집하여 대상자 및 참여 의료진을 선택하는데 도움이 될 수 있다. 둘째, 중재 실행의 초기 경험을 상기하여 해결해야 할 과제와 세부 사항을 이해해야 한다. 마지막으로 평가의 엄격성을 높이기 위해 중재의 시행에 대해 약간의 수정을 제안할 수도 있다. 예를 들면, 기관에서 한 번에 다섯 가지 진료에 한 환자중심메디컬홈(PCMH) 모델을 실행하려는 경우 평가자는 적합성 기준을 충족시키는 사례에서 5가지 진료를 무작위로 선택하도록 제안할 수 있다. 이렇게 하면 의료전달의 변경 내용과 중재 진료의 결과를 나중에 PCMH 모델을 채택할 적절한 진료를 받는 통제 그룹에서의 변화와 비교해 볼 수 있다. 프로젝트에서 평가자와 상의하기 전에 진료방법이 선택되었다면, 평가는 덜 엄격한 비실험적 방법에 의존해야 할 수도 있다.

평가의 목적과 수요자를 고려해야 한다. 평가 결과에 관심이 있는 이해당사자, 평가를 통해 알게 될 것으로 예상되는 결정사항 및 필요한 증거 유형은, 무엇을 질문하고 평가에 어떻게 접근할 것인지를 결정하는 데 중요하다. 예를 들면, 보험회사는 그들이 지급한 총 의료비용에 중재가 미치는 영향과 환자의 만족도에 중점을 둘 수 있고, 고용주는 결근율 및 노동 생산성에 관심을 가질 수 있으며, 일차의료 제공자는 진료 수입·이익과 직원의 만족에 중점을 둘 수 있고, 노동조합은 환자의 기능수준, 만족도 및 본인부담비용에 초점을 맞출 수 있다. 중재의 잠재적 부작용 및 평가로 인한 보고 부담도 고려해야 한다.

평가를 원하는 사람은 누구인가?
수요자들은 어떤 질문에 대한 답을 원하는가?

이해당사자가 필요로 하는 증거의 형식과 엄격성도 고려해야 한다. 방법론적 문제 때문에 결과에 확신을 가질 수 없을 때, 기금 제공자 또는 다른 이해당

사자의 정보 요청에 어떻게 대응해야 할 것인지에 따라 관점은 달라진다. 엄격함과 관련성 간의 상충 관계에 접근하고 이를 조정하는 방법은 계획 단계에서 정하는 것이 좋다. 때로는 가능한 평가 또는 특정 평가 요소에 문제가 많아 (예를 들어 표본수가 적고 검정력의 문제로 비용 데이터를 거의 무의미하게 만드는 경우) 오도될 수도 있는 정보를 생성하는데 리소스를 사용해서는 안 된다.

이때의 질문은 다음과 같다: 이해 관계자가 양적 수치나 질적 설명, 또는 둘 다 필요로 하는가? 이해당사자는 전개되는 모델을 개선하기 위한 지속적인 피드백, 중재가 끝난 시점의 효과 평가 또는 둘 다를 원하는가? 결과는 단기 긍정적인 효과를 암시하는 데 필요한 것인가, 아니면 이해당사자가 그에 맞춰 행동해야 하는데 미치는 강한 영향을 엄격하게 입증해야 하는가? 중재 비용을 정당화하기 위해 효과가 얼마나 커야 하는가? 이러한 문제를 고려하면 측정할 결과, 데이터 수집 방법 및 분석 방법을 선택하는 데 도움이 된다.

일차의로 중재를 평가할 때의 어려움을 파악해야 한다. 일차의로 중재를 평가할 때의 몇 가지 어려움은 다음과 같다.(그림 1의 하단 박스 참고)

▲ **의료(care)를 변화시키는 데 필요한 시간과 강도.** 진료가 변하고 그 변화가 결과를 바꾸는 데는 시간이 걸린다.¹ 많은 연구 결과에 따르면 동기가 부여된 진료가 실제로 의료를 변화시키는 데는 최소한 2~3년이 걸린다.¹²³⁴⁵⁶ 중재가 짧거나 제대로 진행되지 않은 경우에는 결과를 변화시키는 데는 어려울 것이다. 또한, 단기적이거나 부수적인 중재는 알아내기 어려울 정도의 작은 영향이 있는 결과만 조성한다.

▲ **클러스터링이 있을 때 영향 감지력.** 오랫동안 집중적인 중재를 해도 진료 수준에서 결과를 클러스터링 해 보면^d 많은 진료 없이 매우 큰 영향을 미치는 경우를 제외하고는 영향을 감지하기 어려울 수 있다. 예를 들면, 어떤 연구는 PCMH 모델의 비용 효과에 대한 데이터를 수집하고 분석하는 데 많은 시간과 자원을 소비한다. 그러나 클러스터링으로 인해, 너무 적은 진료의 중재는 관측된 변화가 통계적으로 유의미하다는 것을 입증할 수 있는 평가를 위해 70% 이상 비용을 절감해서 진행될 수

d. 클러스터링은 진료 중 환자가 다른 환자가 받는 간호와 비슷한 간호를 받지만 그것이 다른 진료에서 환자가 받은 간호와는 다른 경우에 발생한다. 이것은 주어진 진료에서 환자가 반드시 통계적으로 서로 독립적이라고 간주될 수는 없으며 환자 표본의 유효한 크기가 감소된다는 것을 의미한다(원저).

도 있다.⁷ 결과적으로 비용에 대한 예상 효과가 통계적으로 유의미하지 않은 것으로 평가되면 중재가 비효과적이거나 평가의 검정력이 낮았는지 여부가 명확하지 않게 되는 것이다. (검정력을 계산하는 방법은 부록 A 참조)

- ▲ **데이터 수집.** 정확하고 완전하며 시의적절한 데이터를 얻는 것이 어려울 수 있다. 여러 보험자가 중재에 참여하면 데이터를 공유하지 못할 수 있다. 보험자가 없는 경우, 평가자는 일차의료 진료 이외의 서비스 사용과 지출에 관한 자료를 얻지 못할 수도 있다.
- ▲ **일반화 가능성.** 참여하는 진료가 일반적이지 않은 경우, 결과는 다른 진료와 함께 일반화 될 수 없다.
- ▲ **반사실적 상황(counterfactual) 측정하기.** 중재를 수행하는 의원에서 중재가 없는 경우(반사실적 상황) 무슨 일이 발생했을지 알기 어렵다. 시간이 지남에 따라 추세를 변화시키면 적절한 비교 그룹을 확인하지 않고는 중재에 변화가 있다는 것을 판단하기 어려울 수 있다. 또한 여러 중재가 동시에 발생할 수도 있고, 비교 그룹이 중재에서 발견된 것과 유사한 변화를 경험할 수도 있어 평가가 복잡해질 수 있다.

현실적으로 기대치를 조정하고 평가를 리소스와 일치 시켜야 한다. 평가의 목표는 주어진 리소스 내에서 가능한 고품질의 정보를 생성하는 것이다. 일차의료 중재 평가시의 어려움을 감안하면 어떤 질문에 대해서도 명확하고 유효한 답을 제공하지 못하는 광범위한 질문에 대해 연구하기보다는 좁은 범위의 질문에 대한 답을 잘 하도록 하는 것이 좋다. 위에서 설명한 바와 같이, 영향평가를 가치 있게 만들기 위해 중재 및 평가를 위한 충분한 자원이 필요하다.

리소스가 제한된 경우에는 평가를 축소시키는 것이 좋다. 예를 들어, 중재 실행을 이해하고 개선하는 것에 중점을 둔 평가는, 비용 절감과 보건의료 향상을 위한 초기 단계를 판별할 수 있다. 이 경우 환자, 임상, 직원 및 다른 이해관계자의 경험을 이해하기 위해 타겟 인터뷰를 하는 것과 업무 흐름의 변경 및 의료 정보 기술사용과 같은 몇 가지 중간 프로세스만 측정하는 것을 추천한다. 이처럼 초기 단계에서 문제를 발견하게 되면 많은 노력을 들이기 전에 개선이 가능하다.

중재 결과를 개선할 수 있는 이유와 방법을 보여주는 논리 모델 또는 변화이론을 기술해야 한다.

이 단계에서는 평가자와 실행자가 협력하여 중재의 각 구성 요소, 중재 결과에 영향을 줄 수 있는 경로, 향후 특정 기간 내에 예상되는 영향의 유형을 설명한다. 일차의료 중재는 내부 진료 및 외부

평가자와 실행자는 협력하여 변화 이론(theory of change)을 기술해야 함. 논리모델은 무엇을 측정할 것인지, 언제 할 것인지를 설명해 줌.

보건의료 환경의 맥락에서 이루어지기 때문에 중재의 실행에 영향을 줌으로써 결과에 직접적 혹은 간접적으로 영향을 미칠 수 있는 요소를 논리 모델을 통해 식별해야 한다. 모든 요인에 대해 데이터를 수집할 수 없더라도, 각 요인을 제어할 수 있는 충분한 자료를 할 수 없더라도, 영향을 예측하기 위한 회귀 분석에서는 모든 요인을 고려해야 한다. 의원 또는 기관 관련 특정 요인에는 환자의 인구통계 및 언어, 환자 패널의 크기, 의료기관 소유 여부 및 임상사와 직원의 수 및 유형이 포함된다. 진료 관련 요인의 다른 예시로는 진료 리더십 및 팀 워크가 포함된다.⁸ 더 큰 보건의료 환경을 설명하는 요인에는 전문가 및 병원과 같은 다른 진료 제공자의 진료 패턴, 지역 사회의 의료 자원 및 보험자의 지불 방법 등이 포함된다. 중재 구성 요소에는 진료에 따라 달라지는 중재의 측면들, 즉 질과 안전에 체계적으로 중점을 둔, 환자 중심의 포괄적이며, 조정된, 접근 가능한 진료를 제공하기 위해 전달되는 서비스의 유형 및 양과 같은 정보가 포함되어야 한다. 또한 중재 진료에 따라 달라지는 조치가 포함될 수도 있는데, 이는 진료 개선을 돕기 위한 기술 지원, 의료 제공자와 진료에 대한 추가 지불, 그리고 보건의료 활용도, 질, 비용 산정 등과 같은 선택된 환자의 예후에 대한 정기적 피드백을 주고받는 것 등을 말한다.

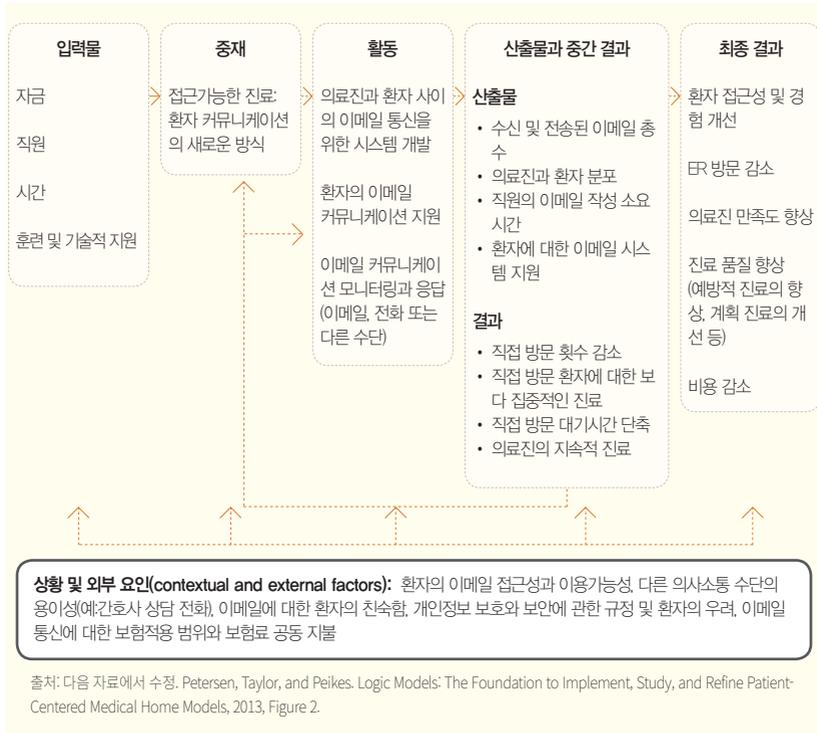
실행 및 연구에 영향을 줄만한 논리 모델 및 도구에 대한 리소스 찾기

논리 모델은 여러 가지 용도로 사용 된다(그림 및 참고 자료는 Petersen, Taylor 및 Peikes 참조⁹). 실행자들이 진료 변화에 대한 논리적 갭을 조기에 인식하도록 도와줌으로써 성공을 보장하기 위해 중재를 수정할 수 있는 적절한 조치를 취할 수 있다. 평가자는 처음부터 논리 모델 접근법을 사용하여 중재에 강력한 기본 논리와 결과를 개선할 수 있는 가능성이 있는지 여부와, 중재가 통계적으로 유의미한 결과를 도출하기 위해 만들어내야 할 효과의 크기를 결정

할 수 있다. 또한 논리 모델은 중재가 의도된 대로 실행되었는지, 개선된 결과와 의도하지 않은 결과를 도출했는지 여부를 보여주기 위해 다른 시점에서 무엇을 측정할지를 결정하는 것과 중재 실행에 대한 촉진 및 제약 요인을 파악하는 데 도움이 된다. 그러나 예상치 못한 정보에 대해서도 개방적이어야 한다. 마지막으로, 중재가 어떻게 진행되는지에 대한 논리 모델의 묘사는 결과를 해석하는 데 도움이 될 수 있다. 예를 들면, 중재가 어려움을 겪고 있는 진료에 대한 더 많은 지원을 목표로 한다면, 결과는 더 많은 지원과 좋지 못한 결과 사이의 상관관계를 보여줄 수 있다. 중재 접근법의 구체적인 내용을 이해하게 되면 기술 지원이 결과를 악화 시킨다는 의미로 결과를 잘못 해석하는 것을 방지할 수 있다.

이 프로세스의 하나의 예로, 가정 의료모델에 필요한 기능 중 하나인 접근성 확장을 위해, 중재 실행과 결과를 연결하는 경로를 설명하는 것을 고려해야 한다. 접근성의 개선은 의료진과의 지속적인 진료를 개선하고 응급실(ER) 및 다른 진료 장소의 사용을 줄이기 위함이다. 만약 평가대상 중재가 이 접근법을 테스트한다면, 당신은 가정진료에 대한 접근을 어떻게 확장시켜 나갈 것인가를 고려할 수 있다. 연장된 진료 시간에 이메일과 전화로 진료를 요청할 것인가, 아니면 진료 시간 이후에는 당직 간호사나 의사가 필요한가? 환자들에게 새로운 선택권과 그것을 사용하는 방법에 대한 세부 사항을 어떻게 알릴 것인가? 거의 모든 중재가 실행 중에 부분적으로 적용되고 많은 중재가 완전히 제대로 실행되지 않기 때문에, 논리 모델에서는 진료에서 실행한 접근법을 문서화하기 위해 과정지표(process indicator)를 명시해야 한다. 이메일을 통해 접근성을 개선한 진료의 몇 가지 과정지표에는 우편 또는 방문 중 선택에 대해 안내를 받은 환자의 수, 다른 의료진과 주고받는 이메일의 총 수, 의료진 전체 수와 환자 한 명당 의료진 수, 의료진이 이메일을 시작하고 응답하는데 소요된 시간 등이 포함된다(그림 2). 사용 가능한 데이터 시스템과 새로운 시스템의 설정 가능성에 따라 어떤 과정지표가 가장 쉽게 수집될 수 있는지를 평가할 수 있다. 수집 방법을 정하려면 의도한 결과에 도달하기 위해 필수 활동을 반영할 수 있는 요소를 고려하고, 데이터 수집에 필요한 리소스와 환자 진료와 의료진의 업무 흐름에 미치는 영향에 따라 균형이 되어야 한다.

[그림2] 이메일 통신과 관련된 PCMH 전략의 논리 모델



향상된 이메일 교신으로 예상되는 **중간결과(intermediate outcomes)**에는 다음 내용이 포함될 수 있다.

- ▲ 환자가 이메일을 통해 간단한 문제를 해결할 시 직접 방문은 감소하나 직접 방문 환자에 대한 더 집중적인 진료 수행
 - ▲ 직접 방문 예약을 위한 대기 시간 단축
 - ▲ 환자가 선호하는 의료진과의 직접 방문 진료를 예약하지 못해도 다른 의료진에 예약을 할 필요가 없다고 느끼므로 같은 의료진과의 지속적인 진료
- 최종 결과에는 다음 내용이 포함될 수 있다:
- ▲ 더 나은 접근성과 경험을 보고하는 환자
 - ▲ 환자가 문제를 느낄 때 의료진이 더 신속하게 중재를 할 수 있기 때문에 ER 방문 횟수 감소

- ▲ 양질의 진료를 제공하고 보다 깊이 있는 직접 방문 진료가 가능함으로 인해 의료진의 경험 향상
- ▲ 연속성, 접근성 및 품질 향상으로 인한 비용 절감

프로그램 기획자가 의도하지 않았지만 발생할 수 있는 결과 또한 추적해야 한다. 예를 들면 보험사가 이메일 교신에 대해 보험금을 지급하지 않고 직접 방문하는 환자 수가 줄어들면 이메일 사용이 진료 수익을 줄일 수 있다. 그 대신, 이메일의 사용 증가는 직접방문이 없으므로 임상이가 핵심 정보를 놓치게 되어 진료과실을 일으키거나, 의료진이 환자와 함께 보내는 전체 시간이 늘어나서 직원의 피로로 이어질 수 있다.

이메일 교신의 기능에 영향을 줄 수 있는 상황별 요인이 의도하거나 의도하지 않은 결과에 영향을 줄 수 있는 가능성에는 다음의 내용이 포함된다. 환자가 이메일에 접속해서 사용할 수 있는지, 보험회사가 교신에 대해 보험금을 지급할 수 있는지 여부와 기밀성, 개인 정보 보호 및 보안에 관한 규정과 환자의 우려, 그리고 ER 사용에 대한 환자부담금 같은 내용이 포함된다.

어떤 결과 지표에 관심이 있는가? 아래 리소스는 측정 항목을 결정하는데 참고할 수 있다.

- ▲ Commonwealth Fund의 PCMH Evaluators Collaborative는 핵심 결과 측정 목록을 제공한다. www.commonwealthfund.org/Publications/Data-Briefs/2012/May/Measures-Medical-Home.aspx
- ▲ The Agency for Healthcare Research and Quality (AHRQ)에서 개발한 The Consumer Assessment of Healthcare Providers and Systems (CAHPS®) PCMH 조사 도구는 환자 경험 측정 자료를 제공한다.

IV. 평가를 어떻게 수행하며 이를 통해 우리는 무엇을 알 수 있는가?

평가 계획을 완료했다면, 설계, 자료 수집 및 분석 작업을 시작한다.

평가는 크게 실험연구와 영향연구 두 가지로 나뉜다. 두 가지 유형의 연구를 병행하여 포괄적인 결과를 도출할 수 있다. 그러나 진료 사례 수가 적으면 비용과 활용 등 결과에 통계적으로 유의미한 "

영향"을 미치는 효과를 탐지하기가 쉽지 않다. 이러한 경우 규모가 큰 연구 사례를 검토하여 중재가 끼치는 영향에 대해 학습하고, 환경에 가장 알맞은 중재 수행 방법을 이해하기 위해, 실험연구의 설계 및 수행에 초점을 맞추어야 한다.

평가를 설계하고 수행하는데 필요한 자료는 리소스 모음집을 참조

실험연구를 설계하고 수행한다. 일부 평가자는 중재의 영향을 측정하는 데 있어 비용, 품질, 그리고 환자, 가족, 임상 의사의 경험에만 초점을 맞추려 할 수 있다. 그러나 중재가 적용된 방식, 각 진료 사례에서 중재가 얼마나 계획에 충실하게 적용되었는가, 그리고 의도적 또는 우발적인 미준수 사례를 설명하는 요인에 관한 연구를 통해서도 많은 것을 배울 수 있다. 여기에는 사용하는 진료참여 및 중재 서비스를 받는 환자, 진료가 이루어지는 방식에 중재가 가져온 변화, 이것이 의도한 측면과 의도치 않은 측면에서 어떻게 다른지, 그 이유는 무엇인지, 그리고 성공적인 중재와 이득이 되는 결과를 달성하는 데 있어서 장애물과 촉진제는 어떤 것이 있는지에 대한 정보 수집 및 정보 분석이 포함된다.

가능하면 정보 수집 과정을 기존의 업무 흐름과 통합하여 임상 의사와 다른 관계자들의 부담을 최소화해야 한다. 응답자의 부담과 평가 비용 측면에서 각 데이터 소스를 수집하는데 드는 비용을 고려해야 한다. 중재가 시행되는 방식을 연구할 때는 양적, 질적 데이터 모두를 포함시키는 것이 좋다. 실험 연구가 질적 데이터에 크게 의존하는 경향이 있더라도 일부 정량적 데이터 소스(혼합 접근법)를 사용하면 결과의 유용성을 증폭시킬 수 있다.^{10 11 12}

리소스가 집약적 데이터 수집을 허용하지 않는 경우, 중재에 관여하거나 영향을 받는 진료 임상 의, 직원, 환자, 가족과의 토론, 추적 시스템 또는 의료

차트, 그리고 이미 수집된 모든 데이터 분석 및 사용 가능한 문서 검토 등의 방법을 활용할 수 있다.

실험연구는 통찰력을 제공하며 대부분 적은 비용으로 연구가 가능하다.

- ▲ **환자와 그 가족과의 면담과 비공식 토론**은 진료에 관한 그들의 인식과 경험을 알려준다.
- ▲ **진료 임상의 및 직원과의 인터뷰와 토론**은 반정형 토론 안내서를 사용한 중재 과정(임상의, 진료 관리자, 간호사, 의료 보조원, 안내 사무실 관계자 및 기타 직원 포함)에 관한 것으로, 이는 어떻게 그들이(그리고 얼마나 일관성 있게) 중재를 실행했는지, 그리고 그에 관한 일반적인 인식, 그리고 어떻게 환자와의 상호 작용과 일을 변화시켰는지, 그들이 환자 진료와 기타 결과를 효과적으로 개선했는지 여부, 중재가 진료 리더십과 직원으로부터 바이-인(buy-in)을 했는지 여부와 재정적 생존 능력, 그리고 그것의 장점과 개선할 영역에 관한 정보를 제공한다.
- ▲ **추적 시스템에서의 데이터**는 운영상의 이유로 시스템이 이미 설치되어 있다면 저비용으로 수집과 분석을 수행할 수 있다. 추적 시스템은 진료관리, 환자 교육 및 과도기의 지원과 같은 새로운 일차의료 모델에서 일반적으로 사용되는 접근법이 의도한대로 실행되었는지(예를 들어 환자가 병원에서 퇴원하거나 만성질환으로 진단 된 후) 여부를 문서화시킬 수 있다. 추적 시스템이 마련되어 있다면, 연구를 위한 유용성을 크게 향상시키기 위해 시스템을 조정하는 편이 상대적으로 쉽다.
- ▲ **의료 기록 검토**는 표본이 적은 경우 저비용으로 진행할 수 있다. 이러한 검토는 일차의료 중재의 특정 측면이 실행되고 있는지 그리고 얼마나 잘 실행되고 있는지를 설명할 수 있다. 예를 들어, 전자의무기록을 검토하여 임상의가 환자 교육을 제공했거나 진료 계획을 개발하고 논의한 환자의 비율을 결정할 수 있다. 또한 각기 다른 특성을 지닌 환자들에게 영향을 미친 중재의 다양한 요소들의 영향을 보다 광범위하게 볼 수 있다. 만성 질환이 있는 환자, 건강에 문제가 없는 환자, 혹은 체중, 흡연 또는 약물 남용에 대한 교육이 필요한 환자와 같이 구체적인 특성을 지닌 환자를 무작위로 검토하거나 집중할 사례를 선택할 수 있다. 환자 진료 시 생기

는 불필요한 편차뿐만 아니라 의료 서비스 제공자간 진료의 차이 또한 살펴볼 수 있다.

- ▲ **문서 검토**는 교육 지침, 프로토콜, 진료에 대한 피드백 보고서, 환자를 위한 치료 계획 등을 포함하고 있어서 중재의 구성 요소에 대한 중요 세부사항을 제공 할 수 있다.

이 정보는 비교적 저비용으로 수집될 수 있으며 중재를 어떻게 개선할지, 어떤 결과 목표는 달성되었는데 일부는 왜 안 되었는지에 대한 통찰력을 제공할 수 있다.

더 많은 리소스가 있다면, 실행 연구에서 아래의 데이터 소스로부터 데이터를 수집할 수도 있다:

- ▲ **환자와 환자 가족을 대상으로 하는 설문 조사**는 많은 환자 표본으로부터 데이터를 수집하는 데 사용될 수 있다. 설문 조사는 환자가 일차의료 진료 시 받는 치료(접근 가능성, 연속성과 포괄성을 포함)에 대한 내용을 비롯하여 진료가 환자 중심으로 이루어졌는지, 주변 지역 타 의료기관과의 조정은 올바르게 이루어졌는지, 그리고 개선 영역에 대한 질문을 포함할 수 있다.
- ▲ **환자 및 가족을 포함하는** 표적집단은 참여자들이 서로의 생각을 토대로 견해, 문제점 및 아이디어에 대한 토론에 적극적으로 참여하게 한다. 이는 가설을 시험하고 환자진료 문제에 대한 새로운 접근법을 개발하는 데 특히 유용할 수 있다.
- ▲ **진료 현장 방문**은 팀의 기능과, 업무 흐름, 환자와의 상호 작용을 직접 관찰하도록 하여 진료진과의 인터뷰를 보완할 수 있게 한다.
- ▲ **진료 임상과의 직원을 대상으로 한 설문 조사**는 많은 임상과의 직원으로부터 얻은 데이터는 중재가 진료 제공에 어떤 영향을 끼치는가에 관한 데이터를 제공할 수 있다.
- ▲ **더 방대한 환자 표본에 대한 의료기록 검토**는 의료진이 진료를 제공한 방법에 대한 보다 포괄적인 평가를 제공할 수 있다.

분석은 각 연구 질문에 답하기 위해서 여러 출처의 데이터를 종합해야 한다.

여러 소스를 교차해서 정보를 비교 및 대조하는 것은 결과를 강화시키고 실행에 대한 보다 완전한 이해를 이끌어낸다. 데이터 소스가 아닌 답을 얻고자 하는 질문으로 정보를 구성하는 것이 이해당사자에게 가장 유용하다.

중재 기간에 따라 중재의 개선을 위한 실행연구 중간 조사 결과를 사용할 수 있다. 이러한 개선 과정에서 중간 경로 수정과 개선이 허용되나, 중재 그 자체가 시간이 지남에 따라 변화한다는 점을 고려할 때, 이러한 과정은 중재의 영향에 대한 연구를 복잡하게 만든다.

대부분의 소규모 시범 프로그램은 그러한 영향을 감지할 수 있는 충분한 진료경험의 부재로 비용에 미치는 효과를 감지할 수 없다. 적은 수의 진료경험만으로 영향력 있는 연구에 자원을 집중하는 것은 좋은 투자가 아니다.

영향에 관한 연구를 설계하고 실시한다. 대부분

의 이해당사자가 갖고 있는 질문은 “건강관리 비용, 환자, 가족, 임상 및 관계자 경험에 미치는 중재의 영향은 무엇인가?”이다. 이는 영향연구에 중요한 질문이다. 안타깝게도 소수의 진료사례의 경우에는 중재 영향의 평가에 내재하는 통계적 문제로 인해 이 질문들에 답하는데 자원을 투자하지 않는 것이 더 낫다. 평가대상 기관이 이런 종류의 대규모

테스트를 지원하거나 더 적은 수의 진료 패턴이 유사한 진료로도 통계력을 충분히 확보할 수 있다면, 본 장의 몇 가지 지침을 참고할 수 있다. 영향 연구를 수행할 수 있을 만큼 충분한 진료를 변화시키고 있는지 여부를 평가하는 방법을 설명하는 것으로 시작한다.

표본 규모가 납득할 만한 영향을 탐지하기 적합한지와 채택을 장려하기에 충분인지 평가한다. 영향 연구를 시행하려는 경우, 우선은 납득할 만큼 충분히 적당한 효과를 감지하기 충분한 정도로 표본의 규모가 크지를 계산해야 한다. 한편, 만약 영향이 입증되는 경우 이해당사자가 중재를 채택하는 것을 고려할 정도로 충분히 규모가 크지도 계산해야 한다. 통계력에 대한 평가는 실제 진료 내에서 환자 결과의 클러스터링을 설명해야 한다. 대부분의 경우, 일차의료 중재에 대한 평가에는 환자 수에 관계없이 매우 많은 수의 진료가 필요한데, 이는 비용 및 이용 측정에 대한 그럴듯하고 적절한 규모의 효과가 통계적으로도 유의미한 것으로 나타날 것이라고 확신하기 위함이다(부록 A 참고). 관찰 결과의

차이가 중재의 진정한 효과라고 확신할 수 있도록 평가에는(진료 패턴이 매우 유사하지 않은 경우) 50 개 이상의 중재 진료가 포함되어야 할 것이다.

대부분의 연구에서, 검정력 추정(power estimates)은 그 효과가 그럴듯하게 생성될 수 있는 것보다 훨씬 큰 경우가 아니면 소수의 진료에 대한 중재의 영향을 탐지하기 어렵다(예외: 유사한 진료 패턴을 가진 진료의 경우). 이 경우 프로그램 효과(영향)를 평가하고 예측하지 않는 것이 좋다. 실제 프로그램 효과와 자연 변동 간 차이를 구별하기 위한 충분한 수의 진료를 하지 못했을 때(분석이 진료 가운데서 환자 결과의 클러스터링을 정확하게 설명한다면) 중재가 효과가 없었다는 잘못된 결론에 이를 수 있다. 이 경우 실험연구 수행에 집중하는 것이 좋다.

영향 연구를 위해 다음 지침을 고려한다. 영향 연구를 포함하는 통계력이 충분하다면 다음의 사항을 고려해야 한다.

1. 중재가 없을 때 환자가 겪었을 결과를 예측할 수 있는 방법이 있어야 한다.

중재와 관련 없는 추세와 외부 요인이 결과에 영향을 미치기 때문에, 시간의 경과에 따른 추세의 변화를 보는 것만으로는 중재의 영향을 정확하게 식별하지 못할 수도 있다. 회의론자들은 시간에 따른 보건의료비용의 변화가 모든 진료에 영향을 미쳤을 수 있으므로, 그러한 연구는 의심스럽다고 할 것이다. 예를 들어, PCMH로 치료 한 환자의 총 비용이 5% 감소했지만 해당 지역의 모든 사례의 의료비용이 같은 기간에 5% 감소한 경우, 평가는 PCMH가 비용에 의미 있는 영향을 미치지 않았을 것이라고 결론을 내려야 한다.

중재 전에 중재 그룹을 비슷한 비교 그룹과 비교하는 것이 중요하다. 평가는 무작위 또는 비실험 디자인을 사용하여 비교 그룹을 선택할 수 있다. 가능한 경우 무작위 디자인을 사용한다.

따라서 중재를 채택하지 않았다면 진료를 제공하는 방식과 환자의 결과에 어떤 변화를 가져왔을지, 즉 반사실적 상황을 고려해야 한다. 중재 진료와 유사한 진료 그룹 간의 결과 변화를 비교하는 것은 중재의 효과를 다른 요인의 효과와 구분하는데 도움이 된다. 가능한 경우 무작위 또는 실험 디자인을 사용하여 비교 진료 그룹을 선택한다. 무작위 디자인은 결과에 더 많은 확신을 주는데 이것은 가능할 때마다 사용해야 한다. 부록 B에는 비교 그룹 선택에 대한 여러 가

지 접근법에 대한 몇 가지 추가 정보가 나와 있지만, 비교 그룹을 선택하는 데 필요한 여러 가지 고려 사항을 다루지는 않는다는 것을 주의해야 한다.^e

2. 중재를 시작하기 전에 비교 그룹과 중재 그룹이 가능한 유사하도록 한다.

환자중심평가연구소
(Patient Centered Outcomes
Research Institute)[®]에서는
연구 방법론에 유용한 조언
을 제공한다. [www.pcori.org/
assets/2013/11/PCORI-
Methodology-Report.pdf](http://www.pcori.org/assets/2013/11/PCORI-Methodology-Report.pdf)

중재가 시작되기 전 중재 그룹과 비교 그룹이 유사하다면 중재가 두 그룹 간의 결과에 차후의 차이를 야기했다는 것을 더욱 확신할 수 있다. 연령, 성별과 인종, 보험의 출처, 만성적인 조건들, 이전 입원비 지출과 여부, 응급실 방문과 숙련된 간호시설에 거주를 포함하여 유사한 환자 패널을 사용한 비교 진료를 선택해야 한다. 이상적으로 진료 규모, 진료가 독립적인지 또는 더 큰 시스템의 일부인지, 비전문의 직원의 수, 유형 및 역할, 도시 / 농촌 지역 등의 진료 수준의 변수들도 비슷해야 한다. 데이터를 활용할 수 있다면, 더 큰 확신을 위해서, 두 그룹 내 환자들이 유사한 비용곡선을 보였는지를 확인하기 위해서 중재 시작 전 지난 몇 년 동안 두 그룹 모두에서 비슷한 결과가 있었는지를 검토해야 한다. 또한 비용 추이에 원래부터 차이가 있던 경우 그 차이를 위해 통제를 할 수 있다. 이들 모두를 사용하여 비교 그룹을 선택할 수는 없더라도 개연성 및 비교 진료의 비교가능성을 위의 여러 관점에서 검토할 수 있다.

3. 신뢰할만한 분석 방법을 사용하여 프로그램의 영향을 추정한다. 유효한 비교 그룹을 선택하고 충분한 사례가 포함 된 경우, 적절한 분석 방법을 통해 정확한 프로그램의 효과 추정치가 산출된다. 여기에는 이중차이분석(difference-in-difference, 중재 시작 전후 중재 그룹에서의 결과 변화를 비교그룹에서의 동일한 기간 동안 결과의 변화에 비교), 환자 및 진료 수준 변수 제어(위험 보정, risk adjustment), 클러스터링 및 다중 비교를 위한 표준 오차(standard errors) 조정이 포함된다. (부록 C 참조)

4. 다른 결과를 측정하기 위해 여러 표본을 사용하여 리소스를 보존한다. 각 결과에 대한 통계력 계산은 다양한 결과에 대한 데이터 수집에 사용할 표본을 결정하는 데 도움이 된다. 조사 데이터를 수집하는 데 일반적으로 비용이 많이

e. 역자주) 환자 중심성 결과에 초점을 두고 비교 임상 연구(comparative clinical effectiveness research, CER)를 수행하는 연구 기관. 환자, 간병인, 임상인, 고용주, 보험사 및 정책 입안자의 건강 관련 의사결정을 돕기 위해 필요한 근거를 생성함. (출처: PCORI 홈페이지. <https://www.pcori.org>)

든다. 일반적으로(대부분의 설문조사기반 결과에 대해) 평가가 의미 있는 효과를 감지할 수 있다고 확신하기 위해서는 평가 한 번에 환자 20~100 명의 데이터가 필요하다. 더 많은 환자의 설문 조사 데이터를 수집하면 자체 환자의 실제 평균 결과 예측의 정확도가 높아질 수 있지만, 전체적으로는 중재에 대한 예상 효과의 정밀도를 아주 약간 향상(즉, 작은 효과를 감지하는 능력을 향상)시킬 것이다. 일반적으로 표본보다는 모든 진료 환자의 청구 또는 전자건강기록(EHR) 데이터를 분석하는데 추가 비용이 상대적으로 적게 든다. 따라서 가능한 많은 환자의 데이터를 사용하여 클레임 및 전자건강기록 기반 결과를 분석하는 것이 좋다. 반면에, 일부 중재는 고위험군 환자에게 더 큰 영향을 줄 것으로 예상되므로, "고위험"을 어떻게 정의할지를 아는 것과 그러한 기준을 충족시키는 환자에 대한 결과를 개별적으로 분석하는 것이 특정 영향을 탐지하는 능력을 향상시킬 수 있다.¹³

실행 및 영향분석 결과를 종합한다. 대부분의 평가는 많은 정보를 생성한다. 평가에 실행 및 영향 분석이 모두 포함되어있는 경우, 두 가지 유형의 결과를 함께 사용하면 상대적으로 테스트되는 모델의 영향에 대해 훨씬 더 정교하게 이해할 수 있다. 논리 모델에서의 외관에 따라 배열된 두 연구 결과의 연관성을 연구하면 일차적으로 중재가 어떻게 작동하는지 밝히고, 개선 방안을 제안하며, 성공 시 이를 다른 방법으로 전파하는 방법을 고려하는데 도움을 줄 수 있다.

이상적으로는 실행과 영향 분석 작업을 통합 할 수 있을 것이며, 그 후 정기적으로 체계적인 방식으로 정보를 교환할 수 있을 것이다. 이러한 유형의 통합된 접근 방식은 진료 운영, 그리고 사업 성공을 위한 장애물과 촉진제에 관한 통찰력을 제공할 수 있다. 또한 통계적 모델을 활용하여 영향을 테스트하기 위한 가설, 그리고 지리적 영역 또는 진료나 환자 유형별 영향의 차이에 대한 설명을 만들어 내는 것을 도울 수 있다. 이 정보는 시행중인 중재를 개선하고 그것이 만드는 변화의 효과에 대한 사례를 알리는데 사용될 수 있다. 실행과 영향 결과를 동시에 수집하는 경우 이를 사용하여 결과를 검증하고 평가 결론의 증거를 강화할 수 있다. 또한 실행과 영향 분석에서 얻은 정보는 성공적인 중재를 개선하고 전파하는 방법을 이해하는 데 유용하다.

V. 결과를 어떻게 활용할 수 있는가?

본 평가 안내서는 PCMH 모델과 같은 일차의료 중재 평가를 계획하고 수행하기 위한 몇 가지 단계를 설명한다. 특정한 중재를 위한 최고의 평가 설계 및 접근법이 무엇인지는 참여 의원이 선별되는 방식과 테스트되는 중재의 유형은 물론, 평가 목적과 이용 가능한 데이터 및 자원에 달려있다.

잘 설계되고 수행된 평가는 중재가 달성한 것 또는 달성하지 못한 것에 대한 근거가 될 수 있다. 구체적으로, (1) 중재가 누구에게 제공되었는지, (2) 보건의료 전달(health care delivery)을 어떻게 바꾸었는지, (3) 임상과의 직원 경험뿐만 아니라 환자 삶의 질, 비용 및 경험 결과에 미치는 영향을 기술할 수 있다. 평가는 실행과 영향 결과가 진료 설정과 환자 하위 그룹에 따라 어떻게 달라지는지도 식별할 수 있는데, 이는 목표하는 중재에 중요한 영향을 미친다. 마지막으로, 평가는 진료에 나타나는 실행 접근법의 변형에 대한 정보를 사용하여 직원 인력 배치의 모범 사례, 팀 구성원의 역할과 서비스 제공의 특정 접근 방식을 식별할 수 있다.

평가 결과는 이해당사자의 질문에 답변을 제공하고 주요 결정을 내리는데 도움이 될 수 있다. 결과는 고려중인 대체 중재의 결과와 비교될 수 있다. 또한 결과가 좋을 경우 중재를 위한 자금 제공을 계속 유지하고, 개량과 함께 더 크고 지속 가능한 규모로 모델을 전파할 계획을 지원할 수 있을 것이다.

VI. 일차 의료 모델 평가를 위한 리소스 모음집 (resource collection)

이 자료 모음에는 평가자가 논리 모델을 개발, 결과 선택, 설계, 실행 및 결과를 통합하고 합성하는데 사용할 수 있는 자료와 도구가 포함되어 있다.

논리모델(logic models)

Innovation Network, Inc. Logic Model Workbook. Washington, DC: Innovation Network, Inc.; n.d. www.innonet.org/client_docs/File/logic_model_workbook.pdf.

Petersen D, Taylor EF, Peikes D. Logic Models: The Foundation to Implement, Study, and Refine Patient-Centered Medical Home Models. AHRQ Publication No.13-0029-EF. Rockville, MD: Agency for Healthcare Research and Quality; March 2013.

W.K. Kellogg Foundation. Logic Model Development Guide. Battle Creek, MI: W.K. Kellogg Foundation; December 2001:35-48. www.wkkf.org/knowledge-center/resources/2006/02/WK-Kellogg-Foundation-Logic-Model-Development-Guide.aspx

결과(outcomes)

Agency for Healthcare Research and Quality. Consumer Assessment of Healthcare Providers and Systems (CAHPS). Rockville, MD: Agency for Healthcare Research and Quality; 2012.

Rosenthal M, Abrams M, Bitton A, the Patient-Centered Medical Home Evaluators' Collaborative. Recommended core measures for evaluating the patient-centered medical home: cost, utilization, and clinical quality. Commonwealth Fund; May 2012(12). www.commonwealthfund.org/~media/Files/Publications/Data%20Brief/2012/1601_Rosenthal_recommended_core_measures_PCMH_v2.pdf.

실행분석(implementation analysis)

Alexander JA, Heard LR. The science of quality improvement implementation: developing capacity to make a difference. *Med Care*. 2011 December;49(Suppl):S6-20.

Bitton A, Schwartz GR, Stewart EE, et al. Off the hamster wheel? Qualitative evaluation of a payment-linked patient-centered medical home (PCMH) pilot. *Milbank Q*. 2012 September;90(3):484-515. doi: 10.1111/j.1468-0009.2012.00672.x.

Bloom H, ed. *Learning More from Social Experiments*. New York: Russell Sage Foundation; 2006.

Crabtree B, Workgroup Collaborators. *Evaluation of Patient Centered Medical*

Home Practice Transformation Initiatives. Washington: The Commonwealth Fund; November 2010.

Damschroder L, Aron D, Keith R, et al. Fostering implementation of health services research findings into practice: a consolidated framework for advancing implementation science. *Implement Sci.*2009;4(1):50.

Damschroder L, Peikes D, Petersen D. Using Implementation Research to Guide Adaptation, Implementation, and Dissemination of Patient-Centered Medical Home Models. AHRQ Publication No.13-0027-EF. Rockville, MD: Agency for Healthcare Research and Quality; March 2013.

Goldman RE, Borkan J. Anthropological Approaches: Uncovering Unexpected Insights About the Implementation and Outcomes of Patient-Centered Medical Home Models. Rockville, MD: Agency for Healthcare Research and Quality; March 2013. AHRQ Publication No.13-0022-EF.

Nutting PA, Crabtree BF, Stewart EE, et al. Effects of facilitation on practice outcomes in the National Demonstration Project model of the patient-centered medical home. *Ann Fam Med.* 2010;8(Suppl 1):s33-s44.

Plsek PE, Greenhalgh T. Complexity science: the challenge of complexity in health care. *BMJ* 2001;323:625-8.

Potworowski G, Green, LA. Cognitive Task Analysis: Methods to Improve Patient-Centered Medical Home Models by Understanding and Leveraging Its Knowledge Work. AHRQ Publication No.13- 0023-EF. Rockville, MD: Agency for Healthcare Research and Quality; March 2013.

Rossi PH, Lipsey MW, Freeman HE. Evaluation: A Systematic Approach. 7th ed. Thousand Oaks, CA: Sage Publications; 2004.

Stange K, Nutting PA, Miller WL, et al. Defining and measuring the patient-centered medical home. *J Gen Int Med.* June 2010;25(6):601-12. www.ncbi.nlm.nih.gov/pmc/articles/PMC2869425/pdf/11606_2010_Article_1291.pdf.

Wholey JS, Hatry HP, and Newcomer KE, eds. Handbook of Practical Program Evaluation. San Francisco: Jossey-Bass/John Wiley & Sons; 2010

영향분석(impact analysis)

개요

Hickam, D, Totten A, Berg A, et al., eds. The PCORI Methodology Report. November 2013. www.pcori.org/assets/2013/11/PCORI-Methodology-Report.pdf.

Jaen CR, Ferrer RL, Miller WL, et al. Patient outcomes at 26 months in the patient-centered medical home National Demonstration Project. *Ann Fam Med*. 2010;8(Suppl 1):s57-s67.

Meyers D, Peikes D, Dale S, et al. Improving Evaluations of the Medical Home. Rockville, MD: Agency for Healthcare Research and Quality; September 2011. AHRQ Publication No. 11-0091. <http://pcmh.ahrq.gov/page/patient-centered-medical-home-decisionmaker-brief-improving-evaluations-medical-home>.

Orr, L. *Social Experiments: Evaluating Public Programs with Experimental Methods*. Thousand Oaks, CA: Sage Publications; 1999.

Shadish, WR, Cook, TD, Campbell, DT. *Experimental and Quasi-Experimental Designs for General Causal Inference*. Boston: Houghton Mifflin Company; 2001.

스텝웨지(무작위 롤아웃) 디자인 (Stepped wedge(randomized rollout) designs)^f

Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. *Contemp Clin Trials*. 2007;28:182-91. <http://faculty.washington.edu/peterg/Vaccine2006/articles/HusseyHughes.2007.pdf>.

데이터가 클러스터된 경우의 검정력 계산(calculating power when data are clustered)

Peikes D, Dale S, Lundquist E, et al. Building the Evidence Base for the Medical Home: What Sample and Sample Size Do Studies Need? Rockville, MD: Agency for Healthcare Research and Quality; October 2011. AHRQ Publication No. 11-0100-EF. White paper prepared by Mathematica Policy Research under contract no. HHS290200900019I TO2. <http://pcmh.ahrq.gov/page/building-evidence-base-medical-home-what-sample-and-sample-sizes-do-studies-need>.

다중 비교를 위한 조정(adjusting for multiple comparisons)

Schochet, PZ. An Approach for Addressing the Multiple Testing Problem in Social

f 역자주) 여러 그룹들 중 일부에 중재를 먼저 받게 하는 연구 설계로, 먼저 중재를 받은 그룹이 이후에 받은 그룹보다 이벤트가 적게 일어났는지 알아보기 위해 수행한다. 실험 단위가 개인이 아닌 클러스터(그룹) 단위인 경우 수행되며 개인간 변동 요인(correlation)뿐만 아니라 그룹의 변동 요인도 고려한다.

Policy Impact Evaluations. Eval Rev. 2009 December;33(6).

직교 설계(orthogonal design)^g

Collins LM, Murphy S, Strecher V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART). New methods for more potent ehealth interventions. Am J Prev Med. 2007;32:112?18.

Zurovac J, Peikes D, Zutshi A, Brown R. Efficient Orthogonal Designs: Testing the Comparative Effectiveness of Alternative Ways of Implementing Patient-Centered Medical Home Models. AHRQ Publication No.13-0024-EF. Rockville, MD: Agency for Healthcare Research and Quality; March 2013.

성향점수매칭(propensity score matching)

Peikes D, Moreno L, Orzol S. Propensity score matching: A note of caution for evaluators of social programs. Am Statistician 2008;62:222-31. http://econpapers.repec.org/article/besamstat/v_3a62_3ay_3a2008_3am_3a2008_3ap_3a222-231.htm.

혼합적 연구방법(mixed methods)

Creswell JW, Klassen AC, Plano Clark VL, et al., for the Office of Behavioral and Social Sciences Research. Best Practices for Mixed Methods Research in the Health Sciences. Bethesda, MD: National Institutes of Health; August 2011.

Miller WL, Crabtree BF, Harrison MI, et al. Integrating mixed methods in health services and delivery system research. Health Serv Res. 2013 December;48(6 Part II).

Wisdom J, Creswell JW. Mixed Methods: Integrating Quantitative and Qualitative Data Collection and Analysis While Studying Patient-Centered Medical Home Models. AHRQ Publication No.13- 0028-EF. Rockville, MD: Agency for Healthcare Research and Quality; March 2013.

g. 역자주) 증재의 효과를 비교 검증하는 방법으로 변수들의 수준 변화에 따른 효과의 차이를 분리하기 위해 개별 속성들 간의 직교성(orthogonality)을 보장하여 변수 간 상호작용을 최소화하면서 추정하는 디자인으로, (1) 실험 개체 간 동질성이 확보되었고, (2) 상호작용은 무시한다고 가정한다. 단일 연구에서 여러 증재(실험자의 수보다는 적어야 함)의 효과를 검증할 수 있다는 장점이 있음²¹

부록 A: 유의미한 효과를 감지할 수 있는 가능성을 평가하기 위한 클러스터링을 통한 통계능력 계산하기

이 부록에서는 영향 연구를 수행할지 여부를 결정하기 위해 통계력을 계산하는 방법과, 영향력 연구를 수행하는 경우 어떤 결과를 검토할 수 있는지에 대해 설명한다. 그렇게 하기 위해서 우선 효과가 어느 정도 크기인지를 고려한다. '어느 정도의 효과가 이해당사자에게 하여금 중재를 채택하도록 만들 것인가?'라는 질문에 확신된 답이 되기 위해서는 목표로 해야 하는 효과크기(effect size)를 결정하여야 한다. 논리 모델/변화 이론(logic model/theory of change)에 대해 설명할 때 이해당사자들이 의미 있다고 여기는 효과크기에 대한 논의가 필요하다. 투자 수익에 초점을 맞춘 이해당사자들은 중재를 위해 지불하는 추가 지급액을 상쇄하기에 충분할 만큼 큰 비용 절감을 필요로 할 수 있다. 다른 이해당사자들은 환자 경험 평가도 향상을 필요로 할지도 모른다. 그런 다음, 작업해야 하는 시간 범주 안에서 그 효과를 낼 수 있는 가능성을 평가한다. 마지막으로, 평가가 효과크기를 감지할 것이라는 데에 확신을 가지려면 얼마나 많은 수의 진료와 환자가 필요한지를 알아보기 위한 통계 검정력을 추정한다. 변환할 진료의 최대 수를 이미 알고 있는 경우에는 마지막 단계로 건너뛰고 평가 결과가 주어진 효과크기가 통계적으로 유의미하다는 것을 파악할 가능성이 얼마나 높은지 확인할 수 있다. 진료 횟수가 너무 적으면 중재가 효과를 창출했는지 여부를 확실히 판단 할 수 없다.

PCM와 같은 일차의료 중재의 평가에서는 종종 큰 표본이 필요하기 때문에, 다른 설계 접근법을 사용하여 평가가 감지할 수 있는 최소 효과를 추정하는 것이 특히 중요하다. 진료 차원의 중재에서 환자 결과는 진료 범위 내에 집중되고 이 클러스터링은 유효표본크기(effective sample size)를 줄인다. 클러스터링은 동일한 진료 중 다른 환자와 비슷한 치료를 받지만, 임상적이 약품과 기타 요인을 실행하는 방식에 차이가 있을 수 있기 때문에, 다른 진료의 환자가 받은 진료와는 다를 경우에 발생한다. 이는 주어진 의원에서 환자가 반드시 통계적으로 서로 독립적이라고 간주될 수는 없으며, 이는 환자의 유효표본크기를 낮추는 것을 의미한다. 결과적으로, 중재에서 의원의 수(환자 수 아님)는 대부분 평가가 높은 확신으로 감지할 수 있는 최소 효과의 크기를 결정한다. 데이터에서의 클러스터링의 양은 환자와 의원의 표본, 그리고 다른 결과에 따라 다르다. 탐지 가능한 최소 효과를 추정할 때(그리고 결과를 분석 할 때), 데이터가 클러스터된 특성을 고려해야 한다.

위험도가 높은 하위 그룹 구성원 간의 영향을 측정하는 전략은 주어진 중재 진료 및 결과에 대한 데이터의 클러스터링 정도에 따라 영향을 탐지하는 역량을 향상시키는데 도움이 될 수 있다. PCM와 같은 모델은 한 의원의 모든 환자를 대상으로 하지만, 아픈 환자를 연

구하면 여러 이유에서 비용 및 서비스 사용과 같은 지속적인 클레임 기반 결과에 대한 영향을 감지 할 수 있는 역량이 증가 할 수 있다. 첫째, 건강한 환자들은 중재 효과에 관계없이 비교적 적은 입원수 및 서비스 사용도 제한적이어서 의료이용 및 비용을 줄일 수 있는 기회가 거의 없을 것으로 예상된다. 그리고 더 증상이 안 좋은 환자들 사이에서는, 비용 및 서비스 이용 횟수가 더 감소할 것으로 예상된다. 결과적으로, 평가는 그러한 더 아픈 환자들의 표본을 사용할 수 있다. 또한 서비스 이용과 비용은 아픈 환자보다 모든 환자들 사이에서 더 다양하기 때문에 모든 환자들 중에서 이러한 결과에 대한 정기적인 변화와 중재의 효과를 구별하는 것은 종종 더 어렵다. 따라서 다양한 결과 및 표본에 대한 최소 감지 가능한 효과를 추정하면 어떤 환자를 위해 어떤 결과 측정치를 추적할지 결정하는 데 도움이 된다.

부록 B: 중재가 없었다면 일어났을 일을 설명하기 위해 비교 그룹 사용하기

이 부록에서는 무작위 또는 비실험 디자인을 사용하여 비교 그룹을 선택하는 복잡한 문제에 대해 간략하게 설명한다. 목표는 중재 진료와 최대한 유사한 비교 진료 그룹을 식별하는 것이다. 디자인 옵션을 더 잘 이해하기 위해서는 숙련된 평가자와 상담하는 것이 종종 가장 효율적이다. 또한(Orr¹⁴ 또는 Shadish, Cook와 Campbell¹⁵ 같은) 좋은 교본을 통해 배경지식을 얻을 수 있다.

가능하면 무작위 디자인을 선정하기

반사실적 상황을 개발하는 데 있어서 가장 철저하면서도 신뢰할 수 있는 방법은 중재에 관여하는 진료를 중재 또는 대조군에 무작위로 배정하는 것이다. 그러면 대조군은 중재 진료 모형을 채택하지 않았더라면 무슨 일이 일어났을 지를 보여주는 좋은 프록시(proxy)의 예를 제공할 것이다. 그러나 많은 이해당사자들은 윤리적인 이유로 혹은 공정에 관한 이유로 무작위 시험을 실시 할 수 없다고 생각한다. 그런 경우, 중요한 질문은 '진료를 바꾸기 위한 리소스보다 변화에 관심을 두는 진료가 더 많은가?'이다. 변화에 관심을 두는 진료가 많은 경우, 진료를 무작위로 추출하는 모든 일차의로 중재의 효과를 연구하는 견고한 무작위 디자인을 제공하는 두 가지 실용적인 접근 방법을 사용할 수 있다.

참여할 진료를 선택하는 첫 번째 접근 방법은 자원하는 모든 진료 중에서 추천하는 것이다. 추천은, 무작위대조시험으로 추천에 의해 선택된 진료가 중재이며, 선택되지 않은 진료는 대조군으로 사용된다.

또 다른 접근방법은 자원해서 참여하는 모든 진료를 허용하면서도, 시행의 롤아웃(rollout)을 동일선상에 놓지 않는 것이다. 이를 “엇갈린 무작위”(staggered randomized) 또는 “스텝 웨지(stepped wedge)” 디자인 이라고 한다. 중재를 시작하기 전에, 선발주자를 위해 대조군은 후발주자의 역할을 한다.¹⁶ 이 디자인의 장점은 (1) 관심 있는 모든 의원이 참여할 수 있는 기회가 있고, (2) 운영 지원은 시스템에 대한 자원 요구를 감소시키며 한번에 소규모 그룹의 진료에 제공될 수 있다는 점이다. 단점은 후발 주자가 중재를 시작할 때까지 순수한 통제 집단으로 대조군의 역할을 한다는 점이다. 예를 들어, 선발 주자보다 1년 늦게 시작하는 경우, 중재와 대조군 간의 결과를 비교하는 데 1년의 데이터만 사용하게 되는데 이것은 일차 진료 변화와 관련된 잠재적 개선을 실현하기에는 너무 짧은 기간일 수 있다.¹⁷ 그러나 “엇갈린 무작위 디자인”을 사용하여 2년간의 경험이 있는 중재 사례와 1

년간의 경험이 있는 사례를 비교하는 등 여러 단계의 결과를 검토할 수도 있다.

평가에서 무작위로 선택한 디자인을 추가하거나 엇갈린 롤아웃으로 사용하는 경우, 무작위로 진료를 선택하는 것이 중요하다. 의료진이 충실하게 중재를 이행했거나, 환자 결과가 좋거나 나빴기 때문에 중재 집단에 대한 의원을 선택하게 되면 진료의 기존 성과 또는 동기 부여의 효과로부터 중재의 효과를 떼어 내어 파악하기 어렵다. 마찬가지로 “staggered randomized” 디자인에서 진료를 롤아웃 기간으로 무작위로 추출하고, 중재를 시작하기에 더 정교하거나 더 열심히 진료로 시작해야 한다는 충동을 피해야 한다.

이해당사자가 모든 진료에 중재를 도입하고 싶다면, 진료 내에서 중재의 구성 요소를 실현하는 다양한 접근법의 효과를 분석하는 것이 또 다른 옵션이다. 연구의 시작 시, 각 진료는 중재 이행에 대한 여러 접근법의 조합을 얻기 위해 무작위로 분류될 수 있다. 예를 들어, 진료는 사회 복지사나 간호사를 사용하여 진료를 조정할 수 있도록 무작위로 지정될 수 있고, 병원에서 퇴원한지 2 일 이내에 직접 또는 전화로 환자에게 후속 조치를 하도록 무작위로 배정될 수 있다. 직교 설계(Orthogonal Design)으로 불리는 이 접근법은 모든 진료가 최소한 일부 구성 요소를 테스트 할 수 있게 하면서(즉, 어떤 진료도 순수하게 통제될 수 없음) 다른 구성 요소를 전달하는 최선의 방법에 대한 중요한 운영 교훈을 생성한다.^{18,19}

무작위 디자인의 사용이 불가능하면 강력한 비교 그룹 디자인을 사용하기

무작위 디자인의 사용이 불가능할 때도 있다. 이 경우 어떻게 참여 진료가 중재에 참여를 위해 선택한(또는 선택받은) 것과 비실험 비교 그룹을 선택할 때 가능한 한 이들 요인을 모방하는 것이 중요하다. 참여를 이끌어내는 요소에는 기관의 공식적, 비공식적인 선발 기준과 에 의한 의사 결정이 포함된다. 예를 들어, 기관이 중재를 테스트하기 위해 특정 도시의 모든 진료를 선택하는 경우, 비교 그룹은 비슷한 시장 및 환자 구성이 있는 도시의 사례를 포함해야 한다. 특정 건강 IT가 시행된 사례만 선택되었다면, 중재 전 비교 그룹을 위해 유사한 건강 IT와 진료, 상태, 환자 구성 및 결과를 선택해야 한다. 이상적으로, 비교 사례 그룹은 중재 사례와 동일한 특성을 가져야 한다. 비교 그룹을 선택하는 두 가지 대중적인 옵션은 회귀절단모형(regression discontinuity, RD)과 성향점수매칭(propensity score matching, PSM)이다. 그러나 소수의 사례만으로는 PSM과 RD 디자인 모두 중재를 위한 동력이 충분하지 않을 수도 있다.

부록 C: 프로그램 영향평가를 위한 분석 방법의 활용

이 부록에서는 프로그램의 영향을 추정하는 데 도움이 되는 이중차이분석, 환자 및 진료 수준 변수를 제어하는 다변량 분석 수행, 클러스터링 및 다중 비교를 위한 표준오차 조정에 대한 위와 같은 분석 방법을 간략하게 설명한다.

이 방법들은 내용이 복잡한데, 아래의 요약은 이 방법들을 구성하는 일반적인 개념을 소개하기 위한 것이다. 다시 말하자면 평가 질문 및 설계에 가장 적합한 분석 방법을 탐색할 때는 숙련된 평가자와 상의하는 것이 가장 효율적이다.

이중차이분석을 사용한 효과 추정

중재 중 동일한 결과의 차이로 인해 중재가 시작되기 전에 중재와 비교 그룹 간의 주어진 결과의 차이를 줄임으로써 프로그램 영향에 대한 이중차분 산정을 추정하는 것이 좋다. 이 접근법은 중재가 일어나지 않았다면, 중재 시작 전 결과의 추세와 두 수준에서 중재와 비교 사례 차이가 중재 후에도 지속되었을 것이라고 가정한다. 예를 들어 그림 2에 설명된 이메일을 통한 접근성 향상 사례의 경우, 중재를 통해 발생한 변화에서 비교군과 같이 시간 경과에 따라 발생한 변화를 제외한 부분을 중재의 영향이라 한다.

환자와 진료 특성의 차이에 대한 통제

중재와 비교 사례의 기존 차이가 결과에 영향을 줄 수 있기 때문에 분석에서 중요한 환자와 진료 수준 변수의 (이전에 설명한) 차이점에 대한 추정을 조정하거나 진료를 교정된 효과로 설정한 뒤(즉, 시간이 지남에 따라 변하지 않는 진료 수준 특성) 다변량 회귀 분석을 사용해야 한다.

클러스터링 및 다중 비교를 위한 표준오차 조정하기

프로그램 효과 추정치의 통계적 유의성을 결정할 때 클러스터링을 고려해야 한다. 클러스터링을 인식하지 못하면 통계적 유의성 테스트를 통해 나타난 중재와 비교 사례 사이의 결과의 차이가 통계적으로 유의미한 것으로 나타날 수 있다. 즉, 클러스터 된 데이터의 특성을 무시하면 위양성(false positive), 즉 존재하지 않는 효과를 찾게 될 수 있다. 마찬가지로, 수많은 결과에 대한 중재의 효과를 검사하면 일부 효과를 우연히 발견할 불확실성이 있다. 따라서 실시하고 있는 검사 횟수에 대해 우연히 예상되는 것보다 통계적으로 유의미한 결과를 찾고 있는 것은 아닌지 여부를 평가해야 한다. 또한 다양한 비교를 위해 표준오차를 조정하는 보다 공식적인 방법들이 존재한다.²⁰

참고문헌

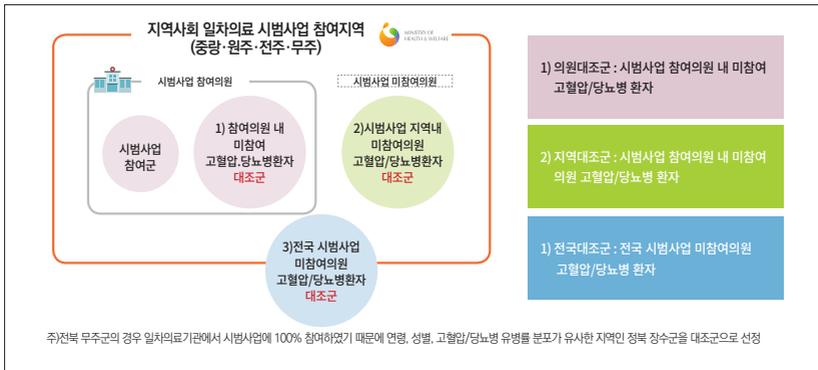
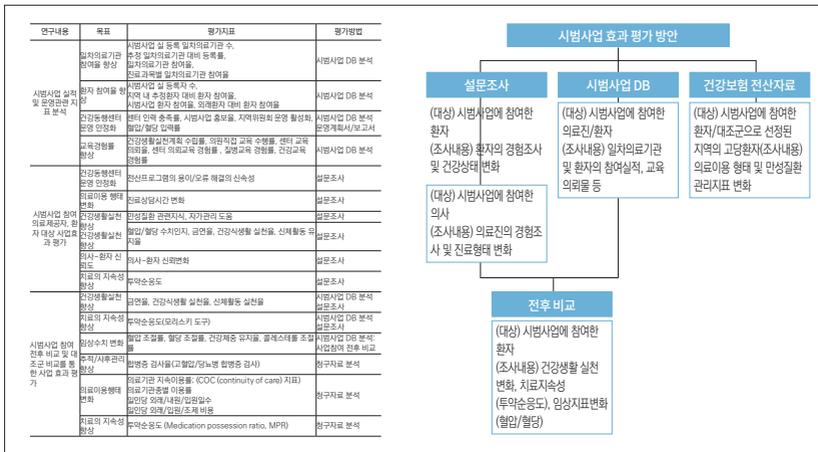
1. McNellis RJ, Genevro JL, Meyers DS. Lessons learned from the study of primary care transformation. *Ann Fam Med*. 2013 May-June;11(Suppl 1):S1-5. doi: 10.1370/afm.1548. PubMed PMID: 23690378; PubMed Central PMCID: PMC3707240.
2. McMullen CK, Schneider J, Firemark A, et al. Cultivating engaged leadership through a learning collaborative: lessons from primary care renewal in Oregon safety net clinics. *Ann Fam Med*. 2013 May-June;11(Suppl 1):S34-40. doi: 10.1370/afm.1489. PubMed PMID: 23690384; PubMed Central PMCID: PMC3707245.
3. Solberg LI, Crain AL, Tillema J, et al. Medical home transformation: a gradual process and a continuum of attainment. *Ann Fam Med*. 2013 May-June;11(Suppl 1):S108-14. doi: 10.1370/afm.1478. PubMed PMID: 23690379; PubMed Central PMCID: PMC3707254.
4. Nutting PA, Crabtree BF, Miller WL, et al. Transforming physician practices to patient-centered medical homes: lessons from the national demonstration project. *Health Aff (Millwood)*. 2011 March;30(3):439-45. doi: 10.1377/hlthaff.2010.0159. PubMed PMID: 21383361; PubMed Central PMCID: PMC3140061.
5. Reid RJ, Johnson EA, Hsu C, et al. Spreading a medical home redesign: effects on emergency department use and hospital admissions. *Ann Fam Med*. 2013 May-June;11(Suppl 1):S19-26. doi: 10.1370/afm.1476. PubMed PMID: 23690382; PubMed Central PMCID: PMC3707243.
6. Reid RJ, Fishman PA, Yu O, et al. Patient-centered medical home demonstration: a prospective, quasi-experimental, before and after evaluation. *Am J Manag Care*. 2009 September;15(9):e71-87. PubMed PMID: 19728768.
7. Peikes D, Dale S, Lundquist E, et al. Building the Evidence Base for the Medical Home: What Sample and Sample Size Do Studies Need? AHRQ Publication No. 11-0100-EF. Rockville, MD: Agency for Healthcare Research and Quality; October 2011. White paper prepared by Mathematica Policy Research under contract no. HHS2902009000191 TO2. <http://pcmh.ahrq.gov/page/building-evidence-base-medical-home-what-sample-and-sample-size-do-studies-need>.
8. Nutting PA, Crabtree BF, Stewart EE, et al. Effects of facilitation on practice outcomes in the National Demonstration Project model of the patient-centered medical home. *Ann Fam Med*. 2010;8(Suppl 1):s33-s44.
9. Petersen D, Taylor EF, Peikes D. Logic Models: The Foundation to Implement, Study, and Refine Patient-Centered Medical Home Models. AHRQ Publication No.13-0029-EF. Rockville, MD: Agency for Healthcare Research and Quality; March 2013.
10. Wisdom J, Creswell JW. Mixed Methods: Integrating Quantitative and Qualitative Data Collection and Analysis While Studying Patient-Centered Medical Home Models. AHRQ

Publication No.13- 0028-EF. Rockville, MD: Agency for Healthcare Research and Quality; March 2013.

11. Creswell JW, Klassen AC, Plano Clark VL, et al., for the Office of Behavioral and Social Sciences Research. Best Practices for Mixed Methods Research in the Health Sciences. Bethesda, MD: National Institutes of Health; August 2011.
12. Miller WL, Crabtree BF, Harrison MI, et al. Integrating mixed methods in health services and delivery system research. *Health Serv Res.* 2013 December;48(6 Part II).
13. Meyers D, Peikes D, Dale S, et al. Improving Evaluations of the Medical Home. AHRQ Publication No. 11-0091. Rockville, MD: Agency for Healthcare Research and Quality; September 2011. <http://pcmh.ahrq.gov/page/patient-centered-medical-home-decisionmaker-brief-improving-evaluations-medical-home>.
14. Orr, L. *Social Experiments: Evaluating Public Programs with Experimental Methods.* Thousand Oaks, CA: Sage Publications; 1999.
15. Shadish, WR, Cook, TD, Campbell, DT. *Experimental and Quasi-Experimental Designs for General Causal Inference.* Boston: Houghton Mifflin Company; 2001.
16. Hussey MA, Hughes JP. Design and analysis of stepped wedge cluster randomized trials. *Contemp Clin Trials.* 2007;28:182-91. <http://faculty.washington.edu/petereg/Vaccine2006/articles/HusseyHughes.2007.pdf>.
17. Jaéén CR, Ferrer RL, Miller WL, et al. Patient outcomes at 26 months in the patient-centered medical home National Demonstration Project. *Ann Fam Med.* 2010;8(Suppl 1):s57-s67.
18. Zurovac J, Peikes D, Zutshi A, Brown R. Efficient Orthogonal Designs: Testing the Comparative Effectiveness of Alternative Ways of Implementing Patient-Centered Medical Home Models. AHRQ Publication No.13-0024-EF. Rockville, MD: Agency for Healthcare Research and Quality; March 2013.
19. Collins LM, Murphy S, Strecher V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART). New methods for more potent health interventions. *Am J Prev Med.* 2007;32:112-18.
20. Schochet, PZ. An Approach for Addressing the Multiple Testing Problem in Social Policy Impact Evaluations. *Eval Rev.* 2009 December;33(6).
21. Jang, DH. Mutual information as a criterion for evaluating the degree of the orthogonality of nearly orthogonal arrays. *Journal of the Korean Society for Quality Management*, 2008, 36.3: 13-21.)

2. 지역사회 일차의료 시범사업 1차 평가연구

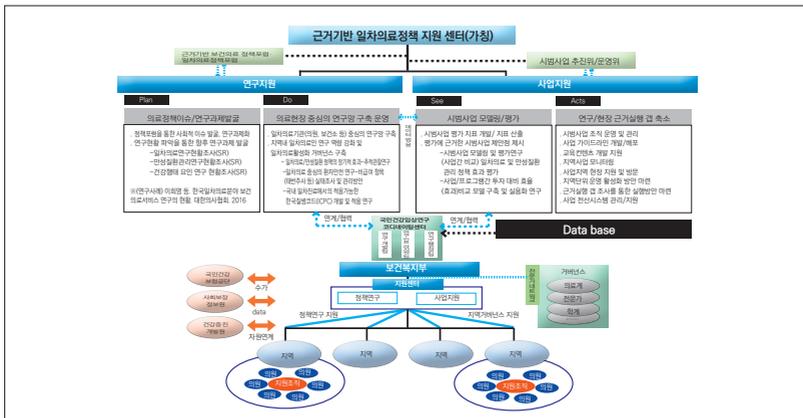
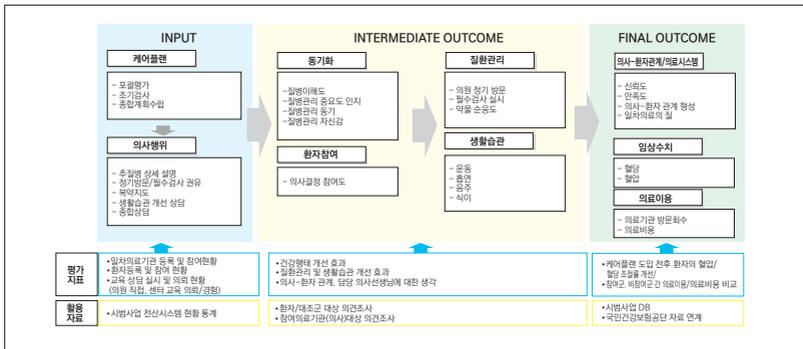
- 연구책임자: 김희선(NECA), 박재현(성균관대학교), 조현호(대한개원내과의사회 보험이사, 중계월내과의원장)
- 연구기간: (1차) 2015.11월 ~ 2016.9월, (2차) 2017.3월 ~ 2017.11월
- 연구목적
 - 시범사업 전·후 비교(3개월, 6개월) 및 체계적인 분석을 통해 시범사업 개선방향 도출
- 연구내용
 - 평가지표 별 산출방안(설문조사, 시범사업 DB, 건강보험공단 자료연계)을 기반으로 사업 효과 평가 실시



* 해당 연구 관련 문의는 연구책임자 또는 한국보건리뷰연구원에 문의하여 주시기 바랍니다.

3. 지역사회 일차의료 시범사업 2차 평가연구

- 연구책임자: 김희선(NECA), 박재현(성균관대학교), 조현호(대한개원내과의사회 보험이사, 중계월내과의원장)
- 연구기간: 2017.3월 ~ 2017.11월
- 연구목적
 - 케어플랜이 추가된 새로운 모형으로 시행 중인 시범사업 전·후 비교(12개월) 및 체계적인 분석(경제적 타당성 평가)을 통해 시범사업의 개선방향 도출
- 연구내용



* 해당 보고서는 국회도서관(www.nanet.go.kr)에서 확인하실 수 있습니다.

일차의료 중재에 관한
리얼월드 평가 안내서

(A guide to real-world evaluations of
primary care interventions)