

A Decision Process of Study Designs for a National Clinical Research Program: Patient-Centered Clinical Research Coordinating Center (PACEN)

Heejung Son¹, Min-Jeong Kim¹, Keunju Yoo¹, Insun Choi¹, In-Hwan Oh^{1,2}, Sang Kil Lee^{1,3}, Jae-Kwan Song^{1,4}, and Dae Seog Heo¹

¹Patient-Centered Clinical Research Coordinating Center (PACEN), National Evidence-based Healthcare Collaborating Agency (NECA), Seoul, Korea

²Department of Preventive Medicine, Kyung Hee University College of Medicine, Seoul, Korea

³Division of Gastroenterology, Department of Internal Medicine, Severance Hospital, Yonsei University College of Medicine, Seoul, Korea

⁴Division of Cardiology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea

공익적 임상연구를 위한 연구설계 결정과정: 환자중심 의료기술 최적화 연구(PACEN) 과제화를 중심으로

손희정¹ · 김민정¹ · 유근주¹ · 최인순¹ · 오인환^{1,2} · 이상길^{1,3} · 송재관^{1,4} · 허대석¹

¹한국보건 의료연구원 환자중심 의료기술 최적화 연구사업단, ²경희대학교 의과대학 예방의학교실,

³연세대학교 의과대학 세브란스병원 소화기내과학교실, ⁴울산대학교 의과대학 서울아산병원 심장내과학교실

Received February 24, 2022

Revised March 29, 2022

Accepted April 26, 2022

Address for Correspondence:

Min-Jeong Kim
Patient-Centered Clinical Research
Coordinating Center (PACEN),
National Evidence-based Healthcare
Collaborating Agency (NECA),
400 Neungdong-ro, Gwangjin-gu,
Seoul 04933, Korea
Tel: +82-2-2174-2881
Fax: +82-2-725-4917
E-mail: bella@neca.re.kr

Health technology assessment (HTA) includes the synthesis of current evidence to support health-care decision-making. In cases of the limited, contradictory, or uncertain evidence, it requires the generation of trustworthy evidence through rigorous health research to support better decision-making. Patient-Centered Clinical Research Coordinating Center (PACEN) is a investigator-initiated research program funded by government, to answer the practical questions raised in real-world settings. It includes systematic process such as topic suggestions from various stakeholders and developing them into research proposals through the reviews of advisory groups involving relevant professionals and the public contributors. The important step in research development is specifying the research questions and determining appropriate study designs to explore the questions. This paper concerns the latter and introduces the process in PACEN that supports the discussion and consensus among the advisory group members to decide study designs for research questions. PACEN supports the four types of evidence-generation studies: randomized controlled trials (RCTs), cluster non-RCTs (cluster non-RCTs), prospective cohorts, and retrospective cohorts. Two diagrams are sequentially applied to determine the appropriate study designs among them once either the evidence synthesis or generation is decided with the first diagram, one of the four study types is decided with the second one. Some cases of utilizing the diagrams in determining of PACEN's clinical research designs are provided. It expects that clinical research developed based on the efforts and collaboration of many stakeholders at PACEN, will be linked in a circular way to HTA and health policy decision-making.

Keywords Health technology assessment; Evidence generation; Patient-Centered Clinical Research Coordinating Center (PACEN); Study designs; Diagrams.

서론

의료기술평가(health technology assessment, 이하 HTA)

는 체계적으로 의료기술의 효과성·안전성·비용효과성을 평가함으로써 보건의료 의사결정을 지원한다. 의료기술평가를 위해 기존 근거를 합성하는 방법(근거합성)은 연관있는

양질의 근거가 가용할 경우에 효율적이다.¹⁾ 그러나 기존 근거가 제한적이고, 상충되고, 불확실성이 큰 경우, 임상연구를 통한 근거생성을 통해 보건의료 의사결정을 지지할 필요가 있다.²⁾

‘공익적 임상연구’는 보건의료현장에서 사용되는 다양한 의료기술의 실제적 효과, 장기 안전성 등 환자에게 중요한 문제를 발굴하고 이를 해결하기 위해 중립적 관점에서 수행되는 국가 지원 연구자 주도 임상연구이다.²⁾ 공익적 임상연구는 의료기술 인·허가를 위한 의뢰자 주도 상업적 임상연구의 형태로는 수행동기가 부족한, 포괄적 인구집단을 대상으로 치료대안 간 효과성·안전성·비용효과성 등을 비교평가하고, 이를 통해 보건의료 현장과 정책적 의사결정에 필요한 ‘근거(evidence)’를 생성한다.³⁾ 이를 통해, 효과적이고 안전한 의료서비스를 제공받을 환자의 권리와 보건의료자원의 효율적인 사용을 옹호하는 등 보건의료서비스의 공공성을 확보할 수 있다.⁴⁾

공익적 임상연구를 지원하는 대표적인 해외기관으로, 영국의 ‘국립보건연구원(National Institute for Health Research, 이하 NIHR)’과 미국의 ‘환자중심성과연구원(Patient-Centered Outcomes Research Institute, 이하 PCORI)’이 있다. 2008년 설립된 NIHR은 국민·환자의 건강 필요와 요구에 부합하는 양질의 근거생성을 지원하며, 연간(2019년 4월-2020년 3월 기준) 310개에 달하는 연구과제를 2.5억 파운드(한화 약 4천억 원) 이상 지원하였다.⁵⁾ PCORI는 2010년 「환자보호 및 적정진료법(Patient Protection and Affordable Care Act)」 하에서 설립된 비영리기관으로, 이해당사자들에게 신뢰할 수 있는 유용한 정보를 제공함으로써 건강 관련 의사결정을 지원하는 것을 목표로 한다.⁶⁾ PCORI는 비교효과연구(comparative effective research)를 지원하며 규모는 2019년 기준 연간 약 3.7억 달러(한화 약 4.2천억 원)에 달하였다.⁷⁾

국내 공익적 임상연구 지원 사업으로서 보건복지부의 재원으로 ‘환자중심 의료기술 최적화 연구사업(Patient-Centered Clinical Research Coordinating Center, 이하 PACEN)’이 2019년 시작되었으며, 한국보건의료연구원(National Evidence-based Healthcare Collaborating Agency, 이하 NECA)이 주관연구기관으로서 총괄 운영하며 지원 규모는 연간 약 230억 원이다.⁸⁾ 동 연구의 목적은 보건의료현장에서 사용되

¹⁾ 근거생성연구는 원자료를 직접 가공하여 지식을 새롭게 생성하는 연구로 통상 일차연구로 불린다. 근거합성연구는 근거생성연구를 통해 생산된 지식을 정량적·정성적으로 종합한 연구로 통상 이차연구로 불린다(박병주 외, 2018, p.31).

⁸⁾ 사업 예비타당성 조사결과 8년간(2019-2026년) 약 1,840억 원의 국고를 지원하기로 하였다(2018년 5월).

는 다양한 의료기술⁹⁾에 대한 효과성·안전성·비용효과성 등에 대한 양질의 근거를 창출하여 보건의료 현장 및 정책과 연계함으로써, 환자에게 최적의 의료서비스를 제공하고 보건의료체계의 지속가능성을 확보하기 위함이다.⁸⁾ 내역사업은 의료기술 간 상대적(비용)효과성에 대한 비교평가연구와, 임상현장에서 활용되고 있지만 안전성·유효성에 대한 의학적 근거가 불충분한 의료기술에 대한 근거생성연구로 구성된다.⁸⁾

NIHR, PCORI 등 공익적 임상연구를 지원하는 기관에서는 사회적 어젠다를 반영하고자 노력을 기울여 왔다. 다양한 이해당사자(국민·환자, 정부·유관기관, 보건의료전문가 등)로부터 광범위하게 보건의료현장의 문제·주제를 확인하고(topic identification),⁴⁾ 연구를 통해 우선적으로 해결이 필요한 주제들을 일련의 절차를 거쳐 선별한다.⁹⁾ 기관별로 약간의 차이는 있지만, 공통적으로 주제확인, 적합성 선별(지원범위, 중복성 등), 주제별 심층검토·과제화 및 주제요약자료 개발(선행 근거 정리, 연구질문 구체화, 연구설계 검토 등), 주제 우선순위 선정 절차로 수행된다.⁹⁾

공익적 임상연구 결과가 환자 개인, 보건의료 실제현장, 정책 등 다양한 맥락에서 건강향상 및 비용절감과 관련된 합리적인 의사결정을 지지하기 위해서는, 높은 수준의 근거(high level of evidence)가 필요하다. 높은 수준의 근거는 중재 및 노출과 건강결과의 인과관계 규명에 있어 비돌림 위험을 최소화한 엄정한 연구설계를 통해 보장된다.¹⁰⁾ 공익적 가치가 높은 주제라 하더라도 근거생성에 있어 연구설계 설정의 오류 및 간과로 인해 신뢰도가 낮은 결과가 도출되면 보건의료기술의 공공성 제고를 위해 결과를 활용하는데 제약이 생긴다. 따라서 공익적 임상연구를 지원하는 기관에서 연구를 통해 높은 수준의 근거가 도출될 수 있도록 연구설계에 대한 명확한 방향성을 제시하는 것은 중요하다.

본고에서 제시하는 PACEN 연구설계 결정 과정은, PACEN 연구주제 검토 및 과제화 과정에서 주제별 연구설계를 결정하기 위한 전문가 논의 및 합의를 지지하고자 마련되었다.⁹⁾ 서두에는 PACEN 지원 임상연구 설계유형을 중심으로 연구

⁹⁾ ‘의료기술’이란 의약품, 의료기기, 진료행위, 기타 보건의료서비스 등 인체의 건강과 생명의 유지·증진에 필요한 상품 및 서비스와 관련되는 기술을 말한다. 이 경우 보건의료서비스는 환자와 일반 국민에게 의료서비스를 전달하는 보건의료 전달체계를 포함한다(환자중심 의료기술 최적화 연구사업 운영관리규정 제1장 제3조1호).

¹⁰⁾ 수요조사 등의 형태로 이해당사자들로부터 연구주제를 제안 받는 것과, 이해당사자들 간 숙의의 과정을 통해 연구주제를 발굴하는 것을 포함한다.

⁹⁾ 환자중심 의료기술 최적화 연구사업에서는 연구주제 제안을 받아 이를 검토하고 과제화하는 과정을 운영한다. 과제화에는 주제별로 다수의 외부전문가 및 국민참여자가 참석하여 논의와 합의를 통해 주요 연구질문과 연구설계를 결정한다.

설계 결정 시 고려할 주요정보를 정리하였다. 또한, 임상연구 유형 중 가장 근거수준이 높은 무작위배정비교임상시험(randomized controlled trials, 이하 RCTs)은 연구목적에 따라 취할 수 있는 두 가지 접근법을 제시하였다. 본문에는 PACEN 연구설계 결정에 필요한 기준과 절차를 두 개의 순차적으로 적용되는 순서도를 기반으로 제시하였다. 포함된 모든 내용은 관련 분야 전문가들의 자문을 받아 최종 완성하였다.

본 론

이론적 배경

연구설계 유형 분류

근거생성을 위한 임상연구 설계 유형은 매우 다양하며, 이를 구분하는 것은 연구결과의 근거 수준을 판단하는 중요한 요소가 된다. 이에, 중복·누락이 없는 분류체계를 수립하고 명확한 기준에 따라 연구설계를 구분하는 도구가 다수 개발되었다. 대표적으로 영국 국립보건임상평가연구원(National Institute for Health and Care Excellence, 이하 NICE)의 연구분류알고리즘¹¹⁾ 및 유사한 도구가 다수 출판된 바 있다.¹²⁻¹⁴⁾ 국내에서는, 2010년 최초 개발된 후 2012년에 한 차례 개정된 'Study Design Algorithm for Medical literature of Intervention (이하 DAMI)'도구가 있다.¹⁵⁾ 이 도구들은 다양한 근거 생성 연구들의 설계 유형을 구분하기 위해 개발된 것으로, 본고에서는 이 도구의 검토·판단 과정을 연구개발 과정에서 연구질문에 적합한 설계를 검토·결정하기 위한 목적으로 적용하였다.

PACEN 연구는 높은 수준의 근거생성을 지원하기 위해 무작위배정비교임상시험(RCTs), 비무작위배정비교임상시험(non-RCTs), 전향적 코호트(prospective cohorts), 후향적 코호트(retrospective cohorts)의 네 가지 유형의 연구를 지원한다. DAMI 도구에 따르면, 이 네 가지 설계 모두 중재와 비교중재에 따라 집단을 나누고 집단 간 주요 결과변수를 비교하며, 연구대상자 배정 가능 여부에 따라 실험연구(배정 가능)와 관찰연구(배정 불가)로 구분된다.

실험연구 중 연구대상자 무작위배정이 가능한 경우 무작위배정비교임상시험, 불가한 경우 비무작위배정비교임상시험으로 구분된다. 다만, 비무작위배정비교임상시험 설계는 연구자의 선호에 의한 중재·비교중재 구분의 편향이 클 수 있어, 군집 수준(의료기관, 시설, 지역 등)에서 적용이 필요한 중재·비교중재를 사용하여 군집 비무작위배정비교임상시험연구(cluster non-RCTs)에 한정한다. 관찰연구로서는 코

호트연구 설계를 지원하는데, 중재 시행과 결과 관찰이 동시에 일어나는 단면연구와 달리, 중재 시행 후 시간을 두고 추적하여 결과변수를 관찰한다. 또한, 비교를 위한 그룹화가 중재 또는 비교중재 노출여부에 따라 이루어져, 그룹화가 결과 발생여부에 따라 이루어지는 환자-대조군 연구와 구분된다. 일반적으로 코호트는 중재와 비교중재에 노출된 시점이 동일한 동시적 코호트를 뜻하며, 노출 시점이 상이한 비동시적 코호트는 지원범위에서 제외된다. 동일한 시점의 비교중재군을 설정하기 어려운 경우, 예외적으로 비동시적 코호트를 설정할 수 있다(예: 현장에서 주로 적용되는 중재가 변경됨에 따라 비교중재군을 과거 시점에서 설정하는 경우 등). 코호트연구 설계는 결과발생 이후 시점에서 이전 자료로 중재·비교중재에 대한 노출과 결과를 순차적으로 관찰한 경우 후향적 코호트, 그렇지 않은 경우 전향적 코호트로 구분된다(Appendix 1).

무작위배정비교임상시험 설계의 두 가지 접근법

무작위배정비교임상시험 연구를 설계하고 수행함에 있어, 연구목적에 따라 설명적(explanatory) 또는 실용적(pragmatic) 접근을 취할 수 있다. 설명적 접근은 이상적인 환경에서 중재의 잠재적 최대 효과를 확인하고 관련 요인에 대한 지식을 생성하기 위함이며, 이를 설명적, 전통적, explanatory RCT, RCT 등으로 일컫는다. 실용적 접근은 현실의 통상적 진료현장에서 중재를 의료서비스로서 제공할 것인지에 대한 의사결정을 지지하는 것이 목적인데 취하며, 실용적, pragmatic RCT, pragmatic clinical trial 등으로 불린다.¹⁶⁾

두 접근법은 각각 다른 목적을 가지고 있으며 연구방법에서도 차이가 있다.¹⁷⁾ 설명적 접근의 경우, 통상 연구기관에서 선정·배제 기준에 따라 엄격하게 정의된 환자군을 연구대상자로서 모집하고, 이들을 관심·비교 중재에 무작위로 배정한다. 연구대상자들은 높은 중재 순응도를 유지하도록 엄격하게 관리되며 연구기관을 방문하여 결과를 측정한다. 이때 측정하는 결과변수들은 대리변수 또는 과정변수인 경우가 많다. 이에 반해 실용적 접근은 다수의 의료기관에서 방문 환자 중 관심 중재 적용이 가능한 환자들을 포괄하여 대표성 있는 연구대상자를 확보한다. 관심 중재 또는 현장에서 사용되는 대안 중재에 무작위배정을 하는데, 환자가 아닌 시점을 배정(stepwise randomization)하거나 의료기관을 배정(cluster randomization)하는 등 실제 의료현장에 적합하도록 다양한 방식을 사용한다. 배정 후 중재 순응도 관리와 자료수집은 통상적인 진료 및 처치 수준에서 벗어나지 않는 선에서 수행된다. 수집되는 자료는 환자의 건강, 삶의 질, 의료비 등 환자와 직접적으로 연관되어, 진료현장 및 정책적

의사결정에 활용할 수 있는 변수를 포함한다. 설명적 또는 실용적 접근은 양분되는 것이 아닌 연속체의 형태로, 모든 무작위배정비교임상시험 연구는 완벽하게 설명적 또는 실용적이라기보다는 경향성을 가진다.¹⁸⁾

무작위배정비교임상시험 연구의 설명적 또는 실용적 성격은 연구의 외적타당도(external validity)와 연관이 있다. 내적타당도(internal validity)가 연구질문에 대한 비뚤림 없는 결과를 낼 수 있음을 의미하는 반면, 외적타당도는 그 결과의 일반화 가능성(generalizability) 또는 적용가능성(applicability)을 뜻한다. 따라서 외적타당도가 높을수록 실용적, 낮을수록 설명적 성격이 강하다.¹⁹⁾ 내적타당도와 외적타당도는 상충될 수 있으므로, 외적타당도를 높이기 위해 실용적 성격이 강한 무작위배정비교임상시험 설계 시 연구의 내적타당도가 감소할 수도 있다.¹⁹⁾ 따라서 둘 간의 균형점을 찾는 것이 중요한데, 일원화된 원칙은 없으므로 실용적 연구 설계 시에는 외적타당도를 높이면서도 내적타당도 감소를 완화하는 방안에 대한 탐색이 필요하다.²⁰⁾ 예를 들자면, 통상적 진료 환경에서 중재 적용으로 인해 연구대상자들에 대한 배정은 폐 및 맹검 유지 가능성이 낮아진다면, 의료기관 단위의 군집 무작위배정을 수행할 수 있다.²⁰⁾ 또한, 통상적 진료자가 결과 평가를 함으로써 평가자의 중립성 유지가 불가능한 경우, 별도 평가자를 고용할 수 있다.²¹⁾ 연구결과와 외적타당도 향상에 대한 일정한 고려 없이 실행 편의에 따라 내적타당도를 타협하는 것은 실용적 무작위배정비교임상시험의 취지에 부합하지 않으므로 주의가 필요하다.

연구계획에 맞게 외적타당도 즉, 적용가능성 수준을 잠정적으로 결정하고 이에 부합하도록 연구를 설계 시, 두 단계로 나누어 검토할 필요가 있다. 먼저, 연구목적이 중재의 최대 효과 확인 또는 진료현장에서 이행여부 결정 중 어디에 가까운지(설명적 또는 실용적 접근법 중 어디에 강조점을 두어야 하는지)를 명확히 한다. 이후 결정된 연구목적에 부합하도록 연구방법을 구체화하고 자가평가를 통해 의도한 적용가능성 수준을 점검한다. 자가평가에 이용할 수 있는 대표적인 도구로서 'Pragmatic-Explanatory Continuum Indicator Summary (PRECIS)-2'가 있는데,¹⁸⁾ 총 9개 항목³⁾에 대해 1-5점으로 평가하여 시각화하며, 점수가 높을수록 실용적 성격·적용가능성이 높다.

³⁾포함기준(eligibility), 연구대상자 모집(recruitment), 환경(setting), 중재 구성(organization), 중재 제공의 유연성(flexibility in delivery), 중재 순응도의 유연성(flexibility in adherence), 추적관찰(follow-up), 주요 결과변수(primary outcome), 주요 결과분석(primary analysis)

PACEN 연구설계 결정 과정

PACEN 연구는 앞서 언급한 네 가지 유형의 근거생성 목적의 임상연구를 지원하므로, 주제검토 시 연구설계는 근거합성 또는 근거생성의 필요성을 먼저 결정한 후, 근거생성이 필요한 주제에 대해 네 가지 설계 중 적절한 하나를 결정하는 두 단계의 순서도를 순차적으로 적용하여 결정한다.

첫 번째 순서도(그림 1)에서는 연구가 필요한 부분을 확인하고 연구방법으로서 근거합성과 근거생성 중 무엇이 보다 적합한지 결정한다. 두 번째 순서도(그림 2)에서는 근거생성이 필요한 것으로 판단된 경우, 무작위배정비교임상시험연구, 비무작위배정비교임상시험연구, 전향적 코호트, 후향적 코호트 중 어떤 것이 적절한지 결정한다. 순서도의 특성상 모든 질문들은 '예' 또는 '아니오'로 답할 수 있도록 구성하였으나, 정도를 판단하여 보다 가까운 쪽으로 검토하는 것이 필요하다. PACEN 연구주제 검토 및 과제화 과정에는 다수의 검토자가 논의를 통해 진행하는데, 검토 대상 주제의 주요 연구질문이 둘 이상이며 각각에 대해 서로 다른 연구방법론 적용이 필요하다고 판단될 경우에는 연구질문별로 분리하여 연구설계를 검토할 수 있다.

본 순서도는 연구설계 유형 분류체계(typology)의 관점에서 개발되었다. 연구설계 결정 외 실제 연구수행 가능성이나 연구결과와 영향력 등 연구결과와 타당성을 보장하기 위한 다른 요소들은 연구주제 개발 과정에서 주제 우선순위 평가 시 고려된다. 따라서 본고의 연구설계 결정 과정을 활용하여 주제별 연구설계를 결정하는데 있어서는 유형 분류체계 관점에서 최선의 결정을 하는 것을 목표로 한다.

근거합성 또는 근거생성 결정 순서도(그림 1)

질문 1. 선행근거의 불확실성 수준이 높은가?

첫 번째 순서도는 두 가지 질문으로 이루어져 있다. 첫 번째 질문은 연구가 필요한 부분을 확인하기 위한 질문으로서, NICE 권고에²²⁾ 따라 선행근거의 불확실성을 검토하여 불확실성이 높은 부분을 파악한다. 불확실성이 낮으면 현재 의사결정을 지지할만한 선행근거가 있다는 것을 뜻하며, 높은 경우 의사결정을 지지할만한 근거가 없어 새로운 연구가 필요함을 뜻한다. 전반적으로 선행근거의 불확실성이 낮다고 판단되면 과제화를 종료한다.

‘예’ 판단 예시:

〈선행근거 없음〉

- 관련 연구가 수행되지 않음

- 관련 연구가 수행되었지만 출판되지 않음

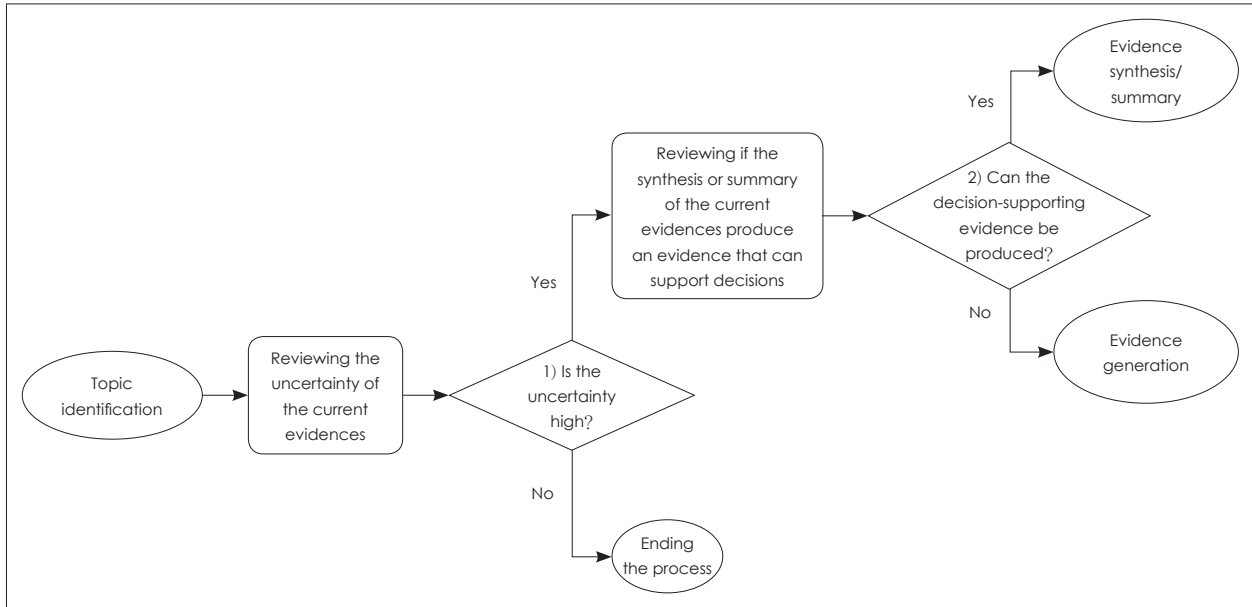


Fig. 1. A diagram to decide evidence-synthesis/summary or evidence-generation.

- 관련 연구가 수행되었고 출판되었지만 확인되지 않음

〈선행연구는 있으나 제한적(불충분·불명확·상충)〉

- 권고의 기반이 되는 정보가 불충분(예: 관심 결과변수에 대한 부적절한 보고)
- 방법론적 제한점: 대상자수가 매우 적어, 연구결과가 우연의 일치에 의한 것이 아님을 통계적으로 확인할 수 없음
- 결과가 결론이 나지 않거나 일관성이 없음
- 결과를 관심 대상 집단에 적용할 수 없음(예: 연구 환경 혹은 사회·문화적 맥락이 비교 가능하지 않음, 사용자 인구집단이 다름, 복용량 등 중재법의 세부사항이 다름)
- 연관성이 있으나 다른 질문(예: 비교중재가 다름)
- 최신의 연구가 아님. 최신 연구를 포함한 체계적 문헌고찰의 업데이트 필요, 진료행태가 변경됨

‘아니오’ 판단 예시:

〈관련 선행근거가 있으며 근거의 수준이 높고 일관됨〉

- 적용 가능한 임상진료지침이 있음
- 연구질문과 대상자, 중재법 및 비교중재법, 주요결과변수 측면에서 일치하는 체계적 문헌고찰/메타분석이 있고, 결과의 권고수준이 높음

질문 2. 선행근거를 활용(합성 및 요약)하여 의사결정을 지지할 결론을 낼 수 있는가?

두 번째 질문을 통해서선 선행근거의 불확실성 수준이 높아 연구를 통해 새로운 근거를 마련하고자 할 때, 근거합성

과 근거생성 중 어느 쪽이 보다 적절한지 판단한다. 이 때 선행근거를 합성·요약함으로써 의사결정을 지지할 결론을 낼 수 있는지 여부를 고려한다.

‘예’ 판단 예시:

〈근거합성·요약 시 불확실성이 현저히 감소함〉

- 선행연구 각각의 방법론적인 제한점에도 불구하고, 선행근거 간 이질성이 크지 않고 일관된 결과가 축적됨
- 연구질문이 일치하는 최신 연구 업데이트가 가능함

‘아니오’ 판단 예시:

〈근거합성·요약이 불가능〉

- 활용할 선행근거가 없음
- 선행연구의 연구질문이 대상자, 중재법 및 비교중재법, 주요결과변수 측면에서 불일치
- 근거합성·요약 시 불확실성이 감소하지 않음
- 선행근거 간 이질성이 크고 결과가 일관되지 않음

이 질문에 있어서는 회색영역(gray zone)이 존재할 수 있다. 연구질문상 충분한 관련성이 있으나 축적된 선행근거가 이질적이고 결과가 일관되지 않을 때 선행근거 활용 가능성에 대한 판단 자체가 어려울 수 있다. 근거합성·요약을 통해 현황을 체계적으로 확인하고 필요한 부분에 대해 근거생성을 순차적으로 하는 것이 방법론적으로 추천되나, 근거의 사회적·정책적 필요성에 따라서 판단이 달라질 수 있다.

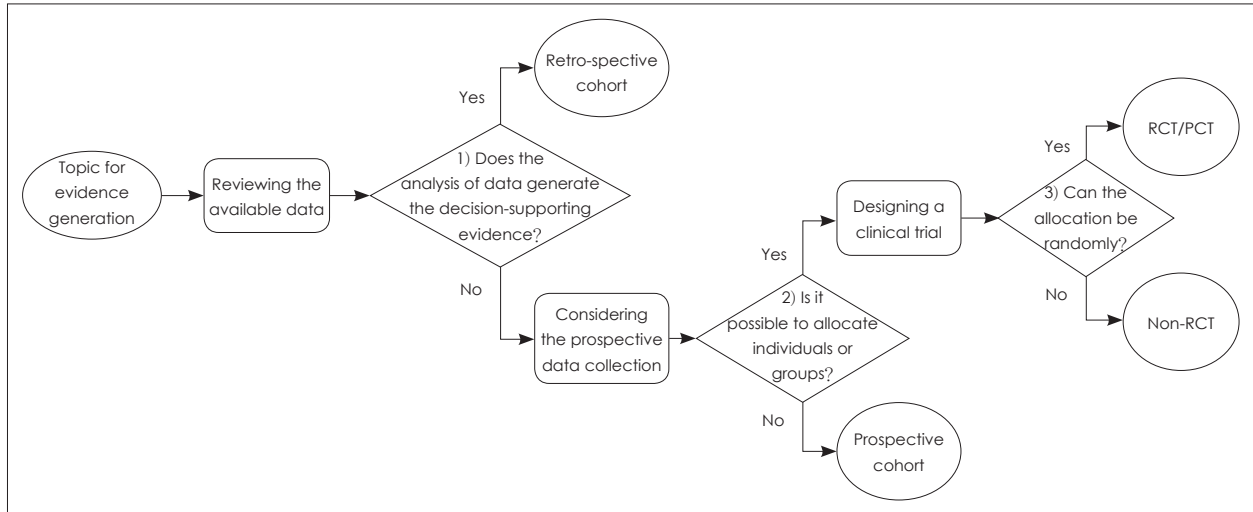


Fig. 2. A diagram to decide one among the four study types for evidence-generation. RCT, randomized controlled trial; PCT, pragmatic clinical trial; Non-RCT; non-randomized controlled trial.

근거생성을 위한 연구설계 결정 순서도(그림 2)

질문 1. 기존 자료원을 활용하여 의사결정을 지지할 근거생성이 가능한가?

첫 번째 질문을 통해서 기존 자료원의 활용가능성을 토대로 전향연구와 후향연구를 구분한다. 이때 기존 자료원은 환자전자기록 및 레지스트리, 건강보험청구자료, 통계청 자료, 국가암등록자료 등 이차자료원을 포함한다. 연구목적에 따라 단독 사용 혹은 둘 이상의 자료를 연계하여 사용하며, 이를 통해 의사결정을 지지하기에 충분한 근거생성이 가능한지 판단한다. 가능한 것으로 판단되면 후향적 코호트 연구를 수행하는 것으로 결정하며, 가능하지 않다면 전향자료 수집을 검토한다.

‘예’ 판단 예시:

- 〈연구질문과 연관성 높은 기존 자료원이 있고 질적·양적으로 충분함〉
- 기존 자료원으로부터 연구대상자 정의, 중재와 비교중재 분류, 주요결과변수 관찰이 가능함
- 필요한 기간 동안 표준화된 형태로 구축되었음

‘아니오’ 판단 예시:

- 〈기존 자료원이 없거나, 있으나 제한적임〉
- 기존 자료원이 존재하지 않음
- 기존 자료원으로 연구대상자 정의를 할 수 없음(예: 자료 입력의 불충분·부정확, 입력 코드가 충분히 세분화되지 않음)
- 기존 자료원으로 중재와 비교중재를 명확히 정의할 수

없음(예: 새로 도입된 서비스 및 프로그램, 시점을 두고 일어나는 경우 에피소드 구분 부정확)

- 기존 자료원으로 주요결과변수 관찰이 어려움(예: 해당 결과변수 관찰이 계획되지 않았거나 자료 구축기간이 짧아 관찰 불가, 관찰 기간 중 결과변수의 자료 형태가 변경되었거나 없어짐)

〈기존 자료원 수집이 불가능〉

- 기존 자료원이 있지만 접근 권한이 없음
- 필요한 자료가 두 개 이상의 자료원에 포함되어 있으나 연계 불가

질문 2. 중재 또는 비교중재에 개인 또는 집단의 배정이 가능한가?

두 번째 질문을 통해서 전향자료를 수집하는 연구 중 관찰연구와 실험연구를 구분한다. 이때 연구자의 의도에 따라 중재군과 비교중재군에 개인(환자) 또는 집단(의료기관 등)을 배정할 수 있는지를 판단한다. 배정이 가능하다면 실험연구로서 임상시험 연구설계를 검토하며, 불가능하다면 관찰연구로서 전향적 코호트연구로 설계를 결정한다.

‘예’ 판단 예시:

- 〈중재 및 비교중재군 배정에 대한 연구자 통제 가능〉
- 연구자가 대상자가 받는 중재 종류, 시점, 과정, 배경을 전적으로 통제할 수 있음

‘아니오’ 판단 예시:

- 〈중재 및 비교중재군 배정에 대한 연구자 통제 불가능〉

- 중재 또는 비교중재가 연구계획 전 이미 수행되었음
- 연구대상자가 중재 또는 비교중재에 자연스럽게 노출됨(지역·의료기관 등 단위가 지정되어 수행되는 시범사업, 제도 등)
- 중재 또는 비교중재가 통상적으로 수행됨(의료필요·요구에 따른 의료처치 등)

질문 3. 무작위배정이 가능한가?

세 번째 질문을 통해 개인 또는 집단의 연구대상자를 무작위로 배정할 수 있는지 판단한다. 개인의 무작위배정이 가능한 경우 무작위배정비교임상시험연구로, 집단으로 무작위배정이 가능한 경우는 군집 무작위배정비교임상시험연구로 결정한다. 불가능한 경우, 군집 비무작위배정비교임상시험연구로 결정한다.

‘예’ 판단 예시:

〈무작위배정 가능〉

- 중재 또는 비교중재군 중 어떤 것을 받더라도 환자의 의료필요·요구·선호 및 가치를 위반하지 않음(환자 동의 동반)

‘아니오’ 판단 예시:

〈무작위배정 불가〉

- 중재와 비교중재의 적응증이 달라 환자의 의료필요·요구에 따라 선택
- 중재와 비교중재를 받는 것이 환자의 생활양식·편의에 영향을 미치므로 선호와 가치에 따라 선택
- (집단 단위) 중재를 수행할 수 있는 집단이 한정적(예: 중재에 필요한 시설·인력을 갖춘 의료기관이 한정적)

〈무작위배정 가능하지만 적절한 방법 사용 불가〉

- 무작위원칙에 부합하는 방법(봉투, 난수표 생성, 소프트웨어 사용) 사용 불가로 부합하지 않는 방법 사용(개인 단위 예: 출생일, 병원 내원일에 따른 배정)

시각화의 복잡성을 피하기 위해 순서도 상에 표현되지는 않았지만 순서도를 통한 연구설계 결정은 반복적(iterative)으로 이루어진다. 모든 단계가 완료된 후, 평가자들은 결정된 연구설계를 통해 연구질문에 의사결정을 지지할 충분한 답을 얻을 수 있는지 검토한다. 답을 얻을 수 있는 경우 연구설계를 확정하고, 답을 얻을 수 없는 경우 이전의 판단으로 돌아가 오류가 없었는지 점검한다. 반복을 통해 다수의 평가자간 논의와 합의가 일어나며 하나의 연구설계로 수렴된다.

PACEN 연구설계 적용 사례

PACEN 연구설계 결정 과정은 2021년 주제개발 과정에서 주제를 검토하는 전문가 집단에 적용되었다. 다음에서 주제개발 과정을 통해 주제별 연구설계가 결정된 사례를 근거합성연구, 무작위배정비교임상시험연구 또는 군집 비무작위배정비교임상시험연구, 전향적 코호트연구, 후향적 코호트연구로 구분하여 제시하였다.

근거합성연구 결정 사례(그림 3)

‘암의 뼈 전이 영상진단검사 비교평가연구’주제의 연구질문은 전문가 검토에 따라 ‘암의 뼈 전이 진단에 있어서 진단 정확도가 높은 검사법은 무엇인가?’, ‘환자의 치료 내용 및 결과 개선에 있어서 효과적인 검사법은 무엇인가?’로 구체화하였다. 먼저 그림 1을 적용하여 선행근거의 불확실성 수준이 높은지(질문 1) 검토 결과, 유방암, 전립선암 등 일부 암종에서 검사법 간 진단정확도의 차이를 관찰한 선행연구가 있었다. 그러나 현재 암의 뼈 전이 영상진단법에 대한 명확한 가이드라인이 부재하며, 특히 검사결과로 인한 치료법 변경, 치료결과 개선의 효과를 관찰한 연구는 제한적이었다. 따라서 선행근거의 불확실성 수준은 높은 것으로 판단하였다. 선행근거를 활용(합성 및 요약)하여 의사결정을 지지할 결론을 낼 수 있는지(질문 2) 검토 결과, 제한적이지만 선행근거가 일부 존재하여 특정 암종에서는 효과성에 대한 판단이 가능할 수 있었다. 이에, 근거합성연구를 수행하여 효과성에 대한 근거를 마련하는 한편, 근거현황을 종합적이고 체계적으로 확인하는 것이 필요하였다. 따라서 동 연구주제는 근거합성연구가 적절한 것으로 판단하였으며, 근거생성을 위한 연구설계 결정 순서도(그림 2)는 추가로 적용하지 않았다.

무작위배정비교임상시험연구 또는 군집 비무작위배정비교임상시험연구 결정 사례(그림 4)

‘노인 환자의 안전을 위한 환자참여 활동 프로그램 개발 및 효과평가 연구’ 주제의 연구질문은 전문가 검토에 따라 ‘노인 환자에서 환자안전을 위한 환자참여 활동 프로그램 수행은 환자안전사건 예방 및 건강결과 향상에 효과적인가?’로 구체화하였다. 그림 1의 선행근거의 불확실성 수준이 높은지(질문 1) 검토 결과, 세계보건기구, 미국 보건의료연구소(Agency for Healthcare Research and Quality, AHRQ) 등에서 환자안전사건 예방을 위한 환자참여가 강조되고 있으며 노인 환자는 환자안전사건에 취약한 집단 중 하나로 선행연구를 통해 잘 알려져 있으나, 노인 환자를 대상으로 한 환자의 안전을 위한 환자참여 활동 프로그램의 안전성 및 효과성에 대한 선행근거는 희소한 것으로 파악되었다. 즉, 선

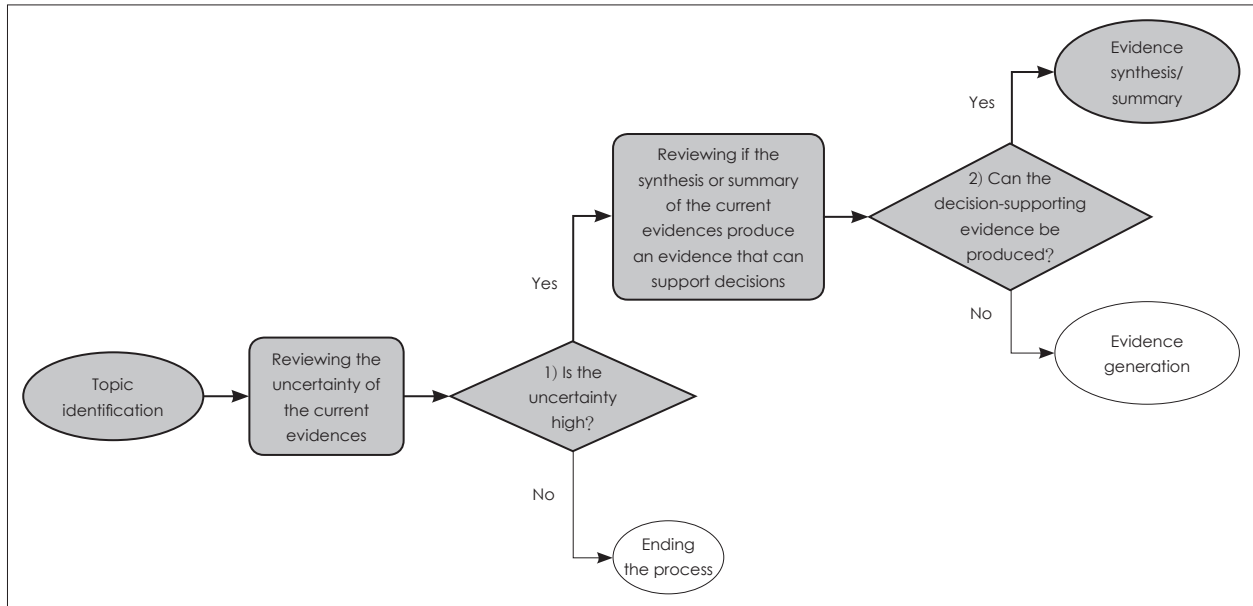


Fig. 3. A decision process of evidence-synthesis study: bone metastasis imaging test. The diagram was presented to decide one between evidence synthesis/summary and generation.

행근거의 불확실성이 높았다. 선행근거를 활용(합성 및 요약)하여 의사결정을 지지할 결론을 낼 수 있는지(질문 2) 검토 결과, 선행근거가 희소하여 근거합성을 통해서 의사결정을 지지할만한 근거 마련 가능성도 낮다고 판단함에 따라 근거생성연구가 적절하였다.

그 다음으로 근거생성을 위한 연구설계 결정 순서도(그림 2)를 적용하여 기존 자료원을 활용하여 의사결정을 지지할 근거생성이 가능한지(질문 1) 검토 결과, 노인 환자의 안전을 위한 환자참여 활동 프로그램은 현재 국내의 진료 현장에서 주로 시행되지 않고 있어, 축적된 기존 자료원이 부재하였다. 따라서 기존 자료원을 활용하여 의사결정을 지지할만한 근거 마련이 어렵다고 판단하였다. 중재 또는 비교중재에 개인 또는 집단의 배정이 가능한지(질문 2) 검토 결과, 전향적으로 데이터 수집을 계획함에 있어 환자(또는 군집)를 환자참여 활동 프로그램(중재군)과 통상적 관리(비교중재)에 배정 가능하다고 판단하였다. 무작위배정이 가능한지(질문 3) 검토 결과, 연구에 참여하는 의료기관 내에서 방문하는 노인 환자를 무작위로 배정하는 것과 의료기관을 무작위로 배정하는 것이 모두 가능하였다. 그러나 후자의 경우, 연구자가 현실적으로 중재군으로 참여할 병원을 모집한 뒤 인력·시설·장비 수준이 비슷한 비교중재군 병원을 모집하는 경우를 고려 시 비무작위로 배정하는 것이 대안적으로 허용되었다. 따라서 본 주제의 연구설계는 무작위배정비교임상시험연구 또는 군집 비무작위배정비교임상시험연구로 결정되었다.

전향적 코호트 결정 사례(그림 5)

‘질식분만 임신부의 산욕기 감염 및 신생아 주산기 감염 예방법의 비교평가연구’ 주제의 연구질문은 전문가 검토 결과 ‘한국인 임신부에서 분만 전 그람양성구균의 군집화 정도는 어떠하며, 감염 여부와 신생아 예후와 어떤 상관관계가 있는가?’로 구체화하였다.⁸⁾ 그림 1을 적용하여 선행근거의 불확실성 수준이 높은지(질문 1) 검토 결과, 그람양성구균은 정상세균총의 하나로 대부분 무증상이지만 신생아와 임신부에서는 현성감염을 일으킬 수 있는데, 임신부를 대상으로 그람양성구균 군집화 정도와 신생아 초기 감염과 예후에 대한 국내 근거는 부족하였다(선행근거의 불확실성 수준이 높았다). 선행근거를 활용(합성 및 요약)하여 의사결정을 지지할 결론을 낼 수 있는지(질문 2) 검토 결과, 선행근거가 희소하여 근거합성을 통해서 의사결정을 지지할만한 근거 마련 가능성도 낮다고 판단하여, 근거생성연구가 적절하였다.

그 다음으로 그림 2의 기존 자료원을 활용하여 의사결정을 지지할 근거생성이 가능한지(질문 1) 검토 결과, 분만 전 임신부의 그람양성구균 집락정도, 임신 주수에 따른 집락의 변화, 신생아 이환율 등의 확인을 위해서는 기존 자료원 이용보다 질식분만 임신부 코호트를 구축하여 관찰 가능하다고 판단하였다. 즉, 기존 자료원을 활용하여 의사결정을 지지할 근거생성이 어려운 것으로 판단하였다. 중재 또는 비교중재에 개인 또는 집단의 배정이 가능한지(질문 2) 검토 결과, 중

⁸⁾복수의 연구질문 중 전향적 코호트로 연구설계가 결정된 연구질문에 한해 기재하였다.

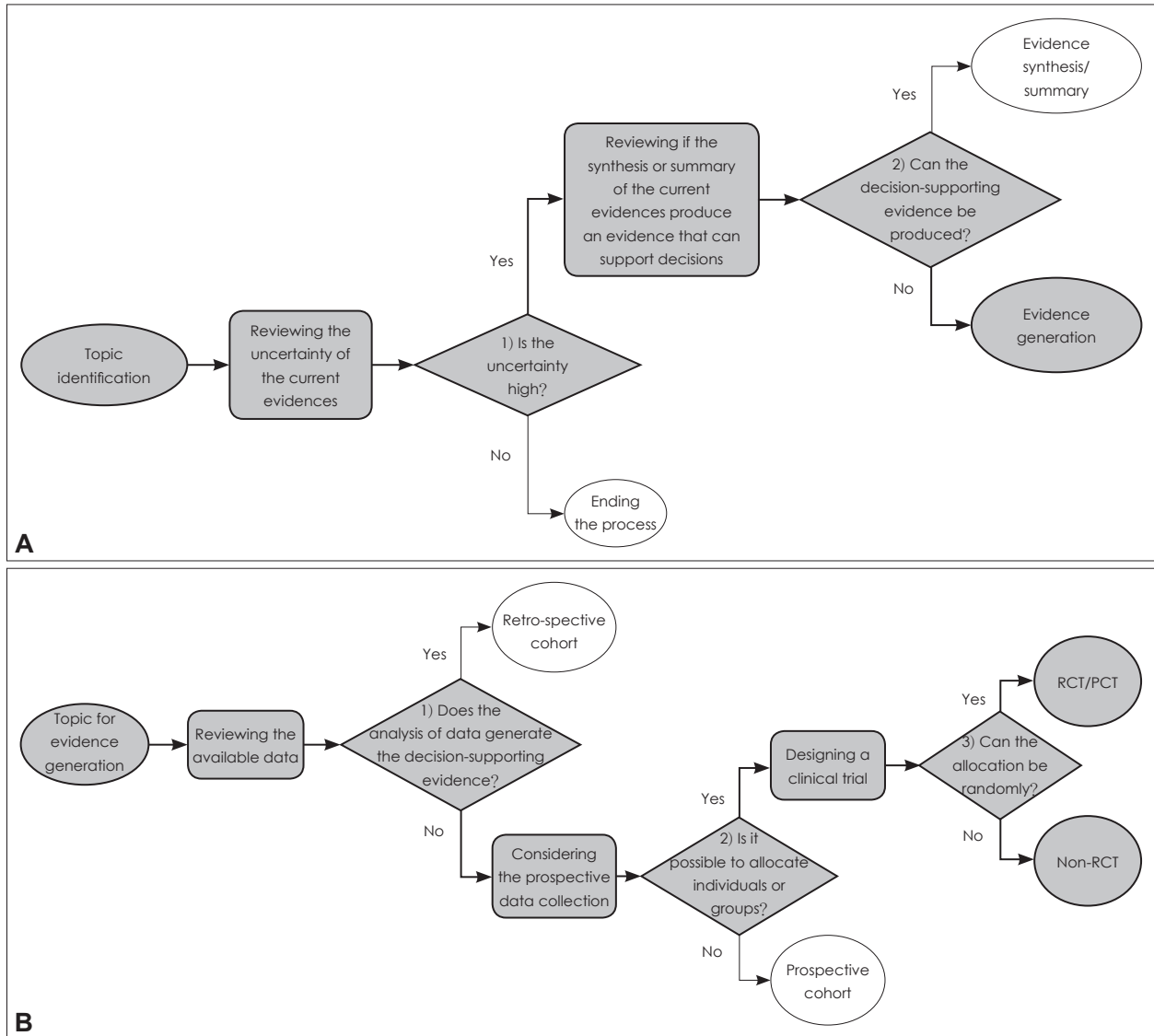


Fig. 4. A decision process of randomized controlled trial study using two consecutive diagrams: development and evaluation of the program for the safety of elderly patients. (A) First diagram to decide evidence generation, (B) second one to decide the best available study design among the four clinical study types.

재 여부가 아니라 노출여부에 따라 군간 비교가 이루어지는 연구로서 연구자에 의한 집단 배정이 불가능하였다. 따라서 ‘질식분만 임신부의 산욕기 감염 및 신생아 주산기 감염 예방법의 비교평가연구’는 전향적 코호트 연구가 적절한 것으로 판단하였다.

후향적 코호트 결정 사례(그림 6)

‘비만을 동반한 2형 당뇨 환자에서 비만대사수술의 비교평가연구’ 주제의 연구질문은 전문가 검토를 통해 ‘낮은 비만도의 2형 당뇨 환자에서 비만대사수술과 최선의 비수술적 치료 중 어떤 치료가 당뇨 관해 및 호전에 더 효과적인가?’, ‘고도 비만의 2형 당뇨 환자에서 당뇨병 개선과 관해율을 높이는

데 더 효과적인 비만대사수술은 무엇인가?’로 구체화되었다. 그림 1의 선행근거의 불확실성 수준이 높은지(질문 1) 검토 결과, 비만을 동반한 2형 당뇨 환자의 비만대사수술에 있어서기에 비해 상대적으로 낮은 비만도를 가진 아시아 국가에서는 혈당 조절이 불량할 경우, 비만도가 높은 환자들에 비해 췌장 베타 세포의 활성이 타진되어 대사수술에 대한 효과가 기대만큼 높지 않을 수 있었다. 따라서 수술과 비수술적 치료 중 최적의 치료를 확인하고 국내 사정에 적절한 비만대사수술 적응증을 확인할 필요가 있었다. 그러나 현재 이에 대해서 선행 근거수준이 낮고 근거 간 이질성이 존재하였다. 고도비만이 있어서 루와이위우회술이 위소매절제술에 비해 우월하거나 비슷하다는 결과가 있었으나, 한국은 높은 위암

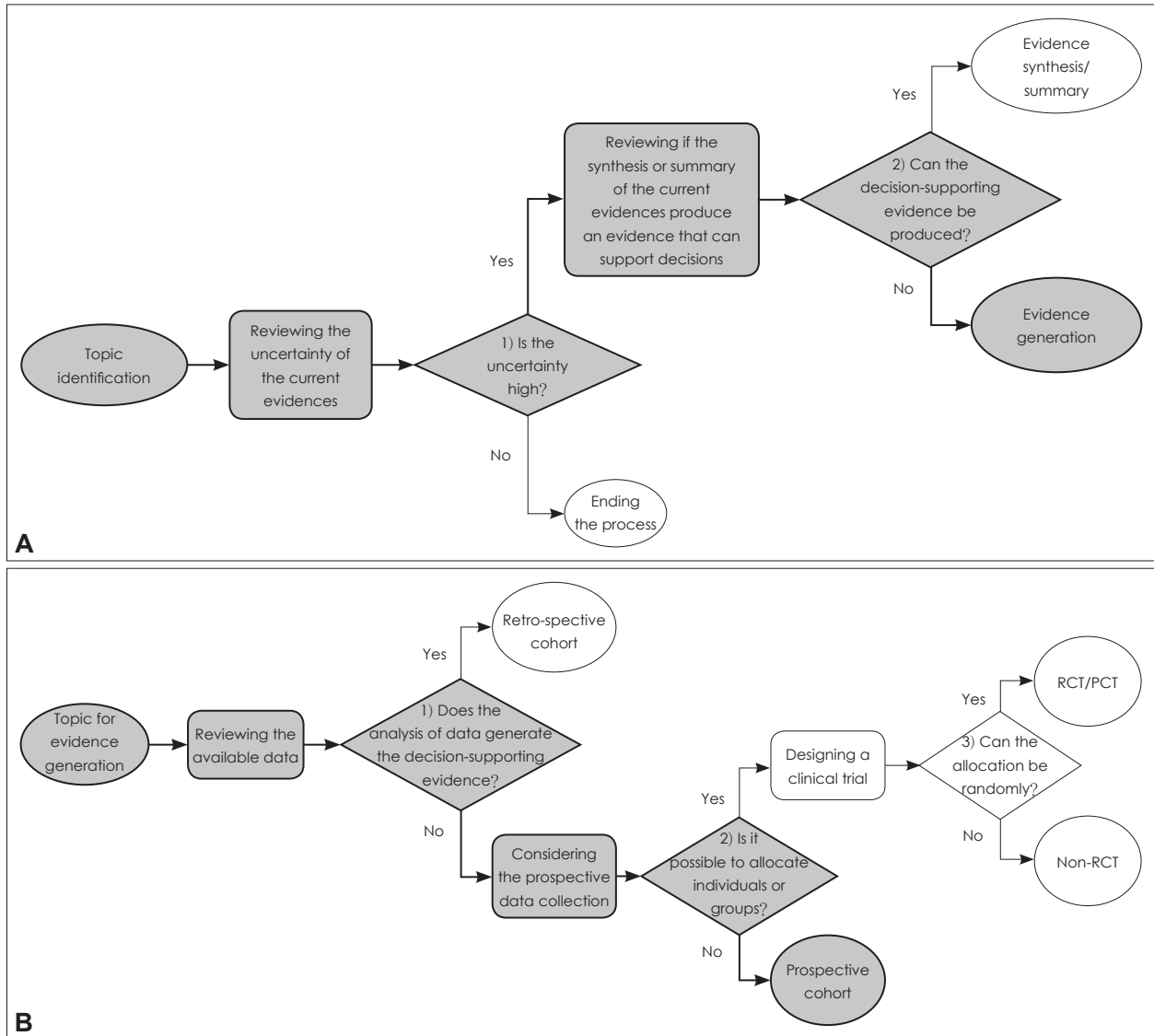


Fig. 5. A decision process of prospective cohort study: prevention for neonatal and maternal infection after vaginal delivery. (A) First diagram to decide evidence generation, (B) second one to decide the best available design among the study types.

발생률로 인해 위소매절제술이 많이 시행되었다. 따라서 국내 실정에 맞는 수술법 확인이 필요하나 한국 등 아시아인 대상 근거가 희박하였다. 따라서 선행근거의 불확실성이 높다고 판단하였다. 선행근거를 활용(합성 및 요약)하여 의사결정을 지지할 결론을 낼 수 있는지 검토 결과, 낮은 비만도의 2형 당뇨병환자 대상 연구질문에 대한 선행 근거수준이 낮고 근거 간 이질성이 존재하며, 고도비만 동반 2형 당뇨병환자 대상 연구질문에 대한 선행근거는 제한적이어서 기존 선행근거를 합성하더라도 불확실성이 감소하지 않을 것으로 판단하여, 근거생성연구가 적절하였다.

그 다음으로 그림 2의 기존 자료원을 활용하여 의사결정을 지지할 근거생성이 가능한지 검토 결과, 근거생성연구 유형을 고려함에 있어, 비만대사수술이 선별급여 의료기술로

서 건강보험청구자료가 축적되어 있었다. 또한, 건강보험청구자료와 의료기관 의무기록자료를 통해 수술 전후의 환자의 임상적 상태에 대한 주요한 정보를 충분히 파악할 수 있을 것으로 판단되었다. 따라서 '비만을 동반한 2형 당뇨병환자에서 비만대사수술의 비교평가연구'는 기존 자료를 이용한 후향적 코호트연구가 적절한 것으로 결정하였다.

PACEN 연구설계 적용 결과 제한점 및 개선방향

2021년도 연구사업 주제개발 시 본고의 연구설계 결정 과정을 적용하여 본 결과, 몇 가지 제한점과 개선사항을 확인하였다.

먼저, 순서도의 질문 중에는 이에 답하기 위해 상당한 사전조사가 필요한 것이 있다. 대표적으로, 근거 합성 또는 생

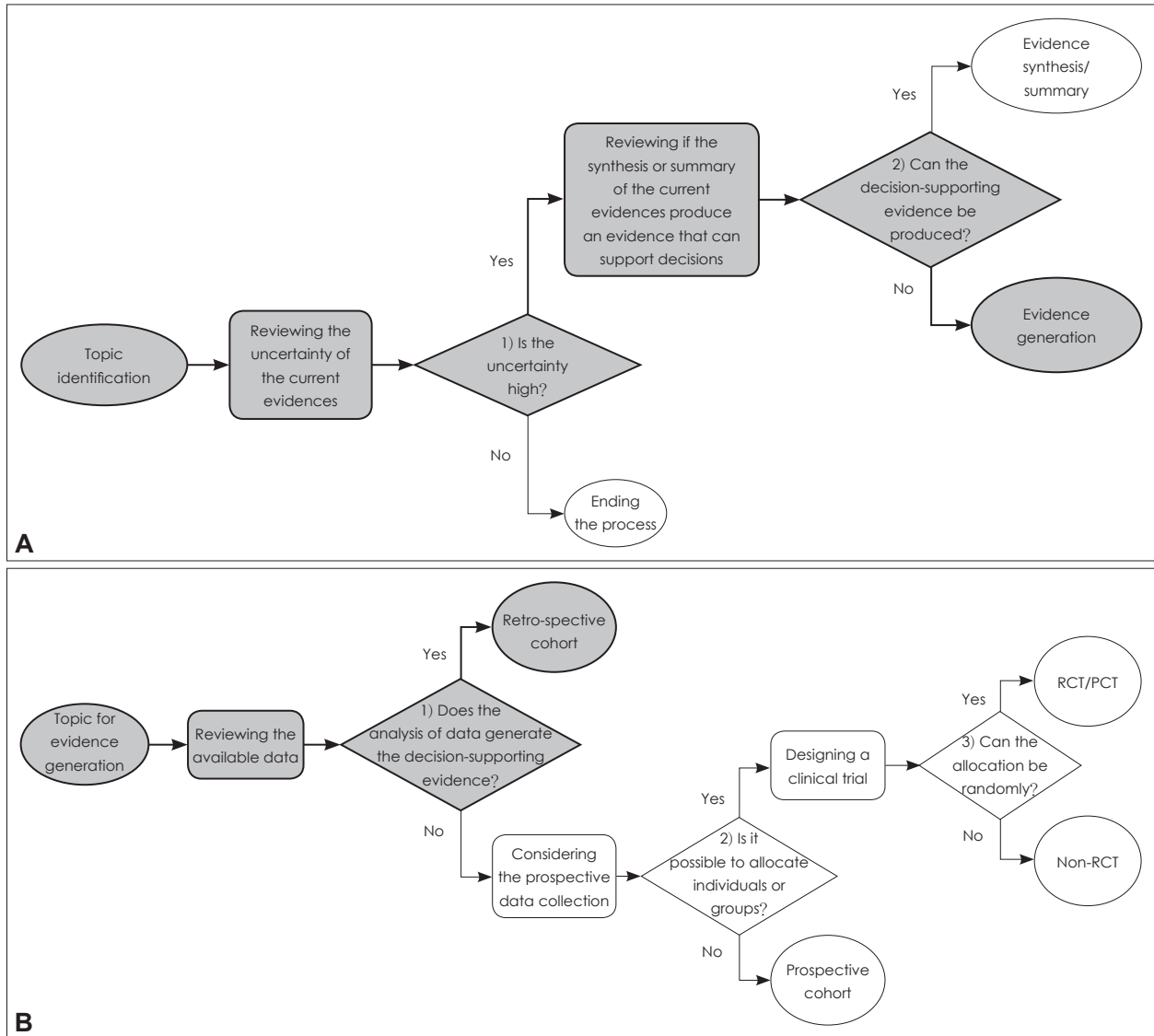


Fig. 6. A decision process of retrospective cohort study: metabolic surgery in type 2 diabetes patients with obesity. (A) First diagram to decide evidence generation, (B) second one to decide the best available design.

성을 결정하는 첫 번째 순서도 질문, '선행근거의 불확실성 수준이 높은가?'는 주제(연구질문)별 선행근거를 파악하는 것이 필요하다. 이 결과에 따라 기존 근거를 합성하는 것이 적절할지 새로운 근거를 생성하는 것이 필요할지 결정을 하게 되는데, 선행근거를 검색하고 검토하는데 소요되는 시간이 사안에 따라 크게 달라진다. PACEN 연구과제화 과정에는 분야별 전문가들이 참여하여 주제와 관련한 선행근거 현황을 효율적으로 파악하고자 하였으며, 사업단은 이를 요약·정리하여 회의자료로 제공하였다.^{b)} 향후 선행근거에 대

^{b)}사업단에서 운영하는 전문가 집단의 연구과제화는 두 차례의 회의를 거쳤다. 1회차에서는 검토대상 연구주제를 검토하고 2회차에서 연구질문 및 연구설계를 확정하였다. 1회차 검토 결과 필요한 근거확인을 사업단에 요청하면, 사업단이 이를 수행하여 회의자료로 정리하여 2회차 회의에 제공하였다.

한 조사·분석을 보다 효율적으로 수행할 수 있는 방안을 마련하는 것이 필요하다.

다음으로, 순서도의 특성 상 질문에 대한 응답을 '예' 또는 '아니오'로 설정하였으나, 포함된 대부분의 질문은 이분법적으로 판단('예' 또는 '아니오')하기보다는 정도를 고려해야 ('예'에 가까운가 '아니오'에 가까운가) 하는 경우가 많다. 따라서 질문별로 제시된 판단 예시에도 불구하고 전문가 간 이견이 발생할 수 있다. 본 순서도에는 2021년 연구과제화 과정에서 관찰된 이견의 내용을 바탕으로 판단 예시를 추가·보완하였다. 향후에는 검토자가 충분한 정보를 바탕으로 논의하여 합의에 이를 수 있도록 지원하는 방안이 필요하겠다.

마지막으로 본 순서도는 연구설계 유형 분류체계의 관점에서 개발되었다. 즉, 순서도에 따라 연구주제별 최선의 연

구설계를 결정하더라도, 타당성 있는 연구결과를 전적으로 보장하기는 어렵다. 연구의 완결성은 연구설계 이외에도 정해진 연구비와 연구기간 내 연구계획에 따른 연구대상자의 원활한 모집, 표준화된 중재·비교중재의 적용, 결과지표의 측정 등 다양한 요소의 영향을 받는다. PACEN 연구개발 과정에서 이들은 과제화 다음 단계인 주제 우선순위 평가 시 고려되었다. 따라서 검토자에게 명확한 안내를 통해 각 단계마다 최선의 결정을 내리는 순차적이고 체계적인 접근을 통해 연구완결 가능성이 높은 공익적 임상연구 주제를 선정할 수 있겠다.

결 론

PACEN에서 지원하는 공익적 임상연구는 사회적 요구에 부합하는 주제를 연구하고 높은 수준의 근거를 마련함으로써 공익적 편익을 높이고자 한다. 이를 위해, 여러 이해당사자들로부터 연구를 통해 해결이 시급한 보건의료 현안 또는 주제를 확인하고, 연구과제로 개발하는 절차를 거친다. 이 과정에서 연구질문이 구체화되고, 연구질문 탐색에 가장 적절하고 연구결과의 설득력의 크기가 큰, 높은 수준의 근거를 마련할 수 있는 연구설계를 결정하게 된다.

본고는 PACEN 연구의 주제 검토 및 과제화 과정에서 연구설계에 대한 논의와 합의를 지지하는 연구설계 결정 과정을 소개하였다. 근거 합성 또는 생성에 대한 결정 및 PACEN 지원 근거생성연구-무작위배정비교임상시험, 군집 비무작위배정비교임상시험연구, 전향적 코호트, 후향적 코호트-에 대해, 주제별로 과제화 과정에 활용하는 기준과 절차, 도구로서 일련의 순서도를 개발하였다. 이를 통해 주제별로 가장 설득력 높은 연구설계를 선택할 수 있도록 지원하였다. 또한, 연구설계상의 주요 쟁점에 초점을 맞추도록 하고 여타 연구완결과 연관된 요소들은 그 다음 주제 우선순위 평가 단계에서 고려하도록 함으로써 연구개발 단계를 체계적으로 진행할 수 있었다. 이에 본고의 연구설계 결정 과정을 향후 과제화 과정에 지속적으로 활용함에 있어, 사용자의 의견수렴을 통해 계속적으로 수정 보완해 나가고자 한다.

끝으로, PACEN의 많은 전문가들의 노력과 협업을 토대로 개발된 공익적 임상연구가 의료기술평가(HTA) 및 보건의료 정책 의사결정에 단계적·순환적으로 연계되기를 희망한다. 공익적 임상연구를 통해 새로운 근거가 마련되면, 그 결과가 의료기술평가를 통한 근거의 통합적 관리 및 확산 단계로 연결되고, 나아가 합리적인 정책의사결정으로 이어지는 것이 바람직하다. 궁극적으로는 이를 통해 효과적이고 안전한 의료서비스 제공을 통해 환자의 건강을 향상시키고 합리적

인 자원 사용을 옹호하는 근거기반 보건의료체계(evidence-based healthcare system)가 지속될 수 있기를 기대한다.

Acknowledgments

This research was supported by a grant of Patient-Centered Clinical Research Coordinating Center (PACEN) funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HI19C0481).

REFERENCES

- 1) McIsaac ML, Goeree R, Brophy JM. Primary data collection in health technology assessment. *Int J Technol Assess Health Care* 2007;23:24-29.
- 2) Kim M, Kim SJ, Lee S, Choi HJ, Lim SE, Park JH, et al. Clinical trials for the public interest: a concept and prospects. Seoul: National Evidence-based Healthcare Collaborating Agency;2019. pp.21-32.
- 3) Kim MJ, Kim SJ, Lee S, HJ Choi, Lim SE, Park JH, et al. Clinical trials for the public interest: a concept and prospects. Seoul: National Evidence-based Healthcare Collaborating Agency (NECA);2019. pp.21-32.
- 4) Heo DS, Yoon YH, Jang YJ, Kim KP, Shin DW, Lee SK. Supporting system modeling of clinical research for improving outcomes. Seoul: National Institute of Health;2007.
- 5) National Institute for Health Research. Annual report 2019/2020. London: National Institute for Health Research;2021.
- 6) Buckley DI, Ansari M, Butler M, Williams C, Chang C. The refinement of topics for systematic reviews: lessons and recommendations from the effective health care program. *Rockville: Agency for Healthcare Research and Quality*;2013.
- 7) Patient-Centered Outcomes Research Institute. 2019 annual report. Washington: Patient-Centered Outcomes Research Institute;2020.
- 8) Regulations on Management of Patient-Centered Clinical Research Coordinating Center (PACEN);Oct 11, 2019.
- 9) Kim MJ, Kang S, Park CY. Research agenda setting processes in the US and UK health technology assessment institutions. *J Health Tech Assess* 2014;2:43-52.
- 10) Park BJ, Heo DS, Lee SM, Ahn HS, Ji SH, Bae EY, et al. Evidence-based healthcare. Seoul: Pakyoungsa;2018. pp.3-118.
- 11) National Institute for Health and Care Excellence. Methods for the development of NICE public health guidance. 3rd ed. London: National Institute for Health and Care Excellence;2012. pp.177-178.
- 12) Hartling L, Bond K, Santaguida PL, Viswanathan M, Dryden DM. Testing a tool for the classification of study designs in systematic reviews of interventions and exposures showed moderate reliability and low accuracy. *J Clin Epidemiol* 2011;64:861-871.
- 13) Zaza S, Wright-De Agüero LK, Briss PA, Truman BI, Hopkins DP, Hennessy MH, et al. Data collection instrument and procedure for systematic reviews in the guide to community preventive services. *Am J Prev Med* 2000;18(1 Suppl):44-74.
- 14) West S, King V, Carey TS, Lohr KN, McKoy N, Sutton SF, et al. Systems to rate the strength of scientific evidence. Rockville: Agency for Healthcare Research and Quality;2002.
- 15) Seo HJ, Kim SY, Lee YJ, Jang BH, Park JE, Sheen SS, et al. A newly developed tool for classifying study designs in systematic reviews of interventions and exposures showed substantial reliability and validity. *J Clin Epidemiol* 2016;70:200-205.
- 16) Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutic trials. *J Chronic Dis* 1967;20:637-648.
- 17) Weinfurt K. Differentiating between RCTs, PCTs, and quality improvement activities [Internet]. Bethesda: NIH Collaboratory; 2014 [cited 2021 Jul 29]. Available from: <https://rethinkingclinicaltrials.org/chapters/pragmatic-clinical-trial/what-is-a-pragmatic-clinical-trial-3/>.
- 18) Loudon K, Treweek S, Sullivan P, Donnan P, Thorpe KE, Zwarenstein M. The PRECIS-2 tool: designing trials that are fit for purpose.

- BMJ 2015;350:h2147.
- 19) Ahn J, Kim Y, Lee HY, Jang B, Jang EJ, Hyun MK. A summary of methods for comparative effectiveness research. Seoul: National Evidence-based Healthcare Collaborating Agency;2013. pp.7-12.
 - 20) Choudhry NK, Avorn J, Glynn RJ, Antman EM, Schneeweiss S, Toscano M, et al. Full coverage for preventive medications after myocardial infarction. N Engl J Med 2011;365:2088-2097.
 - 21) Hansson L, Hedner T, Dahlöf B. Prospective randomized open blinded end-point (PROBE) study. A novel design for intervention trials. Blood Press 1992;1:113-119.
 - 22) National Institute for Health and Care Excellence. Research recommendations process and methods guide. London: National Institute for Health and Care Excellence;2015.

Appendix 1

Characteristics of the study types for evidence-generation

Items of characteristics		Experimental studies (clinical trials)		Observational studies (cohorts)	
		RCTs	Non-RCTs	Prospective cohorts	Retrospective cohorts
Items for all study types	Primary outcomes are compared according to exposure to the intervention or comparator.	Y	Y	Y	Y
	Primary outcomes of different groups are compared.	Y	Y	Y	Y
	Investigators allocate participants (individuals/ groups) to the intervention or comparator group.	Y	Y	N	N
Items for experimental study types only	The groups are randomized.	Y	N	N/A	N/A
Item for observational study types only	The data for exposure to the intervention or comparator and for primary outcomes are collected concurrently.	N/A	N/A	N	N
	Each group is organized on the basis of exposure to the intervention or comparator.	N/A	N/A	Y	Y
	Times of exposure to intervention or comparator are the same.	N/A	N/A	Y	Y
	The data for exposure to the intervention or comparator are collected after the occurrence of primary outcomes, using already existing data.	N/A	N/A	N	Y

Y, yes; N, no; N/A, not applicable