

NECA 비플림위험 평가도구 매뉴얼

- AMSTAR2, ROBIS, RoB2.0, ROBINS-I -



NECA 비돌림위험 평가도구 매뉴얼

- AMSTAR 2, ROBIS, RoB 2, ROBINS-I -



Copyright C 한국보건의료연구원, 2021. All rights reserved. First edition.

본 책자는 공공 소유로서 비영리 목적으로의 인용, 배포 및 복제는

한국보건의료연구원의 별도의 승인 없이도 가능합니다.

그러나 특정 그림이나 표의 경우 원저작권자의 승인 없이는 사용할 수 없습니다.

NECA 비뿔림위험 평가도구 매뉴얼

- AMSTAR 2, ROBIS, RoB 2, ROBINS-I -

인쇄: 2021년 5월 31일 초판 1쇄 발행

발행인: 한광협

발행처: 한국보건의료연구원

서울특별시 중구 퇴계로 173 남산스퀘어빌딩 7층 Tel. 02-2174-2700

Homepage: www.neca.re.kr

ISBN 978-89-6834-725-2

Copyright © National Evidence-based Healthcare Collaborating Agency, 2021. First edition.

Published by NECA

May 2021

All rights reserved. Reproduction of this book by photocopying or electronic means for non-commercial purposes is permitted except those copyrighted materials noted for which further reproduction is prohibited without the specific permission of copyright holders. Otherwise, no part of this book may be reproduced, adapted, stored in a retrieval system or transmitted by any means, electronic, mechanical, photocopying, or otherwise without the prior written permission of NECA.

Comments and suggestions on this book can be made at neca@neca.re.kr



집필진

본 매뉴얼은 다음과 같이 인용될 수 있습니다.

김수영, 박동아, 서현주, 신승수, 이수정, 장보형, 차영주, 최미영, 박종연, 서유신.
NECA 비둘림위험 평가도구 매뉴얼. 한국보건의료연구원. 2021.

Suggested citation:

Kim SY, Park DA, Seo HJ, Shin SS, Lee SJ, Jang BH, Cha YJ, Choi MY, Park JY, Suh YS.
NECA's guidance for assessing tools of risk of bias, National Evidencebased Healthcare
Collaborating Agency, Seoul, 2021.

김수영 교수
한림대학교 의과대학 가정의학교실

박동아 연구위원
한국보건의료연구원

서현주 부교수
충남대학교 간호대학 간호학과

신승수 교수
아주대학교 의과대학 호흡기내과

이수정 연구교수
고려대학교 간호대학

장보형 부교수
경희대학교 한의과대학 예방의학교실

차영주 대리
코반스코리아 임상시험사업본부

최미영 부연구위원
한국보건의료연구원

박종연 선임연구위원
한국보건의료연구원

서유신 연구원
한국보건의료연구원



국민에게 양질의 의료를 제공하고 의료자원을 적절하게 활용하는 것은 대다수 국가들이 주요 보건의료정책 과제로 추구하고 있는 현안입니다. 특히, 최근엔 국제적으로 한정된 의료자원을 효과적으로 투입하여 국민 건강을 향상시키기 위한 전략으로서 과학적·합리적인 근거를 기반으로 한 정책의사결정에 대해 관심이 높아지고 있습니다. 국내에서도 이에 대한 필요성과 공감대가 높아지면서 근거기반 보건의료를 위한 체계적이고 일관된 방법의 수립 및 확산이 요구되고 있습니다.

한국보건의료연구원에서는 개원 이래 지금까지 보건의료분야의 현안들에 대한 의료기술평가연구들을 진행하여 연구 결과를 실제 보건의료 정책 의사결정에 활발하게 적용할 수 있도록 노력을 지속해 왔습니다. 또한 국내에서 우리 현실에 적합한 접근 방안을 고민할 필요가 있다고 판단하여 의료기술평가연구를 진행하는데 있어 여러 필수적인 연구방법론들의 표준화에 노력하는 한편, 활용을 촉진시키기 위한 기관 차원의 다양한 노력을 지속해 왔습니다.

그리고 그간 축적된 연구경험과 고민들을 함께 공유하고 논의하기 위한 노력의 일환으로 한국보건의료연구원에서는 체계적 문헌고찰 매뉴얼, 임상진료지침 개발 매뉴얼 및 NECA 연구방법론 시리즈로써 베이지안 메타분석법, 보건의료분야에서 비용 산출방법, 국내 보건의료 이차자료원 활용, 측정된 교란요인을 고려한 성과분석 방법, 한국적 상황을 고려한 비교효과연구 방법이 출간되어 근거기반 방법론의 정립 및 확산에 기여하였습니다. 최근에는 체계적 문헌고찰의 방법론의 최신 업데이트

이트 세부방법들에 대해 정리한 체계적 문헌고찰 방법론 책자를 발간하였고 올해, 체계적 문헌고찰의 가장 중요한 요소인 문헌의 질평가 영역에서의 새로 개발된 평가도구들에 대한 매뉴얼 책자를 발간하여 체계적 문헌고찰 연구방법의 질적 향상에 기여하고자 하였습니다.

적절한 체계적 문헌고찰을 바탕으로 한 의료기술평가보고서는 근거기반 보건의료의 시금석이라고 할 수 있습니다. 근거에 기반한 타당하고 신뢰할 수 있는 정보 제공은 의사 및 보건의료정책가 등 전문가들이 진료 또는 보건의료정책 수립 중 내리는 의학적 결정의 확신성을 높이고 궁극적으로 보건의료의 질을 향상시키는 결과를 가져올 수 있을 것이라 기대합니다.

한국보건의료연구원의 이번 결과물을 계기로 신뢰도 높은 연구 성과물 및 의료기술평가 보고물의 개발과 이에 대한 현장에서의 활용이 더욱 활성화되고 연구방법론의 지속적인 발전을 앞장서서 선도하기를 바랍니다.

2021년 3월

한국보건의료연구원장





적절한 비뚤림위험 평가 적용을 위해

체계적 문헌고찰 방법론은 근거중심의학의 발전과 함께 눈부신 발전을 거듭해왔으며, 체계적 문헌고찰 과정에서 주요 과정이라고 할 수 있는 비뚤림위험 평가의 방법론도 많은 발전이 있었습니다.

처음부터 비뚤림위험 평가는 체계적 문헌고찰 과정의 주요한 부분이었습니다. 하지만 초창기에 시행되던 비뚤림위험 평가는 변이가 너무 컸고, 평가결과가 체계적 문헌고찰 결과에 적절히 반영되지 않는 문제점이 있었습니다. 비뚤림위험 평가와 관련된 표준, 즉 관련용어, 평가방법, 평가도구, 평가결과에 대한 해석방법, 제시방법, 적용방법 등을 정리한 곳은 코크란 그룹이라고 할 수 있습니다. 2008년 코크란 연합은 코크란 Risk of Bias (RoB) 라는 도구를 발표했으며 이후 수많은 도구가 개발되었고, 이와 함께 비뚤림위험 평가 방법론도 많은 발전이 이루어졌습니다. 코크란 그룹은 정밀도, 외적타당도와 구분되는 비뚤림위험이라는 개념을 수립하였으며, 영역기반 비뚤림위험 평가, 투명성의 강조와 같은 비뚤림위험의 기본적인 원칙을 수립하였습니다.

일반적으로 비뚤림위험 평가는 연구설계에 따라서 이루어집니다. 체계적 문헌고찰, 무작위배정 비교임상시험(randomised controlled trial, RCT), 비무작위연구(non-randomised study, NRS), 진단연구, 경제성평가 연구 등이 주요 비뚤림위험 평가대상 연구설계라고 할 수 있습니다. 최근 들어 거의 모든 연구설계에 대한 비뚤림위험 평가도구가 업데이트되었습니다.

RCT 비뚤림위험 평가 도구 중 2008년 발표한 코크란 RoB 도구가 가장 체계화된

도구라고 할 수 있습니다. 이 도구는 코크란 리뷰 및 비 코크란 리뷰 모두에 널리 사용되었지만, 평가자간 신뢰도, 비뚤림위험 평가 결과의 고려와 같은 문제점이 지적되었고 이를 해결하기 위해 2019년 RoB 2라는 새로운 코크란 비뚤림위험 도구가 개발되었습니다.

중재의 유해성, 장기간 중재결과와 같이 RCT만으로 평가가 어려운 경우에 비무작위 연구에 대한 체계적 문헌고찰이 이루어지고 있습니다. 하지만 비무작위연구 체계적 문헌고찰 과정 중 비뚤림위험 평가에 많은 어려움이 있어 왔습니다. 과거 MINORS (Methodological Index for Non-Randomized Studies)나 국내에서 개발된 RoBANS (Risk of Bias for Nonrandomized Studies) 등이 사용되어 왔지만 여러 가지 문제로 제한적 사용만이 이루어졌습니다. 그러다가 코크란 연합에서 비무작위연구의 비뚤림위험 평가도구로 ACROBAT(2014)를 개발하고, 이것을 발전시켜 2016년 ROBINS-I (Risk Of Bias in Non-randomized Studies - of Interventions)를 발표하면서 많은 진보가 이루어졌습니다.

일차연구만큼은 아니지만 여러 가지 경우에 체계적 문헌고찰의 비뚤림위험 평가가 필요한 경우가 있습니다. 우선 임상진료지침의 개발이나 의료기술평가의 한 과정으로 기존 체계적 문헌고찰을 이용(Using SR)하는 경우에 해당 체계적 문헌고찰이 최소한의 질적 표준을 만족하는지 여부를 검토할 경우가 그러하며, 코크란 그룹에서 시행하고 있는 Overview of reviews의 과정에서 일차연구의 ROB평가와 유사한 내용으로 SR의 비뚤림위험 평가를 하는 경우에도 해당이 됩니다. 또한 체계적

문헌고찰 연구를 진행하거나 논문을 쓰는 경우 필요한 표준적 과정과 내용을 검토할 경우에도 이용이 될 수 있습니다.

2007년에 발표된 AMSTAR (A MeaSurement Tool to Assess systematic Reviews)가 과거 체계적 문헌고찰의 비뚤림위험 평가에 가장 많이 쓰인 도구였습니다. 하지만 AMSTAR는 보고의 내용과 다른 질적인 부분에는 다소 미흡하였고, 비무작위연구의 체계적 문헌고찰에는 이용될 수 없는 문제가 있었습니다. 최근에 변화된 비뚤림위험 평가 방법론과 체계적 문헌고찰 방법론을 효과적으로 반영할 수 없다는 지적이 있어서 AMSTAR를 개발했던 팀에서 AMSTAR를 AMSTAR 2로 업데이트하였으며, 2016년에는 영국 Bristol 대학 등의 연구진이 ROBIS (risk of bias in systematic reviews)라는 도구를 발표하였습니다.

우리나라에서 비뚤림위험 평가가 필요할 때 주로 사용되는 도구는 RCT에서 Cochrane RoB도구, NRS에서 ROBANS, 체계적 문헌고찰에서 AMSTAR이였습니다. 이들 도구들이 광범위하게 사용되면서 전체적인 체계적 문헌고찰이나 진료지침의 수준이 높아진 것도 사실이지만 최근 개발된 도구를 사용하지 않아 새롭게 적용되고 있는 표준에는 다소 미흡한 부분이 있었던 것도 사실이었습니다.

이에 NECA에서는 새롭게 업데이트되었거나 개발된 도구의 사용 용이성을 높이기 위해 해당 도구에 대한 매뉴얼을 만들기로 하였습니다.

이 매뉴얼에 소개할 비뚤림위험 평가 도구는 AMSTAR 2, ROBIS, Cochrane RoB 2, ROBINS-I의 네 가지입니다. 각각의 도구에 대해 적절한 번역과정을 거친 한글판도구와 해당 도구의 평가 서식지, 그리고 도구 매뉴얼, 원본 매뉴얼 번역본의 네 가지 버전을 제공합니다.

첫째, 한글판 도구는 해당 도구의 한글 번역본입니다. 번역은 2인 이상이 시행하였으며, 합의를 거쳐 최종본을 만들었습니다. AMSTAR 2, ROBIS의 두 도구에 대해서는 역번역을 통해 개념적 동등성을 보장하는 절차도 거쳤습니다.

둘째, 평가서식지는 해당 도구를 이용하여 비뚤림위험을 평가할 때 필요한 서식지입니다. 여러 가지 문헌에 대한 평가가 이루어지고 여러 사람이 참여하기 때문에

이를 종합할 수 있는 서식이 필요한 경우가 많이 있습니다.

셋째, 도구 매뉴얼은 각각 도구의 원본 매뉴얼을 바탕으로 국내 연구진들이 비플립위험 평가를 시행하는 도중에 참고할 수 있는 내용으로 만들었습니다. 더불어 실제 체계적 문헌고찰 과정에서 얻은 문헌을 대상으로 비플립위험을 시행하고 이에 대해 연구진이 합의한 평가 결과를 제시하고 평가이유까지 제시하여 이용자들이 쉽게 평가할 수 있도록 하였습니다.

넷째, 원본 매뉴얼은 부록의 형태로 제공되며 원래 도구에 있는 매뉴얼에 대한 한글판입니다.

애초에 한글판과 매뉴얼 번역만을 제공하려는 계획이 있었지만, 비플립위험 평가가 원활히 이루어지기 위해 예시를 제공하여 내용의 이해도와 가독성을 높이는 것이 필요하다는 판단에 따라 도구 매뉴얼과 원본 매뉴얼을 별도로 제시하였습니다.

향후 이 책의 내용은 비플립위험 평가의 경험치가 쌓이면서 정제되고 정리되어 더 나은 형태로 발전할 수 있을 것입니다. 그렇게 하도록 해당 내용을 꼼꼼히 살펴보고 적극적으로 활용하려는 노력이 필요할 것으로 생각합니다. 이 책을 만드는데 많은 수고를 하신 한국보건의로 연구원의 박동아 연구위원, 최미영 부연구위원, 이민 부연구위원, 아주대학교 신승수 교수, 경희대학교 장보형 교수, 충남대학교 서현주 교수, 고려대학교 이수정 연구교수, 코반스코리아 차영주님, 그리고 행정지원을 해주신 서유신 연구원께 감사드립니다, 아울러 이 책의 내용에 대해 꼼꼼히 자문해 주신 가톨릭의과대학 임현우 교수님, 순천향대학교 이유경 교수님, 서울대학교 한서경 교수님께도 감사의 마음을 전합니다. 끝으로 이 책 발행에 물심양면 지원을 아끼지 않으신 한광협 원장님께도 깊은 감사를 드립니다.

2021년 2월

김수영



목차

발간사	6
서문	8

PART 1. 체계적 문헌고찰 비뚤림위험 평가 도구

1장. AMSTAR 2

가. 개요	16
나. 국문 AMSTAR 2 평가서식지	18
다. 국문 AMSTAR 2	20
라. 국문 AMSTAR 2 : 사용자 매뉴얼	24

2장. ROBIS

가. 개요	40
나. 국문 ROBIS 평가서식지	40
다. 국문 ROBIS 평가도구, 사용자 매뉴얼	45

PART 2. RCT 비뚤림위험 평가 도구 : ROB 2

3장. Cochrane ROB 도구 2.0 (RoB 2)

가. 개요	76
나. 국문 ROB 2 평가 서식지	79
다. 국문 ROB 2	87
라. 국문 ROB 2 사용자 매뉴얼	104

PART 3. NRS 비뚤림위험 평가 도구 : ROBINS-I

제 4장. ROBINS-I 도구

가. 개요	134
나. 국문 ROBINS-I 평가서식지	135
다. 국문 ROBINS-I	142
라. 국문 ROBINS-I 사용자 매뉴얼	159

PART 1.

체계적 문헌고찰 비뿔리위험 평가 도구

1장. AMSTAR 2

가. 개요	16
나. 국문 AMSTAR 2 평가서식지	18
다. 국문 AMSTAR 2	20
라. 국문 AMSTAR 2 사용자 매뉴얼	24

2장. ROBIS 지침 문서

가. 개요	40
나. 국문 ROBIS 평가서식지	40
다. 국문 ROBIS 평가도구, 사용자 매뉴얼	45



1장. AMSTAR 2

가. 개요

의학 출판의 급격한 증가와 함께 의학적 혹은 정책적 의사결정에 체계적 문헌고찰의 중요성은 계속 커지고 있다. 하지만 모든 체계적 문헌고찰이 적절히 수행된 것은 아니기 때문에, 체계적 문헌고찰을 비평적으로 검토할 수 있게 해주는 도구를 통해 해당 체계적 문헌고찰의 신뢰성 정도를 평가하는 것이 필요할 수 있다.

Cochrane Collaboration Handbook은 체계적 문헌고찰을 작성 할 때 필요한 포괄적인 가이드를 제공하고 있지만 완료된 문헌고찰에 대한 비판적 평가에 관련된 내용을 다루고 있지는 않다.

2007년에 발표된 AMSTAR (A MeaSurement Tool to Assess systematic Reviews)는 체계적 문헌고찰의 가장 널리 사용되는 도구 중 하나이다. 이 AMSTAR는 의료 전문가, 정책 입안자들이 사용할 수 있는 실질적인 비판 평가 도구로서 역할을 해왔으며, 이 도구를 사용하면, 역학에 대한 지식이 없는 사람도 체계적 문헌고찰의 질을 신속하고 재현 가능하게 평가할 수 있다.

AMSTAR 2는 체계적 문헌고찰의 방법론적 질을 평가하기 위한 도구이며, AMSTAR (A MeaSurement Tool to Assess systematic Reviews)에 대한 갱신판이다.

AMSTAR 도구와 관련 영역은 아래와 같다.

영역	질문
1	'사전에' 체계적 문헌고찰의 계획이 수립되었는가?
2	문헌선택과 자료추출은 여러 명이 수행하였는가?
3	포괄적인 문헌검색을 하였는가?
4	포함기준에 출판상태(예: 회색문헌)가 사용되었는가?
5	포함 및 배제된 연구목록이 제시되었는가?
6	포함된 연구의 특성이 제시되었는가?
7	포함된 연구의 질이 평가되고 기술되었는가?
8	포함된 연구의 질은 결론을 도출하는데 적절히 사용되었는가?
9	개별연구의 결과를 결합한 방법이 적절하였는가?

10	출판비뚤림의 가능성을 평가하였는가?
11	이해상충이 기술되었는가?
판단	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오 <input type="checkbox"/> 대답할 수 없다 <input type="checkbox"/> 적용할 수 없다

AMSTAR 2는 AMASTAR에 비해 다음 사항이 변화하였다.

- ① AMASTAR의 10개 영역은 일부 용어만 바뀌고 그대로 존속하도록 하였다.
- ② 두 영역은 세분화된 평가를 위해 두 문항으로 분리하였다(연구 선택과 자료 추출의 중복성을 두 문항으로, 자금원의 영향을 개발 연구 수준과 전체 리뷰 수준으로 구분).
- ③ 비뚤림위험 평가 여부를 무작위배정 비교 임상시험(RCT)와 비무작위연구(NRS)를 나누어서 평가하도록 하였다.
- ④ 회색문헌 관련 문항을 제거하고 문헌검색에 포함시켰다.
- ⑤ 네 영역이 추가됨 : 두 영역은 ROBINS-I에서 왔으며(PICO 적용 여부 평가, 근거합성 과정에서 비뚤림위험의 적용), 두 영역은 기존 AMSTAR 문항이 변형되었다(이질성의 원인과 의미, 그리고 연구설계 선택의 정당화).
- ⑥ 응답에 대한 판단기준이 바뀌었다 ('대답할 수 없다'와 '적용할 수 없다'가 없어짐, 정보가 없다면 '아니오'로 평가해야 한다, '일부 예' 응답을 추가함)

나. 국문 AMSTAR 2 평가서식지

번호	질문	판단
1	체계적 문헌고찰의 연구질문과 포함기준에는 PICO의 구성요소가 포함되었는가?	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
2	체계적 문헌고찰 방법론이 실제 문헌고찰을 시행하기 전에 확립되었으며, 보고서에는 프로토콜로부터 중대한 이탈이 있는 경우 이에 대한 정당화(합당한 이유)가 제시되었나?’	<input type="checkbox"/> 예 <input type="checkbox"/> 일부 예 <input type="checkbox"/> 아니오
3	문헌고찰 저자는 문헌고찰에 포함될 연구설계 선택에 대해 설명하였나?	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
4	문헌고찰 저자는 포괄적인 문헌 검색 전략을 사용하였는가?	<input type="checkbox"/> 예 <input type="checkbox"/> 일부 예 <input type="checkbox"/> 아니오
5	문헌고찰 저자는 연구 선택을 중복으로 수행하였는가?	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
6	문헌고찰 저자는 자료추출을 중복으로 수행하였는가?	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
7	문헌고찰 저자는 배제 연구에 대한 목록과 합당한 배제사유를 제공하였는가?	<input type="checkbox"/> 예 <input type="checkbox"/> 일부 예 <input type="checkbox"/> 아니오
8	문헌고찰저자는 포함된 연구들의 세부사항을 적절히 기술하였는가?	<input type="checkbox"/> 예 <input type="checkbox"/> 일부 예 <input type="checkbox"/> 아니오
9	문헌고찰저자는 문헌고찰에 포함된 개별 연구의 비뚤림위험(ROB)을 평가하기 위해 만족스러운 도구를 사용하였는가?	<input type="checkbox"/> 예 <input type="checkbox"/> 일부 예 <input type="checkbox"/> 아니오
10	문헌고찰 저자는 문헌고찰에 포함된 연구들의 자금 출처에 대해 보고하였는가?	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
11	메타분석을 수행하였다면, 문헌고찰 저자는 이에 대한 합당한 이유를 제시하였고, 연구결과의 통계학적 결함을 위해 적절한 방법을 사용하였는가?	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오 <input type="checkbox"/> 메타분석 없음

번호	질문	판단
12	메타분석을 수행하였다면, 문헌고찰 저자는 개별 연구의 비독립위험이 메타분석 연구결과나 다른 근거 합성에 미칠 잠재적 영향을 평가하였는가?	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오 <input type="checkbox"/> 메타분석 없음
13	문헌고찰저자가 문헌고찰 결과를 해석/논의할 때 개별 연구의 비독립 위험을 고려하였는가?	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
14	문헌고찰저자는 문헌고찰 연구결과에서 발견된 이질성에 대해 만족스러운 설명을 하였는가?	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
15	양적 합성을 하였다면, 문헌고찰저자는 출판비독립 (소규모연구 비독립)에 대한 적절한 조사를 수행하고, 문헌고찰 결과에 미칠 수 있는 영향에 대해 고찰하였는가?	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오 <input type="checkbox"/> 메타분석 없음
16	문헌고찰 저자는 문헌고찰 수행을 위한 자금지원을 포함하여 잠재적 이해상충에 대해 보고하였는가?	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오

다. 국문 AMSTAR 2

1. 체계적 문헌고찰의 연구질문과 포함기준에는 PICO의 구성요소가 포함되었는가?		
<p>예 :</p> <ul style="list-style-type: none"> <input type="checkbox"/> 인구집단(P) <input type="checkbox"/> 중재(I) <input type="checkbox"/> 비교군(C) <input type="checkbox"/> 중재결과(O) 	<p>선택(권고)사항</p> <ul style="list-style-type: none"> <input type="checkbox"/> 추적관찰 시점 	<p><input type="checkbox"/> 예</p> <p><input type="checkbox"/> 아니오</p>
2. 체계적 문헌고찰 방법론이 실제 문헌고찰을 시행하기 전에 확립되었으며 보고서에는 프로토콜로부터 중대한 이탈이 있는 경우 이에 대한 정당화(합당한 이유)가 제시되었나?		
<p>일부 예 :</p> <p>저자는 아래 모두를 포함하는 서면 프로토콜 또는 가이드를 가지고 있다고 진술하고 있다.</p> <ul style="list-style-type: none"> <input type="checkbox"/> 문헌고찰 질문 <input type="checkbox"/> 검색전략 <input type="checkbox"/> 포함기준 <input type="checkbox"/> 비뮴림위험 평가 	<p>일부 예에 더하여:</p> <p>추가로 프로토콜이 등록되어 있고, 아래 사항이 명시되어 있다.</p> <ul style="list-style-type: none"> <input type="checkbox"/> 해당되는 경우 메타분석/합성 계획, 그리고 <input type="checkbox"/> 이질성 원인에 대한 조사 계획 <input type="checkbox"/> 프로토콜 이탈에 대한 정당화(합당한 이유) 	<p><input type="checkbox"/> 예</p> <p><input type="checkbox"/> 일부 예</p> <p><input type="checkbox"/> 아니오</p>
3. 문헌고찰 저자는 문헌고찰에 포함될 연구설계 선택에 대해 설명하였나?		
<p>예: 아래 중 <u>하나</u> 충족:</p> <ul style="list-style-type: none"> <input type="checkbox"/> 무작위 배정 비교임상시험(RCT)만 포함하는 것에 대해 설명 <input type="checkbox"/> 혹은, 중재 비무작위연구만(NRSI)만 포함하는 것에 대해 설명 <input type="checkbox"/> 혹은, 무작위 배정 비교임상시험과 중재 비무작위 연구 모두를 포함하는 것에 대해 설명 		<p><input type="checkbox"/> 예</p> <p><input type="checkbox"/> 아니오</p>
4. 문헌고찰 저자는 포괄적인 문헌 검색 전략을 사용하였는가?		
<p>일부 예 (모두 포함해야 함):</p> <ul style="list-style-type: none"> <input type="checkbox"/> 최소한 2개 이상의 (연구질문에 적합한) 데이터베이스를 검색하였다. <input type="checkbox"/> 키워드와 검색전략을 제시하였다. <input type="checkbox"/> 출판물 제한(예: 언어)에 대해 정당한 이유를 제시하였다. 	<p>예 : 모두 포함해야 함</p> <ul style="list-style-type: none"> <input type="checkbox"/> 포함된 연구의 참고문헌 목록 검색 <input type="checkbox"/> 임상시험/연구 레지스트리 검색 <input type="checkbox"/> 해당 분야 내용 전문가를 포함하거나 자문을 받음 <input type="checkbox"/> 해당되는 경우, 회색 문헌 검색 <input type="checkbox"/> 문헌고찰 완료 24개월 이내에 검색 수행 	<p><input type="checkbox"/> 예</p> <p><input type="checkbox"/> 일부 예</p> <p><input type="checkbox"/> 아니오</p>

5. 문헌고찰 저자는 연구 선택을 중복으로 수행하였는가?		
예: 아래 중 <u>하나</u> 충족: <input type="checkbox"/> 최소한 두 명의 평가자가 독립적으로 적격 연구를 선택하고, 최종 포함여부를 합의하였다. <input type="checkbox"/> 또는 두 명의 평가자가 일부 표본에 대해 적격연구를 선정하였고, 좋은 일치도를 보여(예 : 80% 이상), 나머지는 한 평가자가 문헌 선정을 하였다.	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오	
6. 문헌고찰 저자는 자료추출을 중복으로 수행하였는가?		
예: 아래 중 <u>하나</u> 충족: <input type="checkbox"/> 최소한 두 명의 평가자가 포함된 문헌의 어떤 자료를 추출할지에 대해 합의에 도달하였다. <input type="checkbox"/> 또는 두 명의 평가자가 일부 적격 연구 표본에서 자료를 추출하였고, 좋은 일치도를 보여(예 : 80% 이상), 나머지는 한 평가자가 자료 추출을 하였다.	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오	
7. 문헌고찰 저자는 배제 연구에 대한 목록과 합당한 배제사유를 제공하였는가?		
일부 예 : <input type="checkbox"/> 전문까지 검토하였지만, 문헌고찰에서 배제된 연구에 대한 목록을 제시	예: 아래 사항도 충족 <input type="checkbox"/> 잠재적으로 관련성이 있는 연구가 문헌 고찰에서 배제된 합당한 이유 제시	<input type="checkbox"/> 예 <input type="checkbox"/> 일부 예 <input type="checkbox"/> 아니오
8. 문헌고찰저자는 포함된 연구들의 세부사항을 적절히 기술하였는가?		
일부 예 (아래 모두) <input type="checkbox"/> 인구집단에 대한 기술 <input type="checkbox"/> 중재에 대한 기술 <input type="checkbox"/> 비교군에 대한 기술 <input type="checkbox"/> 중재결과에 대한 기술 <input type="checkbox"/> 연구설계에 대한 기술	예 : 아래 사항도 모두 충족 <input type="checkbox"/> 인구집단 세부사항에 대한 기술 <input type="checkbox"/> 중재 세부사항에 대한 기술(적절한 경우, 용량 포함) <input type="checkbox"/> 비교군 세부사항에 대한 기술(적절한 경우, 용량 포함) <input type="checkbox"/> 연구 세팅에 대한 기술 <input type="checkbox"/> 추적관찰의 시점	<input type="checkbox"/> 예 <input type="checkbox"/> 일부 예 <input type="checkbox"/> 아니오
9. 문헌고찰저자는 문헌고찰에 포함된 개별 연구의 비돌림위험(ROB)을 평가하기 위해 만족스러운 도구를 사용하였는가?		
RCTs 일부 예: 아래 모두에 대한 ROB 평가 <input type="checkbox"/> 배정은폐, 그리고 <input type="checkbox"/> 중재결과 평가시 환자 및 결과 평가자의 눈가림(모든 원인 사망률과 같은 객관적인 중재결과에는 불필요)	예: 아래에 대한 ROB 평가 <input type="checkbox"/> 진정한 무작위 배정 순서, <u>그리고</u> <input type="checkbox"/> 다양한 측정치나 분석 중에서 보고할 연구결과 선택 여부	<input type="checkbox"/> 예 <input type="checkbox"/> 일부 예 <input type="checkbox"/> 아니오 <input type="checkbox"/> NRSI만 포함
NRSI 일부 예: 아래에 대한 ROB 평가 <input type="checkbox"/> 교란 <input type="checkbox"/> 선택 비돌림	예: 아래에 대한 ROB 평가 <input type="checkbox"/> 노출과 중재결과 확인위해 사용한 방법, <u>그리고</u> <input type="checkbox"/> 다양한 측정치나 분석 중에서 보고할 연구결과 선택	<input type="checkbox"/> 예 <input type="checkbox"/> 일부 예 <input type="checkbox"/> 아니오 <input type="checkbox"/> RCT만 포함

10. 문헌고찰 저자는 고찰에 포함된 연구들의 자금 출처에 대해 보고하였는가?	
예 : <input type="checkbox"/> 문헌고찰에 포함된 연구의 자금출처에 대해 보고해야 한다. 주 : 평가자가 살펴보았지만 일차연구 저자가 보고하지 않은 경우도 해당됨.	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
11. 메타분석을 수행하였다면, 문헌고찰 저자는 이에 대한 합당한 이유를 제시하였고, 연구결과의 통계학적 결합을 위해 적절한 방법을 사용하였는가?	
RCTs 예 : <input type="checkbox"/> 메타분석을 통해 자료를 결합한 합당한 이유를 제시하였다. <input type="checkbox"/> 그리고 적절한 가중치법을 이용하여 연구결과를 결합했으며 이질성이 있다면 이를 보정하였다. <input type="checkbox"/> 그리고 이질성의 원인에 대한 조사를 하였다.	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오 <input type="checkbox"/> 메타분석 없음
NRSI 예 : <input type="checkbox"/> 저자는 메타분석을 통해 자료를 결합한 합당한 이유를 제시하였다. <input type="checkbox"/> 그리고 적절한 가중치법을 이용하여 연구결과를 결합했으며 이질성이 있다면 이를 보정하였다. <input type="checkbox"/> 그리고 NRSI의 교란변수를 보정한 효과추정치를 결합하였거나, 보정 효과 추정치를 결합할 수 없을 때만 원 자료를 결합하였다. <input type="checkbox"/> 그리고 RCT와 NRSI 모두가 고찰에 포함되었을 때, 무작위 배정 비교임상시험과 중재 비무작위연구의 요약 추정치를 별도로 보고하였다.	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오 <input type="checkbox"/> 메타분석 없음
12. 메타분석을 수행하였다면, 문헌고찰 저자는 개별 연구의 비뚤림위험이 메타분석 연구결과나 다른 근거 합성에 미칠 잠재적 영향을 평가하였는가?	
예 : <input type="checkbox"/> 비뚤림위험이 낮은 RCT만을 포함하였다. <input type="checkbox"/> 또는 다양한 비뚤림위험을 가진 RCT나 NRSI를 포함한 경우 효과의 요약 추정치에 비뚤림위험이 어떤 영향을 미치는지 조사하였다.	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오 <input type="checkbox"/> 메타분석 없음
13. 문헌고찰저자는 고찰 결과를 해석/논의할 때 개별 연구의 비뚤림위험을 고려하였는가?	
예 : <input type="checkbox"/> 비뚤림위험이 낮은 RCT만을 포함하였다. <input type="checkbox"/> 또는 RCT에 중등도 이상의 비뚤림위험이 있거나 문헌고찰에 NRSI를 포함한 경우에 비뚤림위험이 연구결과에 어떤 영향을 미쳤는지에 대해 고찰하였다	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오
14. 문헌고찰저자는 연구결과에서 발견된 이질성에 대해 만족스러운 설명과 고찰을 하였는가?	
예 : <input type="checkbox"/> 연구결과에서 유의미한 이질성이 없었다. <input type="checkbox"/> 또는 이질성이 있는 경우, 이질성의 원인에 대해 조사하고, 이질성이 문헌고찰 연구결과에 미치는 영향에 대해 논의하였다.	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오

<p>15. 양적 합성을 하였다면, 문헌고찰저자는 출판비독립(소규모연구 비독립)에 대한 적절한 조사를 수행하고, 문헌고찰 결과에 미칠 수 있는 영향에 대해 고찰하였는가?</p>	
<p>예 :</p> <p><input type="checkbox"/> 출판비독립에 대한 그래프 또는 통계적 검증을 수행하고 출판비독립의 유무와 영향 정도에 대해 고찰하였다.</p>	<p><input type="checkbox"/> 예</p> <p><input type="checkbox"/> 아니오</p> <p><input type="checkbox"/> 메타분석 없음</p>
<p>16. 문헌고찰 저자는 문헌고찰 수행을 위한 자금지원을 포함하여 잠재적 이해상충에 대해 보고하였는가?</p>	
<p>예 :</p> <p><input type="checkbox"/> 저자는 이해상충이 없다고 보고하였다. 또는</p> <p><input type="checkbox"/> 저자가 자금을 기술했으며, 잠재적 이해상충을 관리한 방법을 보고하였다.</p>	<p><input type="checkbox"/> 예</p> <p><input type="checkbox"/> 아니오</p>

라. 국문 AMSTAR 2 사용자 매뉴얼

아래 내용은 AMSTAR 2의 사용자 매뉴얼 (<https://amstar.ca/docs/AMSTAR%202-Guidance-document.pdf>)에 기초하여 개별 문항 평가의 내용과 방법을 설명한 것이다.

평가의 실제에 대한 사례를 위해 두 편의 체계적 문헌고찰에 대해 연구진이 AMSTAR 2에 대한 평가를 하였고, 그 결과를 제시한다.

1) 평가에 사용한 문헌

1	Lee HK, Hwang IH, Kim SY, Pyo SY. The effect of exercise on prevention of the common cold: a meta-analysis of randomized controlled trial studies. Korean Journal of Family Medicine 2014; 35(3): 119-126. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4040429/
2	Howes S, Hartmann-Boyce J, Livingstone-Banks J, Hong B, Lindson N. Antidepressants for smoking cessation. Cochrane Database of Systematic Reviews 2020, Issue 4. Art. No.: CD000031. DOI: 10.1002/14651858.CD000031.pub5. https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD000031.pub5/full

2) 개별 문항 해석

1. 체계적 문헌고찰의 연구질문과 포함기준에는 PICO의 구성요소가 포함되었는가?		
예 : <input type="checkbox"/> 인구집단(P) <input type="checkbox"/> 중재(I) <input type="checkbox"/> 비교군(C) <input type="checkbox"/> 중재결과(O)	선택(권고)사항 <input type="checkbox"/> 추적관찰 시점	<input type="checkbox"/> 예 <input type="checkbox"/> 아니오

설명	- PICO(인구, 중재, 대조군 및 중재결과)는 연구질문 체계화를 위한 틀로 흔히 사용된다. - 만일 추적관찰 시점이 적격연구 선택에 중요한 경우(예: 중재의 영향은 몇 년이 지난 후에 나타날 수 있다), 추적관찰 시점도 제시해야 한다. - 이러한 PICO 틀을 통해 문헌고찰 사용자가 연구결과의 적용 가능성을 평가할 수 있다.
평가 가이드	- 초록, 서론, 방법 파트를 주의 깊게 읽어야 함. '예' : PICO의 4가지 요소가 보고서의 어딘가에 나누어져 기술되어 있는 경우

평가의 예

문헌번호	평가결과	평가 이유
1	예	Method에 "A meta-analysis of randomized controlled trials was performed to investigate the effects of regular, moderate-intensity aerobic exercise on the prevention of the common cold in the general population."로 제시함.
2	예	Methods :Criteria for considering studies for this review Types of participants(We included..)/Types of interventions(pharmacotherapies with antidepressant properties for smoking cessation) /Comparators(placebo, no pharmacotherapy, alternative therapeutic control)/Types of outcome measures (Efficacy, measured as smoking cessation).

2. 체계적 문헌고찰 방법론이 실제 문헌고찰을 시행하기 전에 확립되었으며 보고서에는 프로토콜로부터 중대한 이탈이 있는 경우 이에 대한 정당화(합당한 이유)가 제시되었나?

<p>일부 예 :</p> <p>저자는 아래 <u>모두</u>를 포함하는 서면 프로토콜 또는 가이드를 가지고 있다고 진술하고 있다.</p> <ul style="list-style-type: none"> <input type="checkbox"/> 문헌고찰 질문 <input type="checkbox"/> 검색전략 <input type="checkbox"/> 포함기준 <input type="checkbox"/> 비뚤림위험 평가 	<p>일부 예에 더하여:</p> <p>추가로 프로토콜이 등록되어 있고, 아래 사항이 명시되어있다.</p> <ul style="list-style-type: none"> <input type="checkbox"/> 해당되는 경우 메타분석/합성 계획, 그리고 <input type="checkbox"/> 이질성 원인에 대한 조사 계획 <input type="checkbox"/> 프로토콜 이탈에 대한 정당화(합당한 이유) 	<ul style="list-style-type: none"> <input type="checkbox"/> 예 <input type="checkbox"/> 일부 예 <input type="checkbox"/> 아니오
--	--	---

설명	<ul style="list-style-type: none"> - 체계적 문헌고찰은 관찰 연구의 한 형태이기 때문이며, 문헌고찰이 시작되기 전에 문헌고찰 방법을 합의해야 한다. - 잘 개발된 프로토콜을 준수하면 문헌고찰 시 비뚤림위험이 감소한다. - 문헌고찰저자들은 서면프로토콜을 통해 작업했다는 사실이 객관적으로 입증되도록 하여야 한다. - 프로토콜은 등록처 (예: SPRISPO https://www.crd.york.ac.uk/PROSPERO/), 공개 간행물 저널(예: BMJ Open) 또는 기관 사무처 또는 연구 윤리 위원회 등에서 확인할 수 있다.
평가 가이드	<ul style="list-style-type: none"> - '일부 예' : 최소한 연구질문과 문헌고찰 연구 방법이 미리 계획되었음이 보고서에 명시된 경우 - '예' : 문헌고찰이 진행되기 전에(등록처 또는 다른 독립 기관(예: 연구 윤리 위원회 또는 기관 사무처)에 서면 프로토콜이 있었음이 입증된 경우 - 프로토콜에서 이탈이 있는 경우 문헌고찰 저자는 이를 보고하여야 하며, 이것이 정당한지 평가해야 한다. 설명되지 않은 명백한 불일치가 있는 경우 등급을 하향해야 한다.

평가의 예

문헌번호	평가결과	평가이유
1	아니오	서면프로토콜에 대한 언급 없음.
2	예	<ol style="list-style-type: none"> 1) 프로토콜 있음 : Protocol first published: Issue 3, 1997. 2) 프로코틀과 실제 SR의 차이에 대한 기술 있음 : https://www.cochranlibrary.com/cdsr/doi/10.1002/14651858.CD000031.pub5/information. 3) 상기에 대한 정보 모두 있음.

3. 문헌고찰 저자는 문헌고찰에 포함될 연구설계 선택에 대해 설명하였나?

예: 아래 중 하나 충족:

무작위 배정 비교임상시험(RCT)만 포함하는 것에 대해 설명

혹은, 중재 비무작위연구만(NRS)만 포함하는 것에 대해 설명

혹은, 무작위 배정 비교임상시험과 중재 비무작위 연구 모두를 포함하는 것에 대해 설명

예

아니오

설명	<ul style="list-style-type: none"> - 체계적인 문헌고찰에 포함되는 연구 유형이 자의적이면 안 된다. - 일반적인 규칙은 RCT로 제한하였을 때 치료 효과에 대한 평가가 불완전 할 것인지 자문해 보는 것이다(저자들이 기술한 내용으로 미루어 추측해야 하는 경우도 있음). - 이 영역은 다소 소홀히 다루어진 것도 사실이며, 평가 기준을 충족하는지에 대한 판단이 어려울 수 있다. - 연구설계 선택에 대한 정당성은 전체 연구 보고서를 주의 깊게 읽은 후에 하는 것이 좋다.
평가 가이드	<ul style="list-style-type: none"> - 우선 RCT로 제한하였을 때 치료의 효과에 대한 평가가 불완전 할 것인지 자문해 본다. 이에 대한 판단은 관련 RCT가 없는지, 사용 가능한 RCT에 중재결과가 없는지[보통 위해] 여부, 통계적 검정력이 부족한지 여부, 인구집단이 제한적이거나 중재나 대조군의 대표성이 부족한지 등을 검토할 수 있다. - 만일 이에 대한 대답이 '예'인 경우 중재 비무작위 연구를 포함하는 것이 정당화된다. - 반대로, RCT만으로 문헌고찰을 시행하는 것을 정당화하기 위해서는, RCT만으로 관심 효과에 대한 완전한 자료를 제공할 수 있다고 주장해야 한다. - RCT가 필요한 중재결과 자료를 온전히 제공할 수 없는 경우 또는 RCT 문헌고찰이 완료되고 나서, NRS 문헌고찰을 통해 보완할 경우, NRS로만 문헌고찰을 제한하는 것이 정당화된다. - RCT와 NRS를 모두 포함하는 전략은 이를 통해 중재와 관련된 효과와 위해를 완전히 파악할 수 있다면 정당화될 수 있다. 만일 그렇다면 이 두 가지 유형의 연구를 독립적으로 평가하고 결합할 것을 권고한다(메타 분석이 적절한 경우).

평가의 예

문헌번호	평가결과	평가 이유
1	아니오	RCT만 포함한 이유에 대한 특별한 언급이 없음.
2	아니오	RCT만 포함한 이유에 대한 특별한 언급이 없음.

4. 문헌고찰 저자는 포괄적인 문헌 검색 전략을 사용하였는가?

<p>일부 예 (모두 포함해야 함):</p> <ul style="list-style-type: none"> <input type="checkbox"/> 최소한 2개 이상의 (연구질문에 적합한) 데이터베이스를 검색하였다. <input type="checkbox"/> 키워드와 검색전략을 제시하였다. <input type="checkbox"/> 출판물 제한(예: 언어)에 대해 정당한 이유를 제시하였다. 	<p>예 : 모두 포함해야 함</p> <ul style="list-style-type: none"> <input type="checkbox"/> 포함된 연구의 참고문헌 목록 검색 <input type="checkbox"/> 임상시험/연구 레지스트리 검색 <input type="checkbox"/> 해당 분야 내용 전문가를 포함하거나 자문을 받음 <input type="checkbox"/> 해당되는 경우, 회색 문헌 검색 <input type="checkbox"/> 문헌고찰 완료 24개월 이내에 검색 수행 	<ul style="list-style-type: none"> <input type="checkbox"/> 예 <input type="checkbox"/> 일부 예 <input type="checkbox"/> 아니오
--	---	---

설명	<ul style="list-style-type: none"> - 체계적 문헌고찰에 있어 검색의 중요성은 잘 알려져 있기 때문에 기존 도구와 거의 유사하다. - 적어도 두 개의 서지 데이터베이스를 검색해야 하고, 검색 시점, 데이터베이스 이름(예: Central, EMBASE 및 MEDLINE)이 제시되어야 한다. 검색어에 대한 전체 검색 전략을 제시하고, 기존 문헌고찰, 연구 레지스트리를 검토하여 검색을 보완해야 한다. - 검색의 보완을 위해 특정 연구 분야의 전문가를 포함하거나 자문을 받고, 포함 연구의 참고문헌 목록을 검토한다. 때로 연구결과를 확인 하거나 업데이트나 수정분을 얻기 위해 원 연구 저자들에게 접근해야 할 필요가 있다. - 모든 언어의 출판물을 찾아야 하며, 언어 제한이 있으면 정당화해야 한다. - 경우에 따라 회색 문헌 검색의 필요성이 있는 경우가 있다. 회색 문헌 중 일부는 웹 사이트(예: 정부, 비정부기구, 보건의료 기술 기관)에서만 이용할 수 있는 정책보고서, 프로그램 평가 보고 서와 같은 것들도 있다. 회색문헌은 동료평가가 되지 않은 문제가 있기 때문에 유사한 평가라도 있는지 확인해 보아야 한다. 회색 문헌이 중요하게 여겨지는 경우, 저자들은 적절한 자료를 검색해 보아야 하는데 임상시험 등록처, 학술대회 초록집, 학위 논문, 개별 웹사이트(예: 대학, 리서치게이트)가 그런 곳이다. - 의료 중재에 대한 임상시험 중 상당수는 동료심사 학술지에 출판되지 않아서 회사 스폰서나 연구자로부터 직접 입수해야 할 수도 있다.
평가 가이드	<ul style="list-style-type: none"> - '예'에 체크하려면 문헌고찰을 위해 시행한 검색의 모든 측면이 만족스러워야 한다.

평가의 예

문헌번호	평가결과	평가 이유
1	일부 예	"2개이상 DB, 키워드와 검색전략, 언어제한 없음" 등을 만족하지만, "해당분야 전문가 자문, 임상시험 검색" 등은 시행하지 않음.
2	예	참고문헌 검색, 연구 레지스트리, 회색문헌, 24개월 만족함. RCT만 포함한 이유에 대한 특별한 언급이 없음, 내용전문가 포함함.

5. 문헌고찰 저자는 연구 선택을 중복으로 수행하였는가?

예: 아래 중 하나 충족:

최소한 두 명의 평가자가 독립적으로 적격 연구를 선택하고, 최종 포함여부를 합의하였다.

혹은 두 명의 평가자가 일부 표본에 대해 적격연구를 선정하였고, 좋은 일치도를 보여 (예 : 80% 이상), 나머지는 한 평가자가 문헌 선정을 하였다.

예

아니오

설명	<ul style="list-style-type: none"> - 두 명의 문헌고찰 저자가 체계적인 문헌고찰에 포함시킬 연구의 적격성을 연구의 특성(제목, 초록, 전문으로 부터)을 통해 실행하였는지를 확인해야 한다. - 두 명의 평가자가 일부 표본에 대해 시행한 적격연구에 대한 평가에서 좋은 일치도를 보여(예 : 80%이상), 나머지는 한 평가자가 문헌 선정 하는 경우도 선택가능 한 응답옵션이다.
평가 가이드	- “예” : 최소한 두 명의 평가자가 독립적으로 연구 선정을 하거나, 일부 표본에 대해 선정에 대한 두 명의 일치도를 구하고 좋은 일치도를 보여 나머지를 한 평가자가 선정을 한 경우

평가의 예

문헌번호	평가결과	평가 이유
1	예	연구선택을 모두 두 사람이 수행함.
2	예	연구선택을 모두 두 사람이 수행함.

6. 문헌고찰 저자는 자료추출을 중복으로 수행하였는가?

예: 아래 중 하나 충족:

최소한 두 명의 평가자가 포함된 문헌의 어떤 자료를 추출할지에 대해 합의에 도달하였다.

또는 두 명의 평가자가 일부 적격 연구 표본에서 자료를 추출하였고, 좋은 일치도를 보여(예 : 80% 이상), 나머지는 한 평가자가 자료 추출을 하였다.

예

아니오

설명	<ul style="list-style-type: none"> - 자료 추출을 수행하는 독립적인 평가자가 적어도 두 명 이상 있어야 한다. - 의견 불일치가 생겼을 때는 합의 과정을 이용해야 한다. 한 명이 자료 추출을 수행하고자 하는 경우에는, 두 번째 평가자와 적격 연구 표본에 대한 일치도를 확인하고 카파 점수 0.80 이상을 달성해야 한다.
평가 가이드	- “예” : 최소한 두 명의 평가자가 독립적으로 자료 추출을 하거나 일부 표본에 대해 자료 추출에 대한 일치도가 좋아 나머지 자료 추출을 한 명이 수행한 경우

평가의 예

문헌번호	평가결과	평가 이유
1	예	자료 추출을 모두 두 사람이 수행함.
2	예	자료 추출을 모두 두 사람이 수행함.

7. 문헌고찰 저자는 배제 연구에 대한 목록과 합당한 배제사유를 제공하였는가?

일부 예 : <input type="checkbox"/> 전문까지 검토하였지만, 문헌고찰에서 배제된 연구에 대한 목록을 제시	예: 아래 사항도 충족 <input type="checkbox"/> 잠재적으로 관련성이 있는 연구가 문헌고찰에서 배제된 합당한 이유 제시	<input type="checkbox"/> 예 <input type="checkbox"/> 일부 예 <input type="checkbox"/> 아니오
--	---	---

설명	<ul style="list-style-type: none"> - 문헌고찰 저자가 각 연구의 배제에 대한 정당성을 가지는지, 그리고 잠재적으로 관련이 있는 연구의 전체 목록을 제공하는지를 평가한다. - 부적절한/부적합한 인구집단, 중재 및 비교중재 등의 다양한 이유로 연구가 배제될 수 있다. - 비뚤림위험에 근거하여 배제여부를 결정해서는 안 되며, 비뚤림위험은 별도로, 문헌고찰 과정에서 다루어져야 한다. - 합당하지 않은 배제는 문헌고찰 결과를 편향시킬 수 있으므로, 문헌고찰 초기 단계에서는 포용적인(inclusive) 접근법을 권장한다.
평가 가이드	<ul style="list-style-type: none"> - 문헌고찰 저자가 잠재적으로 관련이 있는 연구의 전체 목록을 제공하고 각 연구의 배제에 대한 합당한 사유를 제공하는지를 평가한다.

평가의 예

문헌번호	평가결과	평가 이유
1	아니오	배제된 연구에 대한 목록이 없음.
2	예	배제 연구 목록과 배제 사유가 제시됨.

8. 문헌고찰저자는 포함된 연구들의 세부사항을 적절히 기술하였는가?

일부 예 (아래 모두) <input type="checkbox"/> 인구집단에 대한 기술 <input type="checkbox"/> 중재에 대한 기술 <input type="checkbox"/> 비교군에 대한 기술 <input type="checkbox"/> 중재결과에 대한 기술 <input type="checkbox"/> 연구설계에 대한 기술	예 : 아래 사항도 모두 충족 <input type="checkbox"/> 인구집단 세부사항에 대한 기술 <input type="checkbox"/> 중재 세부사항에 대한 기술(적절한 경우, 용량 포함) <input type="checkbox"/> 비교군 세부사항에 대한 기술(적절한 경우, 용량 포함) <input type="checkbox"/> 연구 세팅에 대한 기술 <input type="checkbox"/> 추적관찰의 시점	<input type="checkbox"/> 예 <input type="checkbox"/> 일부 예 <input type="checkbox"/> 아니오
---	--	---

설명	- 연구대상자, 중재, 비교중재, 중재결과, 연구설계, 분석 및 세팅에 대한 설명이 제공되어야 한다. - 이러한 설명은 평가자 또는 사용자가 연구가 적절하게 선택되었는지(PICO 구조와 관련하여) 그리고 연구의 인구집단과 중재가 자신의 임상진료와 정책과 관련이 있는지에 대해 판단하기에 충분해야 한다. - 또한 중재효과의 이질성을 연구하기 위한 틀도 제공하여야 한다 (예: 용량, 연령 범위, 임상 세팅 등).
-----------	--

평가의 예

문헌번호	평가결과	평가 이유
1	일부 예	인구집단, 중재 세부사항, 세팅 등에 대한 세부사항이 다소 미흡.
2	예	PICO에 대해 자세히 기록됨.

9. 문헌고찰저자는 문헌고찰에 포함된 개별 연구의 비뚤림위험(ROB)을 평가하기 위해 만족스러운 도구를 사용하였는가?		
<p>RCTs</p> <p>일부 예 :아래 모두에 대한 ROB 평가</p> <p><input type="checkbox"/> 배정은폐, 그리고</p> <p><input type="checkbox"/> 중재결과 평가시 환자 및 결과 평가자의 눈가림(모든 원인 사망률과 같은 객관적인 중재결과에는 불필요)</p>	<p>예 :아래에 대한 ROB 평가</p> <p><input type="checkbox"/> 진정한 무작위 배정 순서, 그리고</p> <p><input type="checkbox"/> 다양한 측정치나 분석 중에서 보고할 연구결과 선택 여부</p>	<p><input type="checkbox"/> 예</p> <p><input type="checkbox"/> 일부 예</p> <p><input type="checkbox"/> 아니오</p> <p><input type="checkbox"/> NRSI만 포함</p>
<p>NRSI</p> <p>일부 예 :아래에 대한 ROB 평가</p> <p><input type="checkbox"/> 교란</p> <p><input type="checkbox"/> 선택 비뚤림</p>	<p>예 :아래에 대한 ROB 평가</p> <p><input type="checkbox"/> 노출과 중재결과 확인위해 사용한 방법, 그리고</p> <p><input type="checkbox"/> 다양한 측정치나 분석 중에서 보고할 연구결과 선택</p>	<p><input type="checkbox"/> 예</p> <p><input type="checkbox"/> 일부 예</p> <p><input type="checkbox"/> 아니오</p> <p><input type="checkbox"/> RCT만 포함</p>

설명	<ul style="list-style-type: none"> - 비뚤림위험 평가는 NRSI(중재 비무작위 연구)를 포함하는 체계적 문헌고찰 평가에 있어 중요한 부분이다. 핵심적인 내용은 문헌고찰 저자들이 연구결과를 요약하고 해석할 때 비뚤림위험을 고려했는지 여부이다. - RCT에 대한 체계적 문헌고찰에서 개별 RCT에 대한 ROB에 대해 적절한 평가를 수행했는지 여부를 코크란 핸드북에 따라 판단해야 한다. - 문헌고찰 저자는 ROB 평가에 체계적인 접근법을 사용해야 하며, 가급적 적절히 개발된 평가 도구를 사용해야 한다. - 비무작위연구의 평가에 문헌고찰 저자가 RoB를 어떻게 평가하였는지 평가할 때, 만일 현재 팀에서 그런 조언을 얻기 힘들면 자문해줄 방법론 전문가와 내용 전문가를 찾아봐야 한다. 그들에게 어떤 교란 요인이 중요한지, 고려 중인 문헌고찰과 관련이 있는 선택 비뚤림과 측정 비뚤림을 확인하는 방법에 대한 구체적인 조언을 들어야 한다. 또한 교란 요인에 대한 보정 방법에 대해 조언을 구해야 한다.
평가 가이드	<ul style="list-style-type: none"> - 비뚤림위험을 평가했는지 살펴보고, 어떤 도구를 사용하였는지 확인한다. - 해당 도구의 비뚤림위험 영역에서 해당 사항이 있는지 확인한다.

평가의 예

문헌번호	평가결과	평가 이유
1	예	<ul style="list-style-type: none"> - RCT만 포함함. - Cochrane ROB tool을 사용하여 모두 포함되어 있음.
2	예	<ul style="list-style-type: none"> - RCT만 포함함. - Cochrane ROB tool을 사용하여 모두 포함되어 있음.

10. 문헌고찰 저자는 고찰에 포함된 연구들의 자금 출처에 대해 보고하였는가?

예 :
 문헌고찰에 포함된 연구의 자금출처에 대해 보고해야 한다.
 주 : 평가자가 살펴보았지만 일차연구 저자가 보고하지 않은 경우도 해당됨.

예
 아니오

설명	- 상업적으로 후원되는 연구가 독립적으로 자금을 지원받은 연구보다 후원자의 제품에 유리한 연구결과가 나올 가능성이 더 높다. - 문헌고찰 저자는 문헌고찰에 포함된 각 연구에 대한 자금 출처를 상세히 기술하거나 연구 보고서에 정보가 제공되지 않았다는 것을 기술하는 것이 중요하다. 이 정보에 따라 상업적으로 자금을 조달한 연구결과와 독립적으로 자금을 조달한 연구결과를 별도로 분석할 수도 있다.
-----------	--

평가의 예

문헌번호	평가결과	평가 이유
1	아니오	포함된 연구들의 자금출처에 대한 언급이 없음.
2	예	연구들의 자금 출처와 이해상충에 대해 자료를 추출하였다.

11. 메타분석을 수행하였다면, 문헌고찰 저자는 이에 대한 합당한 이유를 제시하였고, 연구결과의 통계학적 결합을 위해 적절한 방법을 사용하였는가?

RCTs
 예 :
 메타분석을 통해 자료를 결합한 합당한 이유를 제시하였다.
 그리고 적절한 가중치법을 이용하여 연구결과를 결합했으며 이질성이 있다면 이를 보정하였다.
 그리고 이질성의 원인에 대한 조사를 하였다.

예
 아니오
 메타분석 없음

NRSI
 예 :
 저자는 메타분석을 통해 자료를 결합한 합당한 이유를 제시하였다.
 그리고 적절한 가중치법을 이용하여 연구결과를 결합했으며 이질성이 있다면 이를 보정하였다.
 그리고 NRSI의 교란변수를 보정한 효과추정치를 결합하였거나, 보정 효과 추정치를 결합할 수 없을 때만 원 자료를 결합하였다.
 그리고 RCT와 NRSI 모두가 고찰에 포함되었을 때, 무작위 배정 비교임상시험과 중재 비무작위연구의 요약 추정치를 별도로 보고하였다.

예
 아니오
 메타분석 없음

설명	<ul style="list-style-type: none"> - 문헌고찰 저자는 어떤 원칙에 따라 포함된 연구에 대한 메타 분석을 수행하기로 결정하였는지를 문헌고찰 프로토콜에 명확히 제시해야한다. - 이러한 원칙에는 단일 통합 효과(예 : 다수의 호환 가능한, 검정력이 낮은 연구에서 얻은 것, 연구의 비교성(연구집단, 중재, 비교중재 측면에서)을 평가하여 결합가능한지에 대한 평가 등의 내용이 될 수 있다. - RCT의 경우 고정 또는 변량 효과 모형을 사용하기로 결정한 이유를 설명해야 하고 이질성 조사 방법을 설명해야한다. - NRSI 연구의 연구집단은 수십 또는 수백 명의 소규모 코호트에서 수십만 명 규모, 그리고 사건의 수도 몇 건에서 수천 건에 이르기 까지 규모가 매우 다양하다. 이러한 연구결과를 규모가 작은 RCT의 연구결과와 결합할 경우, 비무작위 연구들의 자료가 통합 효과추정치를 결정할 가능성이 크다 - 또한, NRSI의 결과는 다양한 비뚤림에 의해 영향을 받을 수 있으며, 전체 통합 추정치는 정밀하지만 비뚤림이 있을 수 있다. 문헌고찰 저자는 연구설계가 다르면 통합된 추정치를 별도로 보고해야 한다. - NRSI가 고품질 RCT 정도로 비뚤림위험이 낮은 경우는 매우 드물기 때문에, NRSI를 기반으로 한 통합 추정치와 신뢰구간을 주의 깊게 검토해야 한다. 이 문제는 NRSI의 다양한 비뚤림위험과 비뚤림위험에 대한 불확실성을 고려할 때 중요하다. - 이질성은 모든 메타분석에서 중요한 문제이다. 비무작위 연구의 경우 서로 다른 연구에서 더 다양한 방법이 사용될 가능성이 크기 때문에, 이질성 문제는 NRSI에 대한 문헌고찰에서 특히 중요하다. 일반적인 이질성의 원인[다른 비교군, 중재결과의 기저 위험 변이 또는 연구 집단의 다른 특성, 다른 중재(예 : 용량 효과, 맥락/ 세팅, 실무자 경험) 및 다양한 중재결과 정의] 이외에도, 연구 자료의 출처, 자료의 완전성, 자료 관리 및 분석 방법에서의 이질성을 고려하는 것이 중요하다. - 교란요인의 중재 효과에 대해 통계적 보정을 하면 보정되지 않은 추정치와는 상당히 다른 추정치가 도출되기도 한다. 일반적으로, NRSI 고찰 결과를 결합할 때는 원자료가 아니라 보정된 효과추정치를 대상으로 해야 한다. 만약 저자들이 그렇게 하지 않았다면 이에 대한 합당한 설명이 있어야 한다. - 그러나 연구에 따라서는 공변량 세트(또는 상이한 방식으로 측정되거나 나름대로 적용된 공변량)를 달리해서 보정된 치료효과를 보고할 가능성이 매우 높다. 이러한 다양성은 잠재적 이질성의 또 다른 원인이 되기도 한다.
평가 가이드	<ul style="list-style-type: none"> - 메타 분석을 보고 한 경우에만 평가하고 메타분석이 없으면 '메타분석 없음'에 체크 한다 - 예 : 메타분석을 시행한 이유에 대한 언급이 있으며, 적절한 가중치법과 이질성 조사 방법을 이용하여 메타분석을 수행하였다.

평가의 예

문헌번호	평가결과	평가 이유
1	예	<ul style="list-style-type: none"> - 메타분석을 통해 자료를 결합한 합당한 이유를 제시하지 않았지만 내용으로 미루어 단일 통합효과, 비교성에 대한 고려를 하고 있음. - 이질성에 따라 fixed, random 효과모형을 사용하겠다고 언급하고 있음. - 이질성은 I2를 사용하여 추정하겠다고 밝히고 있음.
2	예	<ul style="list-style-type: none"> - 메타분석을 통해 자료를 결합한 합당한 이유를 제시하지 않지는 않지만 내용으로 미루어 단일 통합효과, 비교성에 대한 고려를 하고 있음. - 이질성에 따라 fixed, random 효과모형을 사용하겠다고 언급하고 있음. - 이질성은 I2를 사용하여 추정하겠다고 밝히고 있음.

12. 메타분석을 수행하였다면, 문헌고찰 저자는 개별 연구의 비뚫림위험이 메타분석 연구결과나 다른 근거 합성에 미칠 잠재적 영향을 평가하였는가?

예 :

비뚫림위험이 낮은 RCT만을 포함하였다.

또는 다양한 비뚫림위험을 가진 RCT나 NRSI를 포함한 경우 효과의 요약 추정치에 비뚫림위험이 어떤 영향을 미치는지 조사하였다.

예

아니오

메타분석 없음

설명	<ul style="list-style-type: none"> - 문헌고찰 저자가 높은 질의 RCT만 선택한 경우, 연구결과에 미치는 비뚫림의 잠재적 영향에 대한 논의가 거의 필요 없을 수 있다. 그러나 다양한 질을 가진 RCT를 포함시킨 경우, 비뚫림위험의 영향을 회귀분석으로 평가하거나, 비뚫림위험이 낮은 연구만으로 통합효과크기를 추정해야 한다. - NRSI의 경우, 문헌고찰 저자들은 비뚫림위험이 낮거나 증등도인 연구, 또는 ROB가 낮은 연구만으로 통합효과크기를 추정해야 한다. 메타 분석 (또는 회귀분석과 같은 다른 자료합성 기술)이 수행되지 않은 경우에도, 저자는 개별 연구결과에 대한 ROB의 영향에 대한 의견을 제시해야 한다.
-----------	--

평가의 예

문헌번호	평가결과	평가 이유
1	아니오	- 비뚫림위험이 메타분석 결과나 근거합성에 미칠 영향에 대한 언급이 없음.
2	예	- 비뚫림위험에 따른 효과 추정치에 대해 고찰하고 있다. (예 : “Removing the three studies deemed to be at an overall high risk of bias did not change the interpretation of the pooled effect estimate.”)

13. 문헌고찰저자는 고찰 결과를 해석/논의할 때 개별 연구의 비뚫림위험을 고려하였는가?

예 :

비뚫림위험이 낮은 RCT만을 포함하였다.

또는 RCT에 증등도 이상의 비뚫림위험이 있거나 문헌고찰에 NRSI를 포함한 경우에도 비뚫림위험이 연구결과에 어떤 영향을 미쳤는지에 대해 고찰하였다

예

아니오

설명	<ul style="list-style-type: none"> - 문헌고찰 저자는 문헌고찰 연구결과에 대한 해석에 있어 ROB의 영향에 대해 논의해야 한다. 이점은 RCT와 NRSI를 모두 포함시킨 경우 더 중요할 수 있다. 비뚫림위험은 효과 추정치, 개별 연구결과들의 차이에 영향을 미칠 수 있다
-----------	---

설명	- 임상 진료나 정책에 영향을 줄 수 있는 권고안을 제시 할 경우, 저자는 ROB를 명시적으로 고려해야한다.
평가 가이드	- 예 : 비뚤림위험이 낮은 RCT만을 포함하였거나, 그렇지 않은 경우 비뚤림위험이 연구결과에 어떤 영향을 미쳤는지 고찰한 경우

평가의 예

문헌번호	평가결과	평가 이유
1	아니오	-비뚤림위험이 고찰 결과 해석에 미칠 영향에 대한 언급이 없음.
2	예	- 근거수준평가에 비뚤림위험을 고려하였음.

14. 문헌고찰저자는 연구결과에서 발견된 이질성에 대해 만족스러운 설명과 고찰을 하였는가?

예 :

- 연구결과에서 유의미한 이질성이 없었다.
- 또는 이질성이 있는 경우, 이질성의 원인에 대해 조사하고, 이질성이 문헌고찰 연구결과에 미치는 영향에 대해 논의하였다.

- 예
- 아니오

설명	- 문헌고찰 저자가 이질성의 주요 원인인 PICO 요소와 연구설계에 따른 비뚤림위험 영역에 대해 조사하였는지 중요하다. - 문헌고찰 저자는 이러한 가능성을 탐색하고 연구결과 결론 및 권고안에 대한 이질성의 영향에 대해 논의해야한다.
평가 가이드	- 예 :이질성이 있는지 여부를 조사하였고, 이질성이 있다면 PICO 요소와 방법론적인 고려사항 등 이질성의 원인을 조사하였고, 이질성이 연구결과에 어떠한 영향을 미쳤는지 조사하였다.

평가의 예

문헌번호	평가결과	평가 이유
1	아니오	- 이질성이 있어서 하위집단 분석을 하였지만 만족스러운 설명이나 고찰을 하지는 못함.
2	예	- I2로 이질성 평가를 하였고 이질성에 따라 연구결과에 미치는 영향을 고찰하였다.

15. 양적 합성을 하였다면, 문헌고찰저자는 출판비돌림 (소규모연구 비돌림)에 대한 적절한 조사를 수행하고, 문헌고찰 결과에 미칠 수 있는 영향에 대해 고찰하였는가?

예 : 출판비돌림에 대한 그래프 또는 통계적 검증을 수행하고 출판비돌림의 유무와 영향 정도에 대해 고찰하였다.

예
 아니오
 메타분석 없음

설명	<ul style="list-style-type: none"> - 출판비돌림은 갈때기 그림 혹은 통계적 검증을 통해 평가하며 그 결과가 유의하면 출판 비돌림이 있다는 의미이다. 하지만 검정이 민감하지 않기 때문에 검정결과가 유의하지 않다고, 출판 비돌림이 없음을 보장하지는 않는다. - 출판 비돌림의 중요성은 맥락과 세팅에 달려 있는데 예를 들어 산업체 후원 연구(예 : 약물, 의기기 등)는 독립적인 연구에 비해 비돌림의 영향을 받을 가능성이 더 크다. - 중요한 것은 문헌고찰 저자가 (필요와 세팅에 따라) 1) 심층적이고 집중적인 문헌 검색을 통해 출판 비돌림을 조사하기 위해 최선을 다했는지, 2) 연구결과에 대한 해석 및 고찰에 출판 비돌림의 영향을 고려하였는지, 3) 현재의 연구결과를 무효화하는데 필요한 누락된 '무효(null)' 연구의 수를 결정하기 위해 민감도 분석을 수행하였는지 여부이다.
평가 가이드	<ul style="list-style-type: none"> - 예 : 출판 비돌림에 대한 검정(그래프 혹은 통계적 검정)을 수행하여 출판비돌림이 있는지 그리고 있다면 그 영향이 어떠한지 고찰하였다.

평가의 예

문헌번호	평가결과	평가 이유
1	아니오	- 선택문헌 10편 미만으로 출판비돌림에 대한 검정을 수행하지 않았음. 출판비돌림의 유무 및 영향에 대해 고찰하지 않았음.
2	예	- 출판비돌림에 대한 검정을 수행하였고 그 영향을 고찰하였음.

16. 문헌고찰 저자는 문헌고찰 수행을 위한 자금지원을 포함하여 잠재적 이해상충에 대해 보고하였는가?

예 : 저자는 이해상충이 없다고 보고하였다. 또는 저자가 자금원을 기술하였으며, 잠재적 이해상충을 관리한 방법을 보고하였다.

예
 아니오

설명	<ul style="list-style-type: none"> - 이해상충이 있는 연구는 그렇지 않은 연구보다 증재에 유리한 연구결과를 도출할 수 있으며, 이는 체계적 문헌고찰도 마찬가지이다. 따라서 체계적 문헌고찰 저자도 직접적인 자금 출처를 보고해야한다.
-----------	---

	<ul style="list-style-type: none"> - 학술지에 있는 일반적 이해상충 정보 이외에 필요하면 다른 유대관계를 보고해야 한다. 문헌고찰은 자금 지원으로부터는 자유로울 수 있지만, 문헌고찰 저자가 체계적 문헌고찰에 포함된 제품을 제조하는 회사와 관련이 있을 수 있기 때문이다. - 전문가적인 이해상충은 강력하지만, 거의 보고되지 않아 식별하기가 어렵다. 연구자들이 연구 분야에 오랫동안 있어서 생기는 신념이 있을 수 있고, 그것으로 인한 이해상충이 생길 수 있다. 이런 이해상충은 평가하기 어려울 수 있지만, 문헌고찰 저자가 문헌고찰 중인 분야에서 다양한 논문을 출판하였으며 그들의 연구가 해당 체계적 문헌고찰에 포함되어 있다는 사실에서 추론될 수 있다. - 이해상충 때문에 다른 비뚤림 영역의 결함이 생길 수 있고 만일 그렇다면 이 문항은 항상 개별적으로 평가해야한다.
평가 가이드	- 예 : 이해상충이 없다고 보고하였거나 재정적 이해상충이 있는 경우 이를 어떻게 관리하였는지 보고한 경우

평가의 예

문헌번호	평가결과	평가 이유
1	예	이해상충이 없다고 보고하였다.
2	예	이해상충이 없다고 보고하였다.(https://www.cochranelibrary.com/cdsr/doi/10.1002/146518w58.CD000031.pub5/information).

3) 체계적 문헌고찰 연구결과의 전반적인 신뢰도 평가 (SR 질 평가)

가) 핵심적 영역의 확인

- 현재로서 다음의 7개의 영역이 핵심적 영역이라고 할 수 있다.

- ① 문항 2 : 문헌고찰 수행 전에 프로토콜 등록
- ② 문항 4 : 문헌검색의 적절성
- ③ 문항 7 : 배제 문헌의 정당화
- ④ 문항 9 : 포함 문헌에 대한 비뚤림위험 평가
- ⑤ 문항 11 : 메타분석 방법의 적절성
- ⑥ 문항 13 : 문헌고찰의 연구결과 해석에 비뚤림위험의 고려
- ⑦ 문항 15 : 출판비뚤림의 존재와 영향 평가

- AMSTAR 2에서 비뚤림위험이 핵심적 영역 중 2개를 차지하였는데 NRS를 포함하는 연구에 있어서는 비뚤림위험이 매우 중요하기 때문임. 만일 질이 높은 RCT만을 포함하는 경우에는 이것이 덜 중요할 수 있다

- 임상시험협력그룹의 결과물을 기반으로 체계적 문헌고찰을 수행하는 경우와 같이 이미 알려진 문헌에 대한 합성을 할 때는, 검색(문항 4), 배제문헌의 정당화(문항 7), 출판비돌림 영역(문항 15)은 핵심적 영역으로 고려되지 않을 수 있다.
- 메타분석을 수행하지 않은 경우 메타분석의 적절성 부분(문항 11)은 적용할 수 없는데 그런 경우에는 문헌고찰 저자가 연구결과를 요약할 때 비돌림위험의 영향을 잘 고려하였는지를 평가해야 한다.
- 위의 7개의 핵심적 영역은 개발자의 제안이며 문헌고찰 평가자가 다른 중요한 문항을 추가하거나 대체할 수 있다. 예를 들어, 치료의 부작용에 관한 체계적 문헌고찰에서 비무작위 연구를 포함하지 않았을 경우 문항 3 (포함된 연구설계 선택에 대한 설명), 연구간 치료효과 크기의 상당한 이질성을 설명하지 못한 경우 문항 14 (결과의 이질성에 대한 설명 및 고찰)이 핵심적 영역에서의 결함으로 평가될 수 있다.

나) SR 질 평가

• 높음

- 약점이 없거나 1개 정도의 핵심적이지 않은 약점만 있음 : SR은 해당 질문에 대해 정확하고 포괄적으로 연구결과를 요약함.

• 중등도

- 둘 이상의 비핵심적 약점 : 체계적 문헌고찰에 두 개 이상의 약점이 있지만 핵심적이지는 않음. 합성결과가 비교적 정확하다고 할 수 있음 (비핵심적인 약점이 상당히 많으면 낮음으로 갈 수도 있음)

• 낮음

- 비핵심적 약점에 관계없이 하나의 핵심적 영역의 약점 : 핵심적인 약점이 있으며 합성결과가 포괄적이거나 정확하지 않을 수 있음.

• 매우 낮음

- 비핵심적 약점에 관계없이 핵심적 영역에서 둘 이상의 약점 : 합성결과가 정확하고 포괄적이라고 하기 어려움.

평가의 예

문헌번호	평가결과	평가 이유
1	매우 낮음	핵심영역 4개(2, 7, 13, 15)에서 약점이 있어 "매우 낮음".
2	높음	1개(3)의 비핵심적인 영역에서만 약점이 있으므로 "높음".

2장. ROBIS

가. 개요

ROBIS (Tool to assess risk of bias in systematic reviews) 는 체계적 문헌고찰(이하 문헌고찰)의 비뚤림위험을 평가하고자 개발된 도구이다. ROBIS는 증재, 진단, 예후 및 병인과 관련된 질문을 다루는 의료 환경 내의 문헌고찰을 위해 사용되도록 설계되었다. 이 도구는 3단계로 이루어져 있다. (1) 관심 질문과의 관련성 평가(해당하는 경우), (2) 문헌고찰 과정에서의 우려사항 확인, (3) 문헌고찰에서 비뚤림위험 판단. ROBIS는 문헌고찰의 비뚤림위험과, (해당하는 경우) 연구질문과 검토하는 체계적 문헌고찰의 관련성을 모두 평가한다. 구체적으로 1) 문헌고찰의 방법이 요약 추정과 체계적 문헌고찰 결론의 비뚤림위험을 최소화한 정도, 그리고 2) 문헌고찰이 다루는 연구질문이 체계적 문헌고찰의 overview 작성자나 진료지침 개발자와 같은 사용자가 다루는 연구질문과 일치하는 정도를 다룬다.

ROBIS는 세 부류의 사용자 그룹을 염두에 두고 개발되었다.

- 체계적 문헌고찰의 overview 저자
- 진료지침 개발자
- 문헌고찰이 완료되면 문헌고찰에서 비뚤림위험을 평가하거나 프로토콜 단계에서 문헌고찰 방법을 계획할 때 비뚤림위험을 최소화하고자 하는 검토자

체계적 문헌고찰에서 비뚤림위험을 평가하고자 하는 사람에게도 도움이 될 수 있지만, 다른 목표 청중을 위해 특별히 설계된 것은 아니다. 다른 ROBIS의 잠재적인 사용자에는 의사 결정을 지원하는 조직(예: NICE, IQWiG), EBM에 관심이 있는 임상, 저널 편집자 및 원고 검토자가 있다.

나. 국문 ROBIS 평가서식지

1) 1단계: 관련성 평가 (해당하는 경우)

ROBIS는 증재, 원인, 진단, 예후 질문에 대한 체계적 문헌고찰의 비뚤림위험을 평가하도록 설계되었다. 질문 종류에 따라 목표질문(즉, overview 또는 진료지침 상의 질문)과 평가 중인 문헌고찰에서 다루는 질문을 기술하시오.

중재 문헌고찰

범주	목표질문 (예: overview 또는 진료지침)	평가 대상 문헌고찰
환자/인구집단		
중재		
비교중재		
중재결과		

원인 문헌고찰

범주	목표질문 (예: overview 또는 진료지침)	평가 대상 문헌고찰
환자/인구집단		
노출과 비교노출		
결과(outcome)		

진단 정확도 문헌고찰

범주	목표질문 (예: overview 또는 진료지침)	평가 대상 문헌고찰
환자군		
중재 검사		
참고 표준 검사		
대상 질환		

예후 문헌고찰

범주	목표질문 (예: overview 또는 진료지침)	평가 대상 문헌고찰
환자군		
예상되는 중재결과		
모델의 사용 목적		
의도한 시점		

문헌고찰의 질문이 목표질문에 부합하는가?	예/아니오/불확실
------------------------	-----------

2) 2단계: 문헌고찰 과정의 우려사항 파악

영역 1: 연구 적격 기준

연구 적격 기준 및 이에 대한 제한, 그리고 연구 목적과 적격 기준이 사전에 설정되었다는 근거가 있는지 여부를 기술하시오:

1.1 문헌고찰이 사전에 정의된 연구 목적과 적격 기준을 준수하였는가?	Y/PY/PN/N/NI*
1.2 적격 기준이 문헌고찰의 질문에 부합하는가?	Y/PY/PN/N/NI
1.3 적격 기준이 모호하지 않은가?	Y/PY/PN/N/NI
1.4 연구 특성에 따른 적격 기준의 제한이 적절한가? (출판 시기, 표본 크기, 연구의 질, 측정된 중재결과 등)	Y/PY/PN/N/NI
1.5 정보의 출처에 따른 적격 기준의 제한이 적절한가? (출판형식(학술지논문, 보고서 등)이나 출판 상태, 언어, 자료이용 가능성 등)	Y/PY/PN/N/NI
연구 적격 기준의 세부사항에 대한 우려	낮음/높음/불확실

우려의 근거:

* Y=YES, PY=PROBABLY YES, PN=PROBABLY NO, N=NO, NI=NO INFORMATION

영역 2: 연구의 검색 및 선택

연구의 검색 및 선택 방법을 기술하시오 (예: 참여한 검토자 수):

2.1 출판/미출판 보고서에 대한 적절한 범위의 데이터베이스/전자자료원이 검색에 포함되었는가?	Y/PY/PN/N/NI*
2.2 관련 보고서를 찾기 위해 데이터베이스 검색에 추가적인 방법이 사용되었는가?	Y/PY/PN/N/NI
2.3 검색 전략의 용어와 구조가 가능한 한 많은 적격 연구를 검색할 가능성을 높였는가?	Y/PY/PN/N/NI
2.4 날짜, 출판 형식 또는 언어에 따른 제한이 적절하였는가?	Y/PY/PN/N/NI
2.5 연구 선택의 오류를 최소화하기 위한 노력이 이루어졌는가?	Y/PY/PN/N/NI
연구의 검색 그리고/또는 선택 방법에 대한 우려	낮음/높음/불확실
우려의 근거:	

* Y=YES, PY=PROBABLY YES, PN=PROBABLY NO, N=NO, NI=NO INFORMATION

영역 3: 자료 수집 및 연구 평가

자료 수집 방법, 연구에서 추출하거나 다른 방법을 통해 수집한 데이터, 비돌림위험 평가 방법 (예: 관련된 검토자 수) 및 비돌림위험 평가에 사용된 도구를 기술하십시오:

3.1 자료 수집 오류를 최소화하기 위해 노력하였는가?	Y/PY/PN/N/NI*
3.2 연구결과를 해석할 수 있기에 충분한 연구 특성 정보가 문헌고찰 저자와 독자에게 제공되었는가?	Y/PY/PN/N/NI
3.3 합성에 사용되도록 타당한 모든 연구결과가 수집되었는가?	Y/PY/PN/N/NI
3.4 비돌림위험(또는 방법론적 질)이 적절한 기준을 사용하여 공식적으로 평가되었는가?	Y/PY/PN/N/NI
3.5 비돌림위험 평가의 오류를 최소화하기위한 노력을 하였는가?	Y/PY/PN/N/NI
자료 수집과 연구 평가 방법 대한 우려	낮음/높음/불확실

우려의 근거:

* Y=YES, PY=PROBABLY YES, PN=PROBABLY NO, N=NO, NI=NO INFORMATION

영역 4: 합성 및 연구결과

합성 방법을 기술하십시오:

4.1 포함되어야 하는 모든 연구가 자료합성에 포함되었는가?	Y/PY/PN/N/NI*
4.2 사전에 정의된 모든 분석을 보고하였거나 이탈(departure) 대해 설명하였는가?	Y/PY/PN/N/NI
4.3 포함된 연구들의 연구질문, 연구설계, 연구결과의 특성과 유사성에 근거해보았을 때 합성이 적절하였는가?	Y/PY/PN/N/NI
4.4 연구간 변이(이질성)가 작거나 합성 중 다루어졌는가?	Y/PY/PN/N/NI
4.5 연구결과가 견고하였는가? 예를 들어 깔대기 그림(funnel plot)이나 민감도 분석으로 연구결과의 견고성을 보여주었는가?	Y/PY/PN/N/NI
4.6 일차 연구에서 비돌림위험이 미미했거나, 비돌림위험이 자료 합성 과정에서 고려되었는가??	Y/PY/PN/N/NI
합성 및 연구결과에 대한 우려	낮음/높음/불확실

우려의 근거:

* Y=YES, PY=PROBABLY YES, PN=PROBABLY NO, N=NO, NI=NO INFORMATION

3) 3단계: 비뚤림위험의 판단

2단계 평가에서 파악된 우려의 요약

영역	우려	우려의 근거
1. 연구 적격 기준의 세부사항에 대한 우려		
2. 연구의 검색 그리고/또는 선택 방법에 대한 우려		
3. 자료수집과 연구평가 방법에 대한 우려		
4. 합성 및 연구결과에 대한 우려		

문헌고찰의 비뚤림위험

결론이 근거에 의해 뒷받침되는지를 기술하십시오:

A. 연구결과를 해석할 때 영역 1~4에서 확인한 모든 우려 사항을 다루었는가?	Y/PY/PN/N/NI
B. 문헌고찰의 연구질문에 대해 포함된 연구의 관련성이 적절히 고려되었는가?	Y/PY/PN/N/NI
C. 문헌고찰 저자는 통계적인 유의성을 보이는 결과만을 강조하는 것을 피하였는가?	Y/PY/PN/N/NI
문헌고찰의 비뚤림위험	낮음/높음/불확실

비뚤림위험의 근거:

다. 국문 ROBIS 평가도구, 사용자 매뉴얼

1) 1단계: 관련성 평가 (해당하는 경우)

- 목표질문(target question; overview 또는 진료지침 상의 질문으로 평가자가 답변하려고 하는 질문)과 평가 중인 문헌고찰의 질문이 관련이 있는지 평가한다. 평가할 문헌고찰의 PICO 또는 이에 준하는 항목을 작성한 후 목표질문과 문헌고찰 질문이 일치하는지를 확인한다.
- 하나 이상의 범주가 일치하지 않을 경우 '아니오'로, 부분 일치하는 경우 '부분적'으로 평가한다. 문헌고찰이 단독적으로 평가되고 있고 목표질문이 없다면 이 단계는 생략할 수 있다.

문헌고찰 종류별 목표질문과 PICO 형식의 예시

종류	PICO 형식	예시
중재	환자/인구집단: 중재: 비교중재: 중재결과:	만성 C형 간염에 감염된 성인 pegylated 인터페론을 이용한 트리플 항바이러스 요법 이중 항바이러스 요법 지속적인 바이러스 반응 (sustained virologic response)
원인	환자/인구집단: 노출과 비교노출: 중재결과:	성인 체질량 지수 (BMI) 대장암
진단	환자군: 중재 검사: 참고 표준 검사: 대상 질환:	직장암 의심 증상을 가진 성인 내시경 초음파 외과적 조직학 검사(surgical histology) 조기 직장암 (T0)
예후	환자군: 예상되는 중재결과: 모델의 사용 목적: 의도한 시점:	태아 성장 제한이 있거나 없는 임신부로서, 막(membrane)의 조기파열 증거가 없고, 선천성 또는 구조적 이상의 증거가 없음. 유해한 임신 결과 (저체중 또는 과체중, 신생아 사망, 주산기 사망률) 예측 임신 후기 (37주 이상)

예시

진단 문헌고찰에서:

범주	목표질문	평가 대상 문헌고찰
환자:	직장암 의심 증상을 가진 성인	불확실
중재 검사:	내시경 초음파	내시경 초음파
참고 표준 검사:	외과적 조직학	수술
대상 질환:	조기 직장암 (T0)	조기 직장암 (T0)

타당성 평가		이유
문헌고찰의 질문이 목표질문에 부합하는가? (예: overview 혹은 진료지침에서)	불확실 (unclear)	이 질문은 중재 검사, 참고 표준 검사 및 대상 질환은 일치하지만 인구집단이 문헌고찰에서 정의되지 않았기 때문에 문헌고찰에서 다른 질문이 목표질문과 일치하는지 여부는 불확실하다.

2) 2단계: 문헌고찰 과정의 우려사항 파악

- 2단계는 비뮴림이 문헌고찰의 결과에 영향을 끼칠 수 있는 영역을 찾아내는 것을 목표로 하며, 연구 적격 기준, 연구 검색 및 선택, 자료 수집 및 연구 평가, 합성 및 결론 등 4개 영역의 평가로 이루어져 있다.
- 각 영역은 i) 판단을 뒷받침하는데 사용되는 정보, ii) 신호 질문(signalling question), iii) 비뮴림위험 우려의 판단 등 세 부분으로 구성된다. 각 영역들은 연결해서 고려되어야 하며 각각 따로 평가해서는 안 된다. 예를 들어 영역 2(연구의 검색 및 선택)를 평가할 때, 평가자가 영역 1에 명시된 연구 문제와 연결해서 검색 전략의 용어와 구조를 고려해야 한다.
- 신호 질문은 ‘그렇다(Y)’, ‘아마도 그렇다(PY)’, ‘아마도 아니다(PN)’, ‘아니다(N)’, ‘정보 없음(NI)’으로 답하며, ‘그렇다’는 낮은 우려를 나타낸다. 이후 각 영역과 관련된 비뮴림에 대한 우려 수준은 ‘낮음’, ‘높음’, 또는 ‘불확실’로 판단한다. 영역의 모든 신호 질문에 대한 답변이 ‘그렇다’ 또는 ‘아마도 그렇다’인 경우, 우려 수준은 낮은 것으로 판단할 수 있다. 신호 질문이 ‘아니다’ 또는 ‘아마도 아니다’인 경우, 비뮴림위험 우려 가능성이 존

재한다. ‘정보 없음’ 범주는 판단을 하기에 보고된 자료가 불충분한 경우에만 사용해야 한다.

- 판단하는 데 사용된 정보(판단에 대한 뒷받침)를 기록함으로써, 투명하게 등급을 매기고 문헌고찰 저자들 사이에서 독립적으로 평가를 완료하여 토론을 용이하게 하도록 한다. ROBIS 사용자는 평가를 완료하기 위해 주제 내용과 방법론적 전문 지식을 모두 필요할 것이다.

가) 영역 1: 연구 적격 기준

- 영역 1은 일차연구 적격 기준이 사전에 지정되고 명확하며 문헌고찰 질문에 적절한지 여부를 평가하는 것을 목표로 한다. 문헌고찰 수행 이전에 어떤 연구를 포함시켜야 할지 결정해야 하며, 이 때 일차연구 자체의 특성과 결론에 따라 이루어지는 것이 아니라 사전에 결정된 적격 기준에 따라 일관성 있게 이루어져야 한다. 적격 기준이 적절하게 사전에 정해졌는지 여부를 평가하기 위해서는 프로토콜이나 등록 문서가 필요하다.
- 적격 기준에는 문헌고찰의 연구질문, 고려하는 연구설계, 인구집단/연구대상자, 중재/노출 및 비교중재/노출이 포함될 수 있으며 가능한 한 구체적이어야 한다. 문헌고찰 저자는 적격 기준에 특정 중재결과, 연구의 질, 또는 특정 날짜 범위 등과 같은 제한사항을 두고 적합 여부를 판단할 수 있는데, 때로는 이러한 제한사항이 연구 선택의 비돌림으로 작용할 수 있다.

(1) 신호 질문과 설명

1.1	문헌고찰이 사전에 정의된 연구목적과 적격 기준을 준수하였는가?
------------	---

설명

문헌고찰 프로토콜의 적격 기준에는 일반적으로 연구설계, 연구대상자 및 중재/노출의 유형을 포함해야 한다. 모든 제외 기준도 미리 정해야 한다.

- Y : 이러한 정보를 제공하는 프로토콜이 있는 경우
- PY : 프로토콜은 없지만 사전에 정의된 목표와 세부적인 적격 기준에 대한 정보가 제공되고, 적격 기준이 사전에 정해져 문헌고찰 중에 준수되었다고 믿을 만한 충분한 이유가 있는 경우
- PN : 프로토콜이 없는데 사전에 정의된 목표와 적격 기준에 대한 정보를 오직 문헌고찰 출판물에서만 확인할 수 있으며 이러한 사항이 사전에 정해져 연구 시작부터 준수되었다고 믿을 만한 이유가 없는 경우
- N : 이러한 사항 중 일부 또는 전부가 누락된 경우

1.2	적격 기준이 문헌고찰의 질문에 부합하는가?
------------	--------------------------------

설명

- 적격 기준은 문헌고찰 질문에서 비롯되어야 하며, 포함된 일차연구가 질문에 적합한지 판단이 가능하도록 인구집단에 대한 명확한 설명과 같은 세부사항을 제공해야 한다.
- 이 질문에 답하기 위해서 연구 내용에 대한 지식이 요구될 가능성이 높다.

1.3	적격 기준이 모호하지 않은가?
------------	-------------------------

설명

- 적격 연구의 특성에 대한 구체적인 정보가 제공되어야 하며 가능한 한 연구 유형, 인구집단, 중재, 비교중재 및 중재결과에 대한 모호성을 피해야 하고, 기준은 문헌고찰을 재현할 수 있을 정도로 충분히 상세해야 한다.
- 질병의 진단방법, 약물 용량, 투여 빈도, 동반 치료 등과 같은 특성, 그리고 중재/노출 및 비교중재/노출에 대한 구체적인 세부사항이 제시되어야 한다.
- PN or N : 문헌고찰에 제시된 적격 기준에 구체적인 의문(query)이 있을 경우

1.4	연구 특성에 따른 적격 기준의 제한이 적절한가?
------------	-----------------------------------

설명

- 연구 특성에 따라 적용되는 모든 제한사항은 명확하게 기술되고 확실한 근거를 제시해야 한다.
- 이 질문은 적격 기준이 문헌고찰 질문에 적합한지 여부를 다루는 위의 질문과는 다르다. 예를 들어 연구설계, 연구 발표 날짜, 연구의 크기, 연구 질 및 이용 가능한 중재결과 측정 등이 있을 수 있다.
- Y or PY : 충분한 정보가 있고 평가자가 제한사항이 적절하다고 판단하는 경우, 제한사항이 명시적으로 없는 경우는 Y
- PN or N : 제한사항이 정당화되지 않고 이러한 제약이 적절한지 판단하기에 충분한 정보가 없는 경우

1.5	정보의 출처에 따른 적격 기준의 제한이 적절한가?
-----	-----------------------------

설명

- 출판 상태, 형식(format), 언어, 자료의 가용성 등과 같은 정보의 출처에 대한 제한사항을 명확하게 기술하고 확실한 근거를 제시해야 한다.
- 이 질문은 검색 제한에 관한 영역 2의 질문과는 다르다.
- Y : 적격 기준이 충분히 상세하고 정보 출처에 대한 제한사항이 명시적으로 없는 경우.

(2) 연구 적격 기준을 결정하는 것에 대한 우려

낮음	문헌고찰 질문과 목표를 명확히 명시하고, 문헌고찰 과정에서 적절하고 상세한 적격 기준을 사전에 구체화하고 정당화하기 위해 상당한 노력을 기울였다.
높음	문헌고찰 질문에 대답하는데 중요하고 관련이 있는 연구들이 사전에 지정된 목표와 적격 기준이 부족하거나 부적절한 제한이 부과되었거나, 문헌고찰 질문을 다루는데 적절하지 않은 연구가 포함되었기 때문에 문헌고찰에서 제외되었을 가능성이 높다.
불확실	비플림위험에 대해 판단하기에 보고된 정보가 불충분하다.

(3) 등급부여 예시**a. 연구 적격 기준에서 높은 우려로 판단된 문헌고찰 ¹⁾****본문 기술 내용:***Abstract - Aim*

"To evaluate the accuracy of EUS in T staging of early rectal cancers."

Study Selection Criteria

"Only EUS studies confirmed by surgical histology were selected. EUS criteria used for T0 was tumour confined to the mucosa. From this pool, only studies from which a 2x2 table could be constructed for true-positive, false-positive, false negative and true-negative values were included."

1) Puli SR, Bechtold ML, Reddy JB, Choudhary A, Antillon MR. Can endoscopic ultrasound predict early rectal cancers that can be resected endoscopically? A meta-analysis and systematic review. Dig Dis Sci 2010;55(5):1221-9.

연구 적격 기준에 대한 평가 예시 ('높음'으로 평가)

신호질문	평가	이유
1.1 문헌고찰이 사전에 정의된 연구 목적과 적격 기준을 준수하였는가?	NI	사전에 정의된 목표와 적격 기준이 존재한다는 근거가 없었다.
1.2 적격 기준이 문헌고찰의 질문에 부합하는가?	PN	보고한 적격 기준은 적절하나 연구대상 인구집단 관련 일부 세부 사항이 부족했다. 따라서 적격 기준이 적절하다는 판단을 하는 것이 불가능하다.
1.3 적격 기준이 모호하지 않은가?	PN	적격 기준에 대한 세부 사항이 불충분했다. 특히 어떤 연구 인구집단이나 어떤 연구설계가 적합한지에 대한 세부사항은 없었다.
1.4 연구 특성에 따른 적격 기준의 제한이 적절한가?	NI	연구 특성에 따른 제한에 대한 세부사항은 제공되지 않았다. 전반적으로 적격 기준에 대한 세부 사항이 불충분하다.
1.5 정보의 출처에 따른 적격 기준의 제한이 적절한가?	PY	문헌고찰은 2x2표를 구성할 수 있는 연구로 제한되었다. 이에 대한 정당성이 제공되지 않았지만, 이는 진단 정확도(diagnostic test accuracy, DTA) 문헌고찰에서 흔히 볼 수 있으므로 합리적일 수 있다.
연구 적격 기준의 규정에 대한 우려	높음	적절한 연구가 문헌고찰에 포함되었는지 여부를 판단할 수 있는 연구 적격 기준에 대한 세부 정보가 부족했으며, 특히 적격 참여자에 대한 세부 정보가 부족했다. 기준이 사전 지정되었는지, 사후 적용되었는지도 불확실했다. 따라서 적격 기준의 규정에 대한 우려가 높다.

b. 연구 적격 기준에서 관련하여 낮은 우려로 판단된 고찰²⁾**본문 기술 내용:***Objective*

“To evaluate the association and predictive value of ultrasound measurements of amniotic fluid volume for adverse pregnancy outcome.”

Methods - Study selection

- ① Population. Pregnant women, with or without fetal growth restriction, no evidence of premature rupture of membranes, no evidence of congenital or structural abnormalities.*
- ② Index test. Any measure of amniotic fluid reported by the authors including AFI, amniotic fluid volume and maximum deepest pocket. Any threshold used to define low or high amniotic fluid as reported by the authors of the included studies was accepted.*
- ③ Outcome. Any reference standard looking at compromise of fetal or neonatal wellbeing; including: abnormal cord pH at birth, Apgar scores, perinatal death and composite outcomes such as adverse perinatal outcome. Any reference standard for fetal growth restriction or small for gestational age: Birthweight <10th, <5th, <3 rd centile, absolute birthweight thresholds, ponderal index.*
- ④ Study design. Observational studies in which the results of the test of interest are compared with the outcome findings as confirmed by a reference standard, allowing generation of a 2x2 table to compute indices of association and test accuracy for each available threshold. Case series of ten or fewer and case-control studies determined by outcome were excluded.”*

Data Extraction

“All studies had to state that they excluded rupture of membranes and congenital/structural anomalies due to the association of renal/urinary tract anomalies and karyotypic anomalies with abnormalities of liquor volume.”

2) Morris RK, Meller CH, Tambllyn J, Malin GM, Riley RD, Kilby MD, et al. Association and prediction of amniotic fluid measurements for adverse pregnancy outcome: systematic review and metaanalysis. BJOG: An International Journal of Obstetrics & Gynaecology 2014;121(6):686-99.

(4) 연구 적격 기준에 대한 평가 예시 ('낮음'으로 평가)

신호질문	평가	이유
1.1 문헌고찰이 사전에 정의된 연구 목적과 적격 기준을 준수하였는가?	Y	저자들은 논문의 초록과 연구배경에서 양수 부피의 초음파 측정과 부정적 임신결과 사이의 연관성과 예측값을 평가하는 것이 목적이라고 명확히 규정했다. 별도의 프로토콜에서는 “태아/신생아 위험과 태아 성장 제한을 예측하기 위한 양수 부피측정의 정확성은 무엇인가?”라는 구체적인 문헌고찰 질문을 제공했다. 저자들은 인구집단, 중재 검사, 중재결과, 연구설계의 세부사항을 적격 기준으로 제공했다.
1.2 적격 기준이 문헌고찰의 질문에 부합하는가?	Y	문헌고찰 질문은 태아/신생아 위험과 태아 성장 제한을 예측하기 위해 양수 부피의 정확도를 조사하는 연구를 파악할 필요가 있음을 제시했다. 논문에서 제시한 연구 적격 기준의 세부 내용은 문헌고찰 질문에 적합한 것으로 보였다.
1.3 적격 기준이 모호하지 않은가?	PY	연구설계의 종류는 분명히 명시되어 있었다. 관찰 연구가 포함되었다. 저자들은 (10례 이하의) 사례군 연구와 결과에 따라 결정된 환자 대조군 연구를 제외했다. 관심 인구집단의 명확한 세부 사항(태아 성장 제한이 있거나 없는 임신부, 조기양막 파열의 증거 없음, 선천적 또는 구조적 이상 징후 없음), 중재 검사 (AFI, 양수 부피 및 최심부 포켓 등 사용된 임계값에 관계없이 보고된 양수 측정치), 중재결과(태아 또는 신생아 건강의 위험을 살피는 모든 참고 표준의 몇가지 구체적인 예가 열거되었다. 태아 성장 제한 또는 작은 임신 연령에 대한 모든 참조 표준의 몇가지 구체적인 예가 열거됨)이 제공되었다.
1.4 연구 특성에 따른 적격 기준의 제한이 적절한가?	PY	연구설계의 유형과 연구 규모 문제에 기초한 제한사항이 명확하게 설명되었다. 추가 제한사항은 자료 추출 섹션에서 제공되었다. 비록 정당성이 제공되지는 않았지만, 그 제한은 적절해 보였다.
1.5 정보의 출처에 따른 적격 기준의 제한이 적절한가?	PY	언어 제한은 적용되지 않았다. 비록 정당성이 제공되지는 않았지만 자료의 가용성에 기초했을 일부 제한이 있었다. 정보의 출처에 대한 다른 제한은 기술되지 않았다. 그러므로 이 질문은 “아마도 그렇다”라고 대답했다.
연구 적격 기준의 규정에 대한 우려	낮음	모든 신호 질문에 대해 “그렇다” 또는 “아마도 그렇다”로 대답되었기 때문에 적격 기준의 규정에 대한 잠재적 우려는 확인되지 않았다. 문헌고찰 질문과 목표를 명확히 명시하고, 적절하고 상세한 적격 기준을 사전에 명시하고 정당화하기 위해 상당한 노력을 기울였다.

나) 영역 2: 연구의 검색 및 선택

- 영역 2는 포함 기준에 맞는 일차연구가 문헌고찰에 포함되지 않았는지 여부를 평가하는 것을 목적으로 한다. 가능한 한 많은 적격 연구를 파악하기 위한 민감도 높은 검색은 모든 문헌고찰의 핵심 요소이다.
- 검색에는 출판되었거나 출판되지 않은 연구결과를 파악할 수 있는 적절한 데이터베이스와 논문, 컨퍼런스 및 임상시험기록 등을 색인한 전자정보원이 포함되어야 하며, 적합한 연구 보고서를 파악하기 위한 데이터베이스 검색(예: 기존 고찰의 참고문헌 확인, 인용 검색, 수기 검색) 및 적절하고 민감한 검색 전략의 사용 포함되어야 한다.
- 검색 전략에는 프리텍스트 용어(예: 제목과 초록)와 관련 연구를 파악할 수 있는 적합한 주제 색인(예: MeSH 또는 Emtree)이 포함되어야 한다. 또한 검색할 주제와 자원에 따라 CAS 레지스트리 번호와 같은 추가 필드, 의약품 또는 기기의 상품명 등을 검색하는 것도 도움이 될 수 있다.
- 전략은 일반적으로 인구집단과 중재/중재검사/노출의 조합이다. RCT의 문헌고찰과 같은 특정 유형의 문헌고찰의 경우, 검색된 데이터베이스의 일부 또는 전체에 방법론적 또는 연구설계 필터를 포함하는 것도 적절할 수 있다. 그러나 진단 정확도 연구 같이 다른 유형의 문헌고찰의 경우, 필터를 사용하면 관련 연구가 누락될 가능성이 높아 피해야 한다.
- 문헌고찰과 관련된 검색 관행 및 콘텐츠 전문지식과 관련된 방법론적 지식이 없으면 검색 전략의 민감도를 평가하기 어려울 수 있다.
- 일반적으로 평가자는 연구질문을 포착하기 위해 사용된 개념이 기술될 수 있는 가능한 모든 방법을 다루기 위해 적절한 범위의 용어가 포함되는지 여부를 고려해야 한다. 예를 들어 유방암에 대한 문헌고찰을 위해 '유방암(breast cancer)'이라는 용어를 키워드로 검색하는 것만으로는 충분하지 않고, 검색 전략에 breast neoplasm, carcinoma, tumor와 같은 용어와 관련 MESH, 그리고 이들의 적절한 조합 및 검색필드 적용 등이 필요하다.

(1) 신호 질문과 설명

2.1	출판/미출판 보고서에 대한 적절한 범위의 데이터베이스/전자자료원이 검색에 포함되었는가?
------------	---

설명

- 평가자는 적절한 범위의 데이터베이스 조합을 판단할 필요가 있는데, 이는 문헌고찰 주제에 따라 달라질 것이다.
- 최소한 MEDLINE과 EMBASE 검색을 해야 하고, 학회발표 자료, 연구 레지스트리도 검색도 검토해야 한다.
- 적절한 데이터베이스 범위에 대한 지침은 코크란 핸드북 같은 체계적 문헌고찰 지침 또는 CRD (Center for Review and Dissemination) 웹 사이트에서 확인할 수 있다.
(http://www.york.ac.uk/inst/crd/finding_studies_systematic_reviews.htm).

2.2	관련 보고서를 찾기 위해 데이터베이스 검색에 추가적인 방법이 사용되었는가?
------------	--

설명

- 인용한 문헌 검색, 전문가 접촉, 참고문헌 확인, 수기 검색 등의 추가 방법이 수행되어야 한다.

2.3	검색 전략의 용어와 구조가 가능한 한 많은 적격 연구를 검색할 가능성을 높였는가?
------------	--

설명

- 검색을 복기하기에 충분하도록 사용된 모든 검색어를 보여주는 전체 검색 전략이 제시되어야 이 질문에 대한 완전한 파악이 가능하다.
- PY or PN : 검색어 조합 방법에 대한 표시가 없이 검색어 목록만 나열하는 식의 제한된 세부사항만 제공되는 경우
- 평가자는 검색 전략에 주제에 대한 적절한 범위의 용어가 포함되었는지 여부, MeSH와 제목 및 초록에 포함된 단어의 조합이 사용되었는지 여부, 적용된 필터가 적절했는지의 여부를 고려해야 한다.
- N : 예를 들어, DTA 문헌고찰의 경우 필터의 사용은 관련 연구를 누락한 것으로 나타나므로 해당 필터를 포함하는 전략인 경우
- 검색 전략의 비판적 평가에 대한 지침은 PRESS 근거 기반 점검표에서 찾을 수 있다.
(<http://ejournals.library.ualberta.ca/index.php/EBLIP/article/view/7402>).

2.4	날짜, 출판 형식 또는 언어에 따른 제한이 적절하였는가?
------------	--

설명

- 이 항목을 완전히 판단하려면 이 질문의 세 가지 구성 요소(즉, 날짜, 출판 형식, 언어) 모두에 대한 정보가 필요하다.
- Y : 검색 전략에 제한사항이 없는 경우 (선정기준에 대한 제한에 관한 영역 1(1.5)의 질문과는 다름에 유의)
- N : 언어에 기반 한 논문 제한(예: 영어 문헌으로 제한)이나 출판 형식(예: 전문(full text)으로 출판된 연구로 제한)으로 제한한 경우 (이러한 제한은 대부분 적절하지 않기 때문임)
- 날짜에 대한 제한은 적절할 수 있지만 이 질문에 “그렇다”로 대답하기 위해 명확하게 설명된 근거가 있어야 한다.

2.5	연구 선택의 오류를 최소화하기 위한 노력이 이루어졌는가?
-----	--

설명

- 선택을 위해 제목과 초록으로 선별하는 과정과 전문(full text)을 평가하는 과정 모두 이 질문에서 다룬다.
- Y : 제목과 초록을 적어도 두 명의 검토자가 독립적으로 심사해야 하며, 전문 평가에도 적어도 두 명의 검토자가 참여해야 한다(각자 독립적으로 하거나 또는 한 명이 평가를 수행하고 두 번째 검토자가 결정을 확인해야 한다).

연구 검색 및 선택에 대한 우려

낮음	영역 1에서 평가한 문헌고찰 질문과 적격 기준을 고려할 때, 민감하고 적절한 검색 전략을 사용하여 다양한 검색 방법을 통해 가능한 한 많은 관련 연구를 파악하고 포함 연구를 선택할 때 비플림과 오류를 최소화하기 위한 조치를 취했다.
높음	일부 적격 연구가 문헌고찰에 제외되었을 가능성이 높다.
불확실	비플림위험에 대한 판단을 할 수 있는 정보가 충분하지 않다.

(2) 평가의 예

a. 연구의 검색 및 선택 방법에서 높은 우려로 판단된 문헌고찰³⁾

본문 기술 내용:

“A research librarian searched Ovid MEDLINE from 1947 to August 2012, the Cochrane Library Database (through the first quarter of 2012), Embase (1976 to August 2012), Scopus (1960 to August 2012), PsychINFO (1806 to August 2012), clinical trials registries, and grants databases. At least 2 reviewers independently evaluated studies for inclusion. Non-English language articles were excluded. Included studies published as conference abstracts were only used in sensitivity analyses.”

3) Chou R, Hartung D, Rahman B, Wasson N, Cottrell EB, Fu R. Comparative effectiveness of antiviral treatment for hepatitis C virus infection in adults: a systematic review. Ann Intern Med 2013;158(2):114-23.

연구 검색 및 선택에 대한 평가 예시 ('높음'으로 평가)

신호질문	평가	이유
2.1 출판/미출판 보고서에 대한 적절한 범위의 데이터베이스/전자자료원이 검색에 포함되었는가?	Y	MEDLINE, EMBASE, Scopus, PsycInfo 및 Cochrane 라이브러리를 검색하였다(Cochrane 라이브러리에 어떤 검색원이었는지는 불확실). 이것은 적절한 범위라고 판단되었다. 임상시험 레지스트리 및 연구비 데이터베이스도 검색하였다.
2.2 관련 보고서를 찾기 위해 데이터베이스 검색에 추가적인 방법이 사용되었는가?	N	아니다
2.3 검색 전략의 용어와 구조가 가능한 한 많은 적격 연구를 검색할 가능성을 높였는가?	NI	전체 검색 전략을 보고하지 않았고 검색어에 대한 세부 정보가 없었다. 따라서 이 질문에 대한 평가의 기초가 되는 정보가 없었다.
2.4 날짜, 출판 형식 또는 언어에 따른 제한이 적절하였는가?	N	영어로만 제한한 문헌고찰로 출판비둘림의 가능성이 있다.
2.5 연구 선택의 오류를 최소화하기 위한 노력이 이루어졌는가?	PY	포함여부 평가는 최소 2명의 검토자가 독립적으로 실시한 것으로 알려졌다. 그러나 검색결과와 선별과 전문평가에 모두 이를 적용했다는 것은 명확하지 않았다. 따라서 이 항목은 “아마도 그렇다”로 평가되었다.
연구 파악과 선택방법에 대한 우려	높음	문헌고찰을 영어로만 제한한 것은 관련 연구가 이번 문헌고찰에 포함되지 않았을 위험이 크다는 것을 의미한다. 검색 전략이 목적에 부합하는지 판단할 수 없다.

b. 연구의 파악 및 선택 방법에서 낮은 우려로 판단된 문헌고찰⁴⁾

본문 기술 내용:

“The following sources were searched from inception to October 2011: MEDLINE; EMBASE; Cumulative Index To Nursing And Allied Health Literature (CINAHL); The Cochrane Central Register of Systematic Reviews; The Cochrane Central Register of Controlled Trials; DARE; MEDION; SIGLE; Index of Scientific and Technical Proceedings, Web of Science and ClinicalTrials.gov database. The search consisted of keywords and MeSH terms relating to the tests under investigation combined with MeSH terms of ‘Prenatal

4) Morris RK, Meller CH, Tamblin J, Malin GM, Riley RD, Kilby MD, et al. Association and prediction of amniotic fluid measurements for adverse pregnancy outcome: systematic review and metaanalysis. BJOG: An International Journal of Obstetrics & Gynaecology 2014;121(6):686-99.

Diagnosis’, ‘Ultrasonography’, ‘Amniotic Fluid’ and ‘Pregnancy Outcome’. The full search strategy is shown in the Appendix S2. The reference lists of all included primary and review articles were examined to identify cited articles not captured by electronic searches. No language restrictions were applied. The database was scrutinised by two reviewers (RKM, CHM) and full articles likely to meet the selection criteria were obtained. Translations were obtained for non-English articles. Three reviewers made the final inclusion/exclusion decisions according to adherence to the following criteria.”

연구 검색 및 선택 평가 예시 (‘낮음’으로 평가)

신호질문	평가	이유
2.1 출판/미출판 보고서에 대한 적절한 범위의 데이터베이스/전자자료원이 검색에 포함되었는가?	Y	MEDLINE, EMBASE, CINAHL, SIGLE, Index of Scientific and Technical Proceedings, ClinicalTrials.gov 데이터베이스, Cochrane 라이브러리, DARE, MEDION을 검색하였다. Web of Science 또한 검색하였으나 이 플랫폼의 어떤 데이터베이스를 검색하였는지는 명확하지 않다. 적절한 범위의 자료원을 출판/미출판 보고서를 검색하려는 시도도 포함되었다. 그러나 Cochrane Central Register of Systematic Reviews는 기록오류일 가능성이 높으며 저자들은 CDSR을 검색했다.
2.2 관련 보고서를 찾기 위해 데이터베이스 검색에 추가적인 방법이 사용되었는가?	Y	포함된 모든 일차 연구와 고찰의 참고문헌 리스트를 조사하고 전자검색으로 파악되지 않은 인용문헌을 파악하였다.
2.3 검색 전략의 용어와 구조가 가능한 많은 적격 연구를 검색할 가능성을 높였는가?	Y	상세한 검색전략은 웹 부록을 통해 제공하였다. 인구 집단(임산부)과 중재(초음파)의 조합은 적절하지 않은 제한(예: 연구설계 필터) 없이 민감해보였다.
2.4 날짜, 출판 형식 또는 언어에 따른 제한이 적절하였는가?	Y	데이터베이스의 시작일부터 검색하였고 언어제한을 두지 않았다. 회색문헌을 파악하는 단계도 포함되었다. 어떤 제한도 두지 않은 것으로 판단되어 “그렇다”라고 평가하였다.
2.5 연구 선택의 오류를 최소화하기 위한 노력이 이루어졌는가?	PY	제목과 초록을 선별하는 과정과 전문을 평가하는 과정은 다수의 평가자를 포함했다고 보고하였다. 그러나 평가자가 독립적으로 수행하였는지는 명확하지 않았다. 따라서 “아마도 그렇다”라고 평가하였다.
연구 파악과 선택방법에 대한 우려	낮음	모든 신호질문에 대해 “그렇다” 또는 “아마도 그렇다”로 평가되었고 따라서 잠재적 비돌림의 영역도 파악되지 않았다. 이 문헌고찰은 관련연구를 포함하고 있을 가능성이 높다고 판단하였다.

다) 영역 3: 자료 수집 및 연구 평가

- 영역 3의 목적은 자료의 수집이나 비뚤림위험 평가 과정 중에 비뚤림이 있었는지 여부를 평가하는 것이다. 엄격한 자료 수집을 위해서는 프로토콜 단계에서 계획하고, 예비로 시행해 본 구조화된 자료 수집 양식을 사용해야 하며, 연구결과의 합성과 해석에 기여할 모든 자료를 수집해야 한다.
- 여기에는 수치와 통계자료 그리고 좀 더 일반적인 일차연구 특성들(연구설계, 연구비출처, 세팅, 연구대상자 특성, 선택 기준, 중재/노출/중재 검사의 세부 정보, 연구대상자 탈락 등)도 포함되어야 한다.
- 수치형 자료의 유형은 문헌고찰에 포함된 연구설계에 따라 달라진다. 합성할 때 필요한 형식으로 자료를 이용할 수 없는 경우, 문헌고찰 저자는 이러한 자료를 어떻게 확보하였는지 보고해야 한다(예를 들어 일차 연구의 저자에게 연락). 필요한 자료를 도출하기 위해 적절한 통계 변환을 사용할 수 있다.
- 자료 추출에는 얼마든지 오류가 있을 수 있기 때문에 문헌고찰 저자간에 중복으로 자료 추출하거나 한 사람이 자료 추출 후 다른 사람이 엄격하게 체크를 함으로써 무작위 오류 및 잠재적 비뚤림을 방지할 수 있어야 한다.
- 포함된 연구의 타당성 평가는 문헌고찰에 포함된 일차 연구의 설계를 고려한 적절한 기준을 통해서 이루어져야 하며, 이를 위해 해당 일차 연구설계에 맞는 검증된 도구를 사용하거나 중요한 잠재적 비뚤림의 원인이 될 수 있는 목록화된 기준을 사용해야 한다. 자료 추출과 마찬가지로 비뚤림위험 평가 과정에서 비뚤림 또는 오류가 발생할 수 있기 때문에 비뚤림위험 평가는 두 명의 검토자가 독립적으로 작업하거나 한 사람이 평가하고 다른 사람이 체크하도록 해야 한다.
- 자료 추출과 비뚤림위험 평가에서 모두 불일치 및 이의 해결 과정이 보고되어야 한다.

3.1	자료 수집 오류를 최소화하기 위해 노력하였는가?
-----	----------------------------

설명

- 자료 수집 과정에서 비뚤림과 오류를 최소화하기 위해서는 최소한 두 명의 검토자가 필요하며 파일럿 과정을 거친 구조화된 자료 추출 양식을 사용해야 한다.
- 이상적으로는 독립적으로 수행해야 하지만 한 명의 검토자가 추출하고 두 번째 검토자의 체크도 허용된다.

3.2 연구결과를 해석할 수 있기에 충분한 연구 특성 정보가 문헌고찰 저자와 독자에게 제공되었는가?

설명

- 연구 특성 정보를 통해 이질성에 대한 적절한 조사와 연구결과의 적용 가능성을 고려할 수 있으므로, 연구 특성 정보는 문헌고찰의 연구결과를 해석하는 데 필수적이다.
- 이 정보는 자료 수집에 대한 정보가 보고되는 방법 섹션에 보고되거나 포함된 연구 표의 특성에서 사용 가능하거나 연구결과 본문에 요약될 수 있다.
- 수집된 정보가 문헌고찰의 분량 문제로 간행물에 제시되지 않을 수 있기 때문에 이 질문에 대답하기 어려울 수 있으며, 때로는 웹 부록과 같이 추가적인 출처에 접근해야 할 수도 있다.

3.3 합성에 사용되도록 타당한 모든 연구결과가 수집되었는가?

설명

- 적절하게 합성할 수 있도록 충분한 연구결과를 추출해야 한다. 가능하면 문헌고찰 저자는 합성에 필요한 자료와 형식을 보고해야 한다. 예를 들어 DTA 문헌고찰을 위해 검사 수행의 2x2 자료를 추출하고, 관찰 연구의 문헌고찰을 위해 95% 신뢰 구간을 갖는 보정된 효과 추정치, 보정된 변수의 세부 사항 및 분석 방법이 추출될 수 있다.
- 이러한 자료가 방법 섹션에 명시적으로 보고되지 않은 경우 연구결과 표, 숲 그림과 같은 그래프 또는 본문에 보고된 자료에서 자료 추출을 할 수 있다.
- 사용 가능한 경우 웹 부록 또는 문헌고찰 프로토콜과 같은 추가 출처에 접근해야 할 수도 있다. 문헌고찰에 포함된 모든 일차 연구에서 합성에 필요한 형식으로 자료를 보고하는 것은 매우 드물다.
- Y : 합성에 필요한 형식으로 보고되지 않은 연구결과 자료를 어떻게 얻었는지 설명하기 위한 자세한 정보가 방법 섹션에 포함된 경우 (예: 보고된 데이터에서 추정/변환, 혹은 추가 정보를 위해 저자에게 문의)

3.4 비뚤림위험(또는 방법론적 질)이 적절한 기준을 사용하여 공식적으로 평가되었는가?

설명

- 공식적인 비뚤림위험 평가는 모든 문헌고찰의 필수 구성 요소이다.
- N : 비뚤림위험이 공식적으로 평가되지 않은 경우
- Y : RCT 평가를 위한 Cochrane 비뚤림위험 도구, 또는 DTA 연구를 위한 QUADAS-2와 같이 출판된 도구가 적절한 설계에 사용된 경우.
- 그러나 문헌고찰이 단순히 평가 질문을 나열하거나 출판되지 않은 도구 또는 더 이상 권장되지 않는 도구가 사용되는 경우에는 복잡해지는데, 평가자는 연구설계에 따라 사용된 기준이 일차 연구에서 잠재적 비뚤림의 원인을 식별할 수 있는지 여부를 판단해야 한다.
- 예를 들어, Jadad 점수는 매우 널리 사용되었지만 RCT에서 주요 잠재적 비뚤림의 가능성을 나타내는 배경 은폐는 포함하지 않는다. 이 질문에 “그렇다”라고 대답하려면 Jadad 척도를 사용한 문헌고찰은 추가로 배경 은폐를 평가해야 한다.

3.5 비뚤림위험 평가의 오류를 최소화하기 위한 노력을 하였는가?

설명

- 자료 수집과 마찬가지로 비뚤림위험 평가에는 최소한 두 명의 검토자가 포함되어야 한다. 이상적으로는 독립적으로 수행해야 하지만 한 명이 평가하고 다른 사람이 체크해도 된다.

자료수집과 연구평가에 대한 우려

낮음	영역 2에서 평가된 문헌고찰에 포함된 연구는 2명의 검토자가 적절한 기준, 자료 추출 및 비뚤림위험 평가를 사용하여 비뚤림위험을 평가했으며 관련 연구 특성과 결과를 추출했다.
높음	자료 수집 또는 비뚤림위험 평가 과정을 통해 일부 비뚤림이 발생할 수 있다.
불확실	비뚤림위험에 대한 판단을 보고하기에는 정보가 충분하지 않다.

(1) 평가의 예시

a. 자료 수집 및 연구 평가에서 높은 우려로 판단된 문헌고찰⁵⁾

본문 기술 내용:

“Tables of 2×2 were constructed with the data extracted from each study. Two authors independently searched and extracted the data into an abstraction form. Any differences were resolved by mutual agreement... There is no consensus on how to assess studies designed without a control arm. Hence, these criteria do not apply to studies without a control arm. Therefore, for this meta-analysis and systematic review, studies were selected based on completeness of data and inclusion criteria.”

코멘트: 포함된 연구의 특성표에는 연구설계(후향적, 전향적), 암의 유형(모든 직장암) 및 확인적 수술(confirmatory procedure, 모든 수술)에 대한 세부 정보만 포함된다.

5) Puli SR, Bechtold ML, Reddy JB, Choudhary A, Antillon MR. Can endoscopic ultrasound predict early rectal cancers that can be resected endoscopically? A meta-analysis and systematic review. Dig Dis Sci 2010;55(5):1221-9.

자료 수집 및 연구 평가에 대한 평가 예시 ('높음'으로 평가)

신호질문	평가	이유
3.1 자료 수집 오류를 최소화하기 위해 노력하였는가?	Y	추출 양식을 사용하여 두 명의 검토자가 독립적으로 데이터 추출을 수행했다. 불일치는 합의에 의해 해결되었다.
3.2 연구결과를 해석할 수 있기에 충분한 연구 특성 정보가 문헌고찰 저자와 독자에게 제공되었는가?	N	요약표는 단지 연구설계(후향적, 전향적), 암의 유형(모든 직장암) 및 확인적 수술(모든 수술)에 대한 정보만 포함한다. 본문에 연구설계나 인구집단에 대한 더 이상의 상세한 설명이 보고되지 않았다.
3.3 합성에 사용되도록 타당한 모든 연구결과가 수집되었는가?	Y	각 연구에서 추출된 2x2 자료는 진단적 정확도 및 관련 신뢰 구간의 모든 측정값을 계산하고 메타 분석을 수행하기에 충분하다.
3.4 비플림위험(또는 방법론적 질)이 적절한 기준을 사용하여 공식적으로 평가되었는가?	N	연구의 질은 공식적으로 평가되지 않았다. 저자들은 이것이 대조군 없는 연구를 평가하는 방법에 대한 합의가 없기 때문이며, 그러나 이용할 수 있는 DTA 연구 평가용 도구가 있었기 때문이라고 진술했다.
3.5 비플림위험 평가의 오류를 최소화하기위한 노력을 하였는가?	N	연구의 질은 공식적으로 평가되지 않았다.
자료수집과 연구평가에 대한 우려	높음	공식적인 질 평가의 부재는 포함된 연구에서 비플림위험이 불확실하다는 것을 의미한다. 독자가 결과를 해석할 수 있는 충분한 연구 세부 사항이 없다. 따라서 이 문헌고찰에 대한 자료 수집 및 연구 평가 모두에서 비플림위험이 높다.

b. 자료 수집 및 연구 평가에서 낮은 우려로 판단된 문헌고찰⁶⁾

본문 기술 내용:

“One investigator abstracted details about the study design, population, setting, interventions, analysis, follow-up, and results. A second investigator reviewed data for accuracy. Two investigators independently applied predefined criteria to assess study quality as good, fair, or poor. Discrepancies were resolved through consensus.”

6) Chou R, Hartung D, Rahman B, Wasson N, Cottrell EB, Fu R. Comparative effectiveness of antiviral treatment for hepatitis C virus infection in adults: a systematic review. *Ann Intern Med* 2013;158(2):114-23.

자료 수집 및 연구 평가에 대한 평가 예시 ('낮음'으로 평가)

신호질문	평가	이유
3.1 자료 수집 오류를 최소화하기 위해 노력하였는가?	PY	한 명의 검토자가 자료 수집을 수행했으며 이는 두 번째 검토자가 정확성을 검사했다. 두 번째 검토자가 추출된 자료가 정확한지 여부만 확인했는지 또는 관련 정보를 놓치지 않도록 논문을 자세히 읽었는지 여부는 확실하지 않다. 따라서 이 질문은 “그렇다”가 아니라 “아마도 그렇다”로 대답했다.
3.2 연구결과를 해석할 수 있기에 충분한 연구 특성 정보가 문헌고찰 저자와 독자에게 제공되었는가?	Y	상세한 연구 특성 및 결과표는 문헌고찰 저자와 독자가 결과를 해석하기에 충분한 정보를 부록에 제공하였다.
3.3 합성에 사용되도록 타당한 모든 연구결과가 수집되었는가?	Y	이분형 자료는 각 중재군의 사건 수 및 총 환자 수로 추출되었다; 이는 상대 위험을 계산하는 데 사용되었다. 이것은 합성에 사용하기에 적합했다.
3.4 비뿔림위험(또는 방법론적 질)이 적절한 기준을 사용하여 공식적으로 평가되었는가?	PY	비뿔림위험 기준에 대한 상세내용은 제공되지 않았지만 문헌고찰 저자들은 적절한 기준을 포함하는 미국 USPSTF (US Preventive Services Task Force) 방법론 지침 (2001)을 인용하였다.
3.5 비뿔림위험 평가의 오류를 최소화하기위한 노력을 하였는가?	Y	두 명의 검토자가 독립적으로 연구의 질을 평가하였다.
자료수집과 연구평가에 대한 우려	낮음	모든 신호질문은 “그렇다” 또는 “아마도 그렇다”로 평가 되었으므로 잠재적인 비뿔림 영역은 확인되지 않았다. 따라서 자료 수집 및 연구 평가의 검토 과정은 문헌고찰에서 비뿔림이 생기지 않을 것이다.

라) 영역 4: 합성 및 연구결과

- 영역 4는 포함된 일차 연구의 자료를 통합(질적 또는 양적 합성)하기로 결정한 경우, 문헌고찰 저자들이 적절한 방법을 사용했는지 평가하는 것을 목표로 한다.
- 질적 또는 양적 합성에서 고려해야 할 가장 중요한 측면은 다음과 같다. (i) 분석 방법이 문헌고찰에서 제안된 연구질문에 적합한지; (ii) 연구 간 변이(이질성)가 고려되는지 여부; (iii) 일차 연구에서 비뚤림이 고려되는지 여부; (iv) 합성되는 일차 연구의 정보가 완전한지 (특히, 예를 들어, 출판 또는 보고 비뚤림으로 인해 누락된 데이터가 사용가능한 데이터와 체계적으로 다를 위험이 있는지); (v) 문헌고찰 저자들이 연구결과를 보고하는 방식에 비뚤림이 있는지 여부.
- 추정 방법의 선택과 같은 메타분석 방법의 기술적 측면은 중요한 고려사항이 될 것 같지 않다. 그러나 표준오차를 표준편차로 해석하거나, 짝지은(matched) 또는 군집화된 데이터와 같은 설계 문제에 대해 보정하지 못하거나, 로그가 아닌 위험 비율에 표준 가중 평균 접근법을 적용하는 등의 실수가 중요할 수 있다.

4.1	포함되어야 하는 모든 연구가 자료합성에 포함되었는가?
------------	--------------------------------------

설명

- 합성할 때 일반적으로 연구질문과 관련된 데이터를 수집했다고 알려진 모든 연구를 포함하려고 해야 한다.
- 이 질문은 연구를 찾지 못해서 합성에 누락된 경우 외에(영역 2와 신호질문 4.5에서 다룸), 문헌고찰에 포함된 연구에 대해 다음의 세 가지 상황을 다룬다. (i) 포함된 연구의 특정 연구결과를 문헌고찰 저자가 이용할 수 없다. (ii) 문헌고찰 저자가 이용할 수 있는 자료를 수집하거나 처리하지 못했다. (iii) 문헌고찰 저자가 의도적으로 연구결과를 배제하였다.
- 각각의 자세한 사항은 부록의 내용 참고. 신호질문 4.5 참조.

4.2	사전에 정의된 모든 분석을 보고하였거나 이탈(departure)에 대해 설명하였는가?
------------	--

설명

- 이 질문의 목적은 문헌고찰 저자가 선호하지 않은 연구결과를 보고하지 않거나 대체하는 방식으로 분석하거나 분석 방법을 선택하여 발생하는 비뚤림을 확인하는 것이다.
- Y : 문헌고찰이 공개되거나 접근가능한 프로토콜을 따른 경우(예: PROSPERO, Cochrane 리뷰 프로토콜).
- PY : 사전에 정의된 분석이 수행되었다고 보고한 경우, 예를 들어 방법 섹션이 엄격해 보이고 언급된 모든 분석을 연구결과에서 다루고 있는 경우.
- N : 사전에 정해진 프로토콜이 명확히 없는 경우.
- NI : 문헌고찰에 프로토콜의 유무에 대한 언급이 없는 경우

4.3	포함된 연구들의 연구질문, 연구설계, 연구결과의 특성과 유사성에 근거해보았을 때 합성이 적절하였는가?
------------	---

설명

- 이 질문은 주로 합성 방법론을 다룬다. 양적 합성이 수행되는 경우(메타 분석), 이 질문은 이를 수행하는 것이 적절한지 여부와 사용된 통계적 방법을 다룬다.
- 서술적 방법으로 합성하였다면 이 질문은 양적 합성이 더 적절한지 여부와 서술적 접근 방식이 적절한지 여부를 다룬다. 양적 합성이 적절한지에 대한 결정은 주로 연구질문, 설계 및 중재결과(측정되는 경우)에서 연구들의 유사성에 의해 결정된다('임상적 이질성', '임상적 다양성').
- Y : 각 연구에서 적절한 가중치를 포함하여 연구를 통계적으로 통합하는 '허용된' 방법이 사용된 경우
- N : 일반적으로 모든 연구에 동일한 가중치를 부여하는 방법을 권장하지 않음에도 그러한 방법이 사용된 경우
- 진단적 검사 정확도에 대한 연구에서 일반적인 방법은 민감도와 특이도 사이의 상관관계를 설명하기 위한 이변량 접근법이다. 다변량, 단단계, 베이지안(Bayesian) 또는 네트워크 메타 분석 방법을 사용하는 메타 분석에는 통계적 조언이 필요할 수 있다.

4.4	연구간 변이(이질성)가 작거나 자료의 합성 중 다루어졌는가?
------------	--

설명

- 이 질문은 연구 특성에서의 변이보다 연구결과에서의 변이에 대한 질문이다. 연구 간 변이는 숲 그림처럼 시각적으로 평가될 수 있고, Cochran의 x^2 검증 또는 Q 같은 통계적 검사를 사용하고, τ^2 같은 연구 간 변이의 측정값을 사용하거나, I² 비일관성 통계량 같이 신뢰 구간 겹침에 대한 측정값을 사용한다.
- N : 변이의 원인을 탐색하기 위해 추가 분석 없이 고정 효과 메타 분석을 사용한 경우
- Y : 이질성을 조사하기 위해, 그리고 하위그룹/메타 회귀 분석을 실행하기 위해 랜덤 효과 모델을 적절하게 사용한 경우 혹은 임상적 이질성으로 인해 통계적 통합이 부적절하다는 것을 기초로 서술적 합성이 수행된 경우(통합하지 않는 것으로 이질성이 처리된 것으로 봄).

4.5	연구결과가 견고하였는가? 예를 들어 깔대기 그림(funnel plot)이나 민감도 분석으로 연구결과의 견고성을 보여주었는가?
------------	--

설명

- 이 질문은 자료의 합성에 의한 연구결과의 불안정성과, 합성 방법을 변경함으로써 연구결과가 변할 수 있는지 여부를 다룬다. 깔때기 그림은 효과 크기와 연구 크기 사이의 관계를 조사하기 위해 사용되는데, 깔때기 그림이 비대칭일 때 표준 고정 효과와 랜덤 효과 방법은 서로 다른 결과를 산출하므로, 민감도 분석의 한 형태로 두 가지 방법을 수행하는 것도 좋다.
- 또는 절삭과 채움 방법(trim-and-fill method), 회귀 접근법 또는 선택 모델의 사용을 통해 다양한 유형의 출판 비뮴림이 연구결과에 미치는 잠재적 영향을 조사할 수 있다.
- 민감도 분석의 다른 예로는 누락된 데이터의 대체(impute), 영향이 큰 단일 연구의 배제, 비뮴림위험이 높은 연구 제거 등이 있다.
- 민감도 분석을 했다는 것 자체만으로는 이 질문에 "그렇다"의 답하는 데 적절하지 않은 것으로 간주된다. 평가자는 민감도 분석으로 연구결과의 견고성을 입증했는지 여부로 결정해야 한다.
- 서술적 합성이 수행된 경우, 평가자는 연구 요약에 대한 다양한 접근방식이 서로 다른 결론을 도출할 수 있는지와, 단일 연구가 결론을 이끌어내고 있는지 여부를 고려해야 한다.

4.6	일차 연구에서 비뚤림위험이 미미했거나, 비뚤림위험이 자료 합성 과정에서 고려되었는가?
-----	--

설명

- N : 문헌고찰 저자가 연구 구성에 중요한 비뚤림이 있다고 판단하는 경우(예 : 체계적 문헌고찰에서 비뚤림위험이 평가되지 않은 경우). 혹은 비뚤림위험이 평가되었지만 문헌고찰 저자들이 그것을 연구결과/결론에 포함시키지 않은 경우, 혹은 비뚤림이 연구결과의 토론의 일부로만 다루어지는 경우,
- Y : 문헌고찰 저자가 “비뚤림위험 낮음” 등급을 받은 모든 연구를 포함하거나 비뚤림의 위험이 높은 연구가 사용된 경우 민감도 분석/보정 방법을 쓴 경우

연구결과를 합성하는데 사용된 방법과 관련된 우려

낮음	합성은 비뚤린 연구결과를 낳지 않을 것 같다. 왜냐하면 자료의 한계를 극복했거나, 그 연구결과가 설득력이 있어 그 한계가 거의 영향을 미치지 않았다.
높음	합성은 비뚤린 연구결과를 초래할 가능성이 있다. 왜냐하면 (i) (연구 내 그리고/또는 연구 간) 잠재적 비뚤림이 무시되었거나, (ii) 연구 간 중요한 변동이 설명되지 않았거나, (iii) 방법론에 중요한 부적절성이 있었거나, (iv) 연구결과가 우려를 유발하는 방식으로 불완전하게 보고되었다.
불확실	보고된 결과가 불충분하기 때문에 비뚤림위험을 평가할 수 없다.

(1) 평가 예시

a. 합성에서 높은 우려로 판단된 문헌고찰⁷⁾

합성에 대한 평가 예시 ('높음'으로 평가)

신호질문	평가	이유
4.1 포함되어야 하는 모든 연구가 자료 합성이 포함되었는가?	NI	저자는 35개의 연구가 문헌고찰과 관련이 있으며, 그 중 29개는 메타 분석에서 사용할 수 있다고 말했다. 이를 확인하기 위한 다른 정보는 제공되지 않았다.
4.2 사전에 정의된 모든 분석을 보고하였거나 이탈(departure)에 대해 설명하였는가?	NI	참고한 프로토콜에서는 어떤 분석도 미리 정의되지 않았다. 본문에서는 더 이상의 정보가 제공되지 않았다.
4.3 포함된 연구들의 연구질문, 연구설계, 연구결과와 특성과 유사성에 근거해보았을 때 합성이 적절하였는가?	N	저자는 “메타 분석이 수행되었다”고만 진술했다(p.11). 저자는 사건 및 인구집단을 요약하고 합계를 사용하여 요약 오즈비를 생성하였고, 연구 내 비교를 활용하지 않았다. 저자는 연구를 통합한다고 기술했지만, 연구 전반에 걸쳐 임상적 다양성이 있는 것으로 보인다.
4.4 연구간 변이(이질성)가 작거나 합성 중 다루어졌는가?	PN	통계적 이질성에 대한 자세한 내용은 제공되지 않았다. 그러나 저자는 높은 비플림에 대한 하위군 분석, 파트너 연구 및 무작위 인구집단 연구에 대해 논의하였다. 또한 연구들은 지역사회 포피 절제(circumcision) 유병률에 따라 분류되었다. 하위그룹에 대한 자세한 내용은 제시되지 않았다.
4.5 연구결과가 견고하였는가? 예를 들어 깔대기 그림(funnel plot)이나 민감도 분석으로 연구결과의 견고성을 보여주었는가?	PN	저자는 문제가 될 수 있는 몇몇 연구에 대해 논의했다. “이러한 가정은 잘못된 것일 수 있지만, 이러한 연구의 포함 또는 배제는 전체 결과에 영향을 미치지 않았다.”(p.11). 이 분석에 대한 오즈비 및 신뢰 구간은 보고되지 않았다.
4.6 일차 연구에서 비플림위험이 미미했거나, 비플림위험이 자료 합성 과정에서 고려되었는가?	N	그 연구는 질이나 비플림위험에 대해 명시적으로 평가되지 않았다. 비플림은 합성에서 명시적으로 다루어지지 않았다.
합성과 연구결과에 대한 우려	높음	개별 연구는 비플림위험에 대해 평가되지 않았으며, 합성에서 잠재적 비플림도 설명되지 않았다. 저자는 메타분석에서 위험한 접근 방식인 연구 내 비교를 분석하기보다는 연구 전반에 걸쳐 중재 집단을 요약하였다. 분석에서는 이질성에 대한 논의나 평가가 없었다.

7) Van Howe RS. Circumcision and HIV infection: review of the literature and meta-analysis. Int J STD AIDS 1999;10(1):8-16.

b. 합성에서 낮은 우려로 판단된 문헌고찰⁸⁾

합성에 대한 평가 예시 ('낮음'으로 평가)

신호질문	평가	이유
4.1 포함되어야 하는 모든 연구가 자료 합성이 포함되었는가?	NI	문헌선정 흐름도는 고찰에 포함된 43개의 연구를 보고했다. QUADAS-2의 등급을 보고한 별도의 수치에서도 43개의 연구결과가 나왔다. 또한 43개의 연구를 보고한 연구 특성의 세부사항이 포함된 부록이 포함되었다. 그러나 연구들이 하나 이상의 중재결과를 보고할 수 있으므로, 연구별로 연구결과를 제시하지 않는 요약 숲 그림을 보고하였다. 모든 연구가 이 합성에 포함되었는지 여부는 확인할 수 없다.
4.2 사전에 정의된 모든 분석을 보고하였거나 이탈(departure)에 대해 설명하였는가?	PY	프로토콜은 부록 정보로 이용 가능했다. 어떠한 이탈도 명시되거나 관찰되지 않았다. 프로토콜은 비돌림위험 평가나 분석을 위한 전략에 대한 세부사항을 명시하지 않았지만, 논문의 방법 섹션은 각각 엄격한 방식으로 다루어졌다. 방법 섹션의 모든 분석은 결과에서 다루어졌다.
4.3 포함된 연구들의 연구질문, 연구설계, 연구결과와 특성과 유사성에 근거해보았을 때 합성이 적절하였는가?	Y	저자들은 "양수(amniotic fluid)의 수치의 랜덤 효과 메타 분석 모델을 사용하여 결과를 합성했다"(688페이지). 예측 능력을 평가하기 위해 2x2 표의 데이터를 사용하여 요약 민감도, 특이도 및 우도비를 계산하고 이변량 랜덤 효과 메타 분석 모델을 사용하여 합성했다.
4.4 연구간 변이(이질성)가 작거나 합성 중 다루어졌는가?	Y	통계적 이질성은 이 고찰에서 예상되었으며 랜덤 효과 모델을 사용하여 다루었다. 잠재적 효과 추정치를 조사하기 위해 메타 회귀 또는 하위군 분석을 수행했다. 12 >0%일 때도 예측 구간이 사용되었다.
4.5 연구결과가 견고하였는가? 예를 들어 깔대기 그림(funnel plot)이나 민감도 분석으로 연구결과와 견고성을 보여주었는가.	PN	저자들은 민감도 분석이 연구결과와 견고성(robustness)을 평가하는 데 사용되었는지 여부를 밝히지 않았다. 그러나 복합 중재결과(composite outcome)의 사용에 대한 논의가 있다.

8) Morris RK, Meller CH, Tamblin J, Malin GM, Riley RD, Kilby MD, et al. Association and prediction of amniotic fluid measurements for adverse pregnancy outcome: systematic review and metaanalysis. BJOG: An International Journal of Obstetrics & Gynaecology 2014;121(6):686-99.

<p>4.6 일차 연구에서 비뚤림위험이 미미했거나, 비뚤림위험이 자료 합성 과정에서 고려되었는가?</p>	<p>PY</p>	<p>비뚤림은 QUADAS-2 툴을 사용하여 평가되었다. 이 평가는 참고 표준에 대한 적절한 설명과 기준 표준에 대한 눈가림에서 비뚤림위험이 높다는 것을 나타냈다. QUADAS-2의 다른 측면은 크게 우려되지 않았다. “질”은 하위군 분석으로 명시되었지만 하위군 분석은 두 가지 주요 결과 중 하나에 대해서만 보고되었다. 양수과소증과 테아 웰빙 중재결과에 대해 저자들은 “하위군 분석에서 차이가 없었다”(P.691)고 말했다.</p>
<p>합성과 연구결과에 대한 우려</p>	<p>낮음</p>	<p>저자는 그들의 분석에서 이질성을 충분히 다루고 하위군 분석을 사용하여 탐구했다. 개별 연구의 비뚤림위험을 다루었고 저자들은 그것이 하위군 분석으로 포함되었다고 언급했다.</p>

5) 3단계: 비뚤림 판단

- 최종 단계에서는 문헌고찰이 전체적으로 비뚤림위험이 있는지 검토한다. 이 평가는 2 단계 영역과 동일한 구조를 사용하며, 여기에는 판단을 뒷받침하는 데 사용되는 신호 질문과 정보가 포함되지만, 비뚤림 우려에 대한 판단은 비뚤림위험의 전반적인 판단으로 대체된다.
- 이 단계의 첫 번째 신호 질문은 연구결과의 해석이 영역 1에서 4까지 확인한 모든 문제를 해결하는지 여부를 묻는다. 문제가 발견되지 않으면 “그렇다”로 대답할 수 있다. 이전 영역 중 하나 이상의 문제가 확인되었지만 결과를 해석하고 결론을 도출할 때 적절히 고려되었다면, 이 또한 “그렇다”로 평가될 수 있으며, 다른 신호 질문의 등급에 따라 문헌고찰은 여전히 “비뚤림위험 낮음”으로 평가될 수 있다.
- 이 단계에는 문헌고찰 연구결과의 해석과 관련된 추가적인 세 가지 신호 질문이 포함된다. 문헌고찰 결론이 제시된 근거에 의해 뒷받침되었는지 여부, 포함된 연구의 관련성을 고려했는지 여부, 문헌고찰 저자가 통계적 유의성에 근거하여 연구결과를 강조하지 않았는지 여부 등이다. 이들은 문헌고찰 연구결과에 비뚤림이 초래될 수 있는 해석의 중요한 측면이다.

가) 3단계에 대한 신호 질문 (각 질문에서 어떻게 등급을 매길지에 대한 지침 포함)

신호질문 a.	연구결과를 해석할 때 2단계에서 확인한 모든 우려 사항을 다루었는가?
----------------	---

설명

- 이 질문은 2단계에서 수행되고 3단계 시작 시 요약된 문헌고찰 과정에 대한 우려의 평가를 다시 언급한다.
- Y : 2단계에서 평가된 모든 영역이 “낮은 우려”로 평가된 경우
- 만약 하나라도 “높음” 또는 “불확실”로 평가되었다면, 평가자는 문헌고찰 저자가 연구결과의 해석에서 2단계 동안 확인된 우려를 적절히 다루었는지 여부를 고려해야 한다. 예를 들어 검색에 연구설계 필터를 적용했기 때문에 비뚤림위험이 높다고 판단했을 때, 이 문헌고찰이 이를 잠재적 한계로 인정하고 문헌고찰의 토론과 결론에서 문헌고찰 결과에 어떤 영향을 미칠 수 있었는지를 논의한 경우, 이 신호 질문은 “그렇다”로 평가될 수 있다.

신호질문 b.	문헌고찰의 연구질문에 대해 포함된 연구의 관련성이 적절히 고려되었는가?
----------------	--

설명

- 문헌고찰 결과를 해석할 때 중요한 측면은 확인된 (포함된) 연구와 문헌고찰의 연구질문에 대한 관련성(적용성/외부 타당성)을 고려하는 것이다. 문헌고찰에 포함된 연구가 연구질문에 직접 적용되지 않고 문헌고찰 연구결과를 해석할 때 이 점을 고려하지 않는 경우 비뚤림이 초래될 수 있다.
- 일부 문헌고찰에서는 공식적인 질 평가의 일부로 포함된 연구의 관련성을 고려할 수 있다. 이는 문헌고찰 질문에 대한 일차 연구의 적용 가능성에 대한 공식적인 평가와 일차 연구의 비뚤림위험 평가를 포함하는 QUADAS-2 도구를 사용한 진단 정확도 문헌고찰의 경우 특히 그렇다.
- 다른 문헌고찰은 토론 섹션에서 연구의 목적 적합성에 대한 논의와 함께 덜 공식적인 접근법을 사용할 수 있다. 문헌고찰의 결론이 포함된 연구의 관련성을 포함하여 근거를 적절하게 반영하는 한 어떤 방법이든 적절할 수 있다.

신호질문 c.	문헌고찰 저자는 통계적인 유의성을 보이는 결과만을 강조하는 것을 피하였는가?
----------------	---

설명

- 단일 문헌고찰에서 복수의 분석이 보고되는 경우 문헌고찰 저자는 모든 분석의 균형 잡힌 설명을 제시하는 것이 중요하다.
- 때때로 문헌고찰 저자는 통계적 유의성에 기초하여 연구결과를 강조하도록 선택할 수 있다. 예를 들어, 문헌고찰에서 10개의 중재결과를 평가하여 9개의 중재결과에 대한 효과가 없고 1개의 유의한 효과가 있는 경우, 해당 중재결과를 지나치게 강조하고 싶을 수 있다. 그러나 이것은 문헌고찰 연구결과와 실제 연구결과에 대한 오해의 소지가 있을 수 있으므로, 이렇게 해서는 안된다.

연구를 확인하고 선택하기 위해 사용된 방법의 비뚤림위험

낮음	이 문헌고찰의 연구결과는 믿을 만한 것 같다. 2단계에서 문헌고찰 과정에 대한 우려를 제기하지 않았거나 문헌고찰 결론에서 우려를 적절하게 고려하였다. 결론은 근거에 의해 뒷받침되었고 포함된 연구의 관련성에 대한 고려가 포함되었다.
높음	2단계 평가 중 제기된 우려 중 하나 이상이 문헌고찰 결론에서 다루지 않았거나, 문헌고찰 결론이 근거에 의해 뒷받침되지 않았거나, 결론이 문헌고찰 질문과 포함된 연구의 관련성을 고려하지 않았다.
불확실	보고된 정보가 불충분하기 때문에 비뚤림위험을 평가할 수 없다.

나) 평가 예시

a. 비플립위험 높음으로 판단된 문헌고찰⁹⁾

본문 기술 내용:

"This meta-analysis and systematic review shows that the pooled sensitivity and specificity of EUS for T0 tumor invasion is very high (around 97%)... Heterogeneity among different studies was evaluated not only with test of heterogeneity but also by drawing SROC curves and finding the AUC, since different studies might use slightly different criteria for staging. An AUC of 1 for any diagnostic test indicates that the test is excellent. SROC curves for EUS showed that the value of AUC was very close to 1, indicating that EUS is an excellent diagnostic test for T0 stage of rectal cancers...EUS has excellent sensitivity and specificity, which helps to accurately diagnose T0 stage of rectal cancers.."

2단계 평가 중 확인된 우려 사항 요약

영역	우려	우려에 대한 근거
1. 연구 적격 기준의 세부사항에 대한 우려	높음	적격기준이 명확하지 않고 애매모호하다. 적격 연구집단에 대한 세부사항은 보고되지 않았다.
2. 연구의 검색 그리고/또는 선택 방법에 대한 우려	높음	검색에서 관련 연구가 누락되었을 가능성이 높다. 세부 정보가 제한적으로 보고되었지만, 적절한 범위의 데이터베이스가 검색된 것으로 보이지 않고, 검색 전략에는 진단 연구설계 용어, 데이터베이스 검색에 추가되는 방법이 보고되지 않았으며, 1980년 이후로의 검색 제한은 정당화되지 않는다.
3. 자료수집과 연구평가 방법에 대한 우려	높음	공식적인 질 평가가 없다는 것은 포함된 연구에서 비플립의 위험이 불확실하다는 것을 의미한다. 독자가 결과를 해석할 수 있도록 하기에는 연구 세부 정보가 충분하지 않았다. 따라서 데이터 수집과 연구 평가 과정 모두에서 비플립위험이 높다.
4. 합성 및 연구결과에 대한 우려	높음	이 문헌고찰에서 이용한 분석 방법이 통계적으로 가장 견고하지는 않지만 신뢰할 수 있는 것으로 보인다. 이 문헌고찰의 주요 관심사는 이질성, 특히 하나의 특이한(outlying) 연구결과는 조사되지 않았다는 것이다. 매우 제한된 연구 세부사항이 제시되고 비플립위험 평가 수행되지 않았기 때문에 독자는 이 연구가 다른 연구와 다른지 여부를 평가할 수 없다. 만약 이 연구가 문헌고찰에 포함된 다른 연구들보다 더 신뢰할 수 있을 것 같으면, 문헌고찰의 전반적인 연구결과는 편향될 것이다.

9) Puli SR, Bechtold ML, Reddy JB, Choudhary A, Antillon MR. Can endoscopic ultrasound predict early rectal cancers that can be resected endoscopically? A meta-analysis and systematic review. Dig Dis Sci 2010;55(5):1221-9.

비독립위험 평가 예시 ('높음'으로 평가)

신호질문	평가	이유
a. 연구결과를 해석할 때 2단계에서 확인한 모든 우려 사항을 다루었는가?	N	2단계 평가에서 확인된 한계 중 어느 것도 문헌고찰 저자에 의해 한계점으로 확인되지 않았으며, 따라서 연구결과 해석에서 다루어지지 않았다.
b. 문헌고찰의 연구질문에 대해 포함된 연구의 관련성이 적절히 고려되었는가?	N	문헌고찰 질문에 포함된 연구의 관련성을 고려하지 않았으며, 특히 인구집단과 관련하여 독자가 이 평가를 하기에는 세부 사항이 불충분했다.
c. 평가자는 통계적인 유의성을 보이는 결과만을 강조하는 것을 피하였는가?	Y	토론에서 강조된 수치와 이로부터 얻어진 결론은 모두 1차 분석에서 도출한 요약 추정치들이다.
비독립위험	높음	2단계 평가에서 저자가 다루지 않은 문헌고찰 절차와 관련된 많은 우려가 있다. 여기에는 포함 기준의 명확성 부족, 연구 누락 가능성, 질 평가의 결여, 포함된 연구에 대한 불충분한 세부 사항, 합성에서 연구 간의 차이를 적절히 고려하지 않는 것 등이 포함된다. 또한 문헌고찰은 선택된 연구와 문헌고찰의 연구 질문의 관련성을 고려하지 않았다. 따라서 이 문헌고찰은 비독립위험이 높다.

b. 비플림위험 낮음으로 판단된 문헌고찰

본문 기술 내용:

Conclusions

“Current evidence suggests that oligohydramnios is strongly associated with being small for gestational age and mortality, and polyhydramnios with birthweight >90th centile. Despite strong associations with poor outcome, they do not accurately predict outcome risk for individuals”

Discussion

“Our quality assessment revealed concerns regarding possibility of bias through patient selection, performance of the index test and reference standard. We were unable to perform subgroup analysis for preterm versus term pregnancies and some studies reported insufficient data to determine whether thresholds for amniotic fluid measurement were adjusted for gestation. Where possible we used the results obtained closest to delivery and have performed subgroup analysis for those where the test was performed within 7 days of delivery. In particular, there was very poor reporting regarding the exact methods of the reference standards and whether there was any treatment used between the performance of the index and reference standard. A major concern therefore is in how many pregnancies was induction of labour performed due to the finding of oligohydramnios, which influences the results for pregnancy outcome, i.e. intervention bias. This bias can only truly be removed by performing an RCT, this would be impossible to perform as measurements of amniotic fluid volume have become the standard in fetal surveillance and management of high-risk pregnancies and so recruitment to such a trial would be very difficult. Finally, the outcome measures used in this review were those that were reported by the authors of the included studies, it is recognised that many of the outcome measures are subjective (e.g. admission to neonatal intensive care unit, need for resuscitation). The only real objective measure of poor fetal outcome is paired samples of cord pH and longer-term outcomes such as cerebral palsy, which were not reported.”

2단계 평가 중 확인된 우려 사항 요약

영역	우려	우려의 근거
1. 연구 적격 기준의 세부사항에 대한 우려	낮음	모든 신호 질문은 “그렇다” 또는 “아마도 그렇다”로 답했기 때문에 적격 기준에 대한 잠재적 우려는 파악되지 않았다. 문헌고찰 질문과 목표를 명확히 명시하였고, 적절하고 상세한 적격 기준을 사전에 구체화하고 정당화하기 위해 상당한 노력을 기울였다.
2. 연구의 검색 그리고/또는 선택 방법에 대한 우려	낮음	제목과 초록을 통해 선별하고, 전문을 평가하는 과정에서 복수의 검토자를 포함했다. 그러나 검토자가 독립적으로 행동했는지는 명시되지 않아 이 항목은 “아마도 그렇다”로 평가되었다.
3. 자료수집과 연구평가 방법에 대한 우려	낮음	모든 논문은 최소 2명의 검토자가 독립적으로 평가하고 적절한 자료를 독립적으로 추출했다. 연구 질문은 적절한 도구를 사용하여 공식적으로 평가되었다.
4. 합성 및 연구결과에 대한 우려	낮음	저자는 분석에서 이질성을 만족스럽게 다루고 하위 그룹 분석을 사용하여 탐색하였다. 개별 연구의 비뚫림위험을 다루었고 저자들은 그것을 하위 그룹 분석으로 포함하였다고 기술하였다.

비뚫림위험 평가 예시 (‘낮음’으로 평가)

신호 질문	평가	이유
a. 연구결과를 해석할 때 2단계에서 확인한 모든 우려 사항을 다루었는가?	Y	2단계 평가 중에 확인된 우려는 없다.
b. 문헌고찰의 연구질문에 대해 포함된 연구의 관련성이 적절히 고려되었는가?	PY	명시적으로는 아니지만, 개별 수준에서 문헌고찰 결과의 함의가 구체적으로 논의되었다. 포함된 연구에서 이용할 수 있는 인구집단과 중재결과의 측면에서 비뚫림의 잠재적 관련성에 대해 세부적으로 논의되었다.
c. 평가자는 통계적인 유의성을 보이는 결과만을 강조하는 것을 피하였는가?	Y	문헌고찰 결론은 통계적으로 유의하거나 유의하지 않은 문헌고찰의 연구결과를 모두 반영하였다.
비뚫림위험	낮음	2단계 평가에서는 문헌고찰 과정과 관련하여 어떠한 문제도 발견되지 않았다. 비뚫림위험 측면에서 문헌고찰에 포함된 연구의 잠재적 한계는 논의에서 자세히 논의되었다. 결론은 문헌고찰의 연구결과를 적절하게 반영하였다.

6) ROBIS 사용에 있어 현실적인 이슈

체계적 문헌고찰에서 비풀림위험을 판단하는 것이 항상 간단한 것은 아니다. ROBIS는 적용 가능성에 대한 평가를 통해 비풀림위험을 평가하기 위한 구조화된 접근방식을 제공하도록 설계되었다. ROBIS는 다른 배경을 가진 검토자가 사용할 수 있도록 설계되었으며, 일부 방법론적, 내용적 전문지식이 요구될 가능성이 높다. 이상적으로 ROBIS 평가는 독립적인 두 명의 검토자에 의해 수행될 것을 추천하지만 적어도 한 명의 검토자가 평가를 점검한 상태에서 다른 검토자에 의해 확인할 수 있다. ROBIS 평가를 위해 모든 신호 질문을 고려할 것을 권고한다. ROBIS는 다양한 주제를 다루는 문헌고찰에 적용하기 위해 일반적으로 개발되었으며, 따라서 모든 신호 질문은 모든 문헌고찰 유형에 관련되어야 한다. ROBIS는 의사결정을 내리는 것을 포함한다. 우리는 검토자들에게 판단을 뒷받침하는 정보, 신호 질문, 전반적인 우려 사항 판단에 대한 근거 등을 기록하도록 요구함으로써 가능한 한 이 과정을 투명하게 하는 것을 목표로 했다.

ROBIS 평가에서는 모든 영역에 대한 모든 신호 질문을 평가해야 한다. 그러나 특정 상황에서는 특정 등급이 달성되면 “정지”하는 것이 적절할 수 있다. 이것은 평가 목적에 따라 달라진다. 만약 ROBIS의 목적이 비풀림위험의 평가와 비풀림의 이유를 제공하는 것이라면 완전한 평가가 필요할 것이다. 그러나 전반적으로 비풀림위험이 높는지 또는 특정 영역에 대한 우려가 있는지를 단순히 확인하는 것이 목적이라면, 비풀림의 높은 우려/위험이 확인되면 중단하기로 선택할 수 있다. 예를 들어 적절한 범위의 데이터베이스(신호 질문 2.1)를 검색하지 않은 경우, 검토는 이 영역에 대한 높은 우려로 고려될 수 있다. 비풀림위험에 있는지 여부에만 관심이 있는 평가자는 단순히 편향의 “높은” 위험의 등급을 부여하면서 이 영역 또는 다른 영역의 나머지 신호 질문을 평가하지 않을 수 있다.

PART 2

RCT 비뚤림위험 평가 도구 : ROB 2

3장. Cochrane ROB 도구 2.0 (RoB 2)

가. 개요	76
나. 국문 ROB 2 평가 서식지	79
다. 국문 ROB 2	87
라. 국문 ROB 2 사용자 매뉴얼	104



3장. Cochrane ROB 도구 2.0 (RoB 2)

가. 개요

2008년 코크란 연합은 코크란 Risk of Bias(RoB) 도구를 발표했으며 2011년에 약간 수정되었다. RoB 도구는 몇 가지 원칙이 있었는데 그것은 1) 보고 질, 정밀도(연구결과에 무작위 오류가 없는 정도) 또는 외적타당성(직접성, 적용 가능성 또는 일반화 가능성)과 같은 개념과 거리를 둔 비뚤림위험의 정의, 2) 영역 기반(또는 컴포넌트) 접근 방식, 3) 이론적 고려와 경험적 근거의 결합을 바탕으로 둔 영역 설정, 4) 비뚤림의 '낮음', '높음' 또는 '불확실' 위험으로 평가된 각 영역에 대한 판단을 제공 등이었다. 하지만 이 도구는 일부 영역의 평가자 간 신뢰도가 낮았으며, '선택적 결과 보고' 영역에 대한 이론적 배경에 대한 재고가 필요했고, '기타 비뚤림위험 영역'의 오용, 분석과 결과 해석에서 비뚤림위험 평가의 적절한 고려 결여 같은 문제들이 제기되었다.

이러한 문제를 해결하기 위해 RoB 2라는 새로운 코크란 비뚤림위험 도구가 개발되었다. RoB 2의 주요 특징은 아래와 같다

- ① 비뚤림 평가는 연구나 증재결과(outcomes) 수준보다는 개별 연구결과(results)에 대해 시행한다.
- ② 비뚤림 영역의 이름을 보다 명확하게 기술하여, 다른 의미로 사용될 위험이나(예: '선택 비뚤림', '실행 비뚤림' 등), 익숙하지 않은 용어로 인해 발생하는 혼란을 줄인다.
- ③ 각 영역에 대한 비뚤림위험 판단에 도움이 되는 알고리즘을 제시하고, 신호 질문을 도입한다.
- ④ 증재 배정 효과와 증재 준수 효과를 구분하며, 이러한 개념은 '의도한 증재에서 이탈'로 인한 비뚤림위험 평가에 사용한다.
- ⑤ 분석에서 연구대상자를 제외함(예: naïve 'per protocol' analysis)으로써 발생하는 비뚤림위험에 대한 평가는 '증재결과 결측으로 인한 비뚤림'보다는 '의도한 증재에서 이탈로 인한 비뚤림'의 영역에서 하게 된다.
- ⑥ '연구결과 선택적 보고' 개념을 '연구결과 선택적 비보고' 개념과 구별하며, 후자의 개념은 도구에서 제거되어 합성 수준에서 다루어진다.
- ⑦ 새 영역을 추가할 수 있는 옵션이 제거되었다.
- ⑧ 전반적인 비뚤림위험을 평가하는 명시적 과정이 도입되었다.

이 매뉴얼에서는 ROB 2와 도구와 평가서식지 그리고 실제 사용례와 해당 내용을 기술한 매뉴얼을 제공한다. ROB 2의 원본 매뉴얼 번역본은 부록으로 제시한다.

#ROB 2 도구 전체 일러두기#

- 이 도구는 연구결과(results)에 비플림이 적용될 수 있는 5개의 영역으로 구성되어 있다.
 - (1) 무작위 배정 과정에서 발생하는 비플림
 - (2) 의도한 중재에서 이탈로 인한 비플림
 - (3) 중재결과 자료의 결측으로 인한 비플림
 - (4) 중재결과 측정의 비플림
 - (5) 보고된 연구결과 선택의 비플림
- 이 문서는 RoB 2 도구의 주요 특징을 설명하고 개별 무작위배정 평행군 임상시험에 적용할 수 있는 지침을 제공한다.
- 비플림의 각 영역 내에 신호 질문을 포함하는 것은 RoB 2의 핵심적인 부분이다. 신호 질문은 비플림 위험 평가와 관련된 정보를 도출하는 것을 목표로 한다. 신호 질문에 대한 응답 옵션은 다음과 같다.
 - (1) 그렇다(Y)
 - (2) 아마도 그렇다(PY)
 - (3) 아마도 아니다(PN)
 - (4) 아니다(N)
 - (5) 정보없음(NI)
- 신호 질문의 단순성과 명확성을 극대화하기 위해, 'Y'의 응답은 질문을 하는 가장 자연스러운 방법으로써 비플림위험이 낮거나 높음을 나타낼 수 있다.
- 'NI' 응답은 (i) 'PY' 또는 'PN'으로 응답하기에는 불충분한 내용이 보고된 경우, (ii) 해당 내용이 없고, 임상시험 상황에서 'PY' 또는 'PN'으로 응답하기는 것이 합리적이지 않은 경우에만 국한해야 한다.
- 이전 질문에 대한 응답이 필요한 경우에만 응답하는 신호 질문의 경우, 응답 옵션으로 '해당되지 않음(NA)'을 사용할 수 있다.
- 각 영역의 신호 질문에 대한 응답을 맵핑하는 알고리즘이 있으며 이에 따라 비플림위험 판단은 (1) 낮은 위험, (2) 일부 우려, (3) 높은 위험의 세 가지로 판단한다. '판단'이라는 한 이유는 알고리즘은 제안된 판단을 제공하지만, 사용자는 이를 검증하고 이것이 필요하다면 변경할 수도 있다는 뜻이다.

- 비뚤림위험에 대한 영역 수준 판단은 결과의 신뢰도에 대한 비뚤림이 미치는 영향에 대해 6개 영역은 각각 동일한 의미를 가져야 한다. 어떤 개별 영역에 대한 ‘높은’ 비뚤림위험으로 판단되면 결과에는 전반적으로 ‘높은’ 비뚤림위험을 가진 것으로 판단된다.
- 이 도구에는 각 영역과 전체 영역의 비뚤림 방향에 대한 판단이 있다. 일부 영역의 경우, 비뚤림이 무효점(null)을 향하는지, 무효점에서 벗어나는지를 쉽게 판단될 수 있다. 예를 들어, 연구대상자가 배정된 증재에서 다른 증재로 전환되는 비율이 높은 경우, 증재 준수 추정효과는 무효점을 향하는 쪽으로 비뚤림 방향이 발생할 수 있다.
- 전반적인 비뚤림위험 판단에 대한 반응 옵션은 개별 영역과 동일하다. 아래 표는 증재 결과에 대한 영역 내의 비뚤림위험 판단을 전체 영역에 걸쳐 판단하는 기본 접근방식을 보여준다.

전반적인 비뚤림위험 판단	기준
낮은 비뚤림위험	임상시험에서 현재 결과의 모든 영역에 대해 낮은 비뚤림위험으로 판단됨
일부 우려	임상시험에서 적어도 한 개의 영역에서 일부 우려가 있었지만 어떠한 영역도 높은 비뚤림위험이 없음
높은 비뚤림위험	임상시험에서 결과의 적어도 한 영역에서 높은 비뚤림위험을 보였거나, 다수의 영역에서 일부 우려사항이 있었고 이러한 우려는 결과의 신뢰도를 상당히 낮출 것으로 예상함

- 의도한 증재에서 이탈로 인한 비뚤림’ 영역에 대한 평가는 문헌고찰 저자가 관심을 기울이는 효과가 기저상태에서 이루어진 증재에 대한 증재 배정 효과(ITT 효과), 임상임상시험 프로토콜에 명시된 대로 증재가 이루어 졌는지에 대한 증재 준수 효과(PP효과) 중 어떤 것인가에 따라 이루어진다.
- 문헌고찰 저자는 관심 증재 효과를 정의하고, 이 효과에 대해 비뚤림위험 도구를 적절히 적용해야 한다. 증재 준수 효과를 평가할 때, 문헌고찰 저자는 (i) 비프로토콜 증재의 발생, (ii) 결과에 영향을 미칠 수 있는 증재 실행 실패, (iii) 임상시험 배정 증재에 대한 비준수 중 어떤 유형의 의도한 증재로부터 이탈을 조사할 것인지 미리 정해야 한다.
- 영역 중 하나 이상에 대해 ‘높은’ 비뚤림위험으로 판단될 경우, 해당 결과에 대한 전반적 비뚤림위험은 ‘높음’이기 때문에, 도구 사용자는 한 영역이 ‘높음’로 판단되면, 평가를 중지하고 싶을 수 있지만 이는 여러 이유로 추천되지 않는다.
- 비뚤림위험 평가를 완료하기 전에, 평가 대상의 설계, 평가할 증재 결과(평가되는 특정 연구결과뿐만 아니라), 관심 효과가 증재 배정 효과인지 증재 준수 효과인지와 같은 평가의 중요한 특성을 문서화하는 것이 필요하며, 이는 사전고려를 통해 수행하도록 한다.

나. 국문 ROB 2 평가 서식지

1) 사전고려사항

연구설계

- 개별 무작위배정 평행군 임상시험
- 군집 무작위배정 평행군 임상시험
- 개별 무작위배정 교차(혹은 기타 매칭방법을 적용한) 임상시험

평가 목적의 관점에서 실험 중재와 비교중재를 기술하십시오.

실험 중재 비교중재

비뚤림위험평가를 시행하는 중재결과를 구체적으로 기술하십시오.

평가할 수치형 연구결과를 구체적으로 기술하십시오.

분석이 여러번 시행된 경우 어떤 연구결과인지(예 : RR = 1.52 (95% CI 0.83 to 2.77)), 어디에서 얻었는지 명시할 것(표, 그림, 문장 등)

현 결과에 대한 평가 목적은 무엇인가?

- 중재 배정의 효과를 평가(중재 배정(ITT) 효과)
- 중재 준수의 효과를 평가(중재 준수(PP) 효과)

평가 목적이 중재 준수 효과라면, 다루어야 할 중재 이탈의 종류는? (최소한 한 개는 체크해야 함)

- 비프로토콜 중재의 발생
- 중재결과에 영향을 미칠 수 있는 중재 실행 실패
- 연구대상자가 배정된 중재 비준수

비뚤림위험 평가 정보는 어디에서 얻었나? (모두 체크하십시오)

- 학술지 논문
- 임상시험 프로토콜
- 통계분석 계획서(Statistical analysis plan, SAP)
- 비상업적 임상시험 레지스트리 자료(예: ClinicalTrials.gov 자료)
- 회사 소유 임상시험 레지스트리 자료(예: GSK Clinical Study Register record)
- 회색 문헌 (예: 출판되지 않은 학위 논문)
- 학회 초록
- 규제 문서(예: 임상연구 보고서, 약물승인 패키지)
- IRB 지원 서류
- 연구비 지원 데이터베이스 요약
(예: NIH RePORTER 혹은 Research Councils UK Gateway to Research)
- 임상시험 연구자와 사적 교신
- 스폰서와 사적 교신

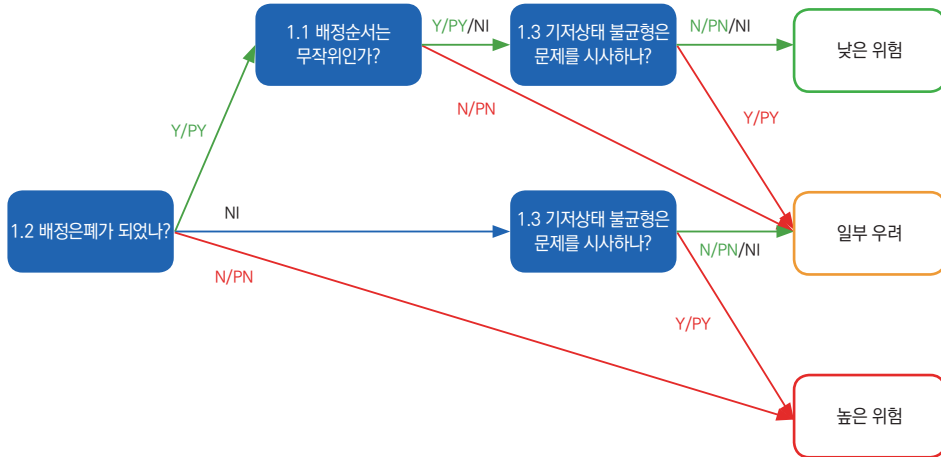
2) 평가도구

가) 영역1. 무작위과정에서 생기는 비뚤림

신호질문	응답 옵션
1.1 배정 순서는 무작위였나?	Y/PY/PN/N/NI
1.2 연구대상자가 모집되어 중재에 배정될 때까지 배정순서가 은폐되었나?	Y/PY/PN/N/NI
1.3 중재군간 기저상태 차이가 무작위 과정의 문제를 시사하나?	Y/PY/PN/N/NI

비뚤림위험 평가	낮음 /높음/일부 우려
선택사항: 무작위 배정 과정에서 발생하는 비뚤림의 예측 방향은 어떠한가?	적용불가/실험중재번호/비교중재번호/무효점을 함함/무효점에서 멀어짐/예측불가

무작위 과정에서 생기는 비뚤림위험 판단을 위한 알고리즘



나) 영역 2. 의도한 중재에서 이탈로 인한 비밀리 (중재 배정 효과)

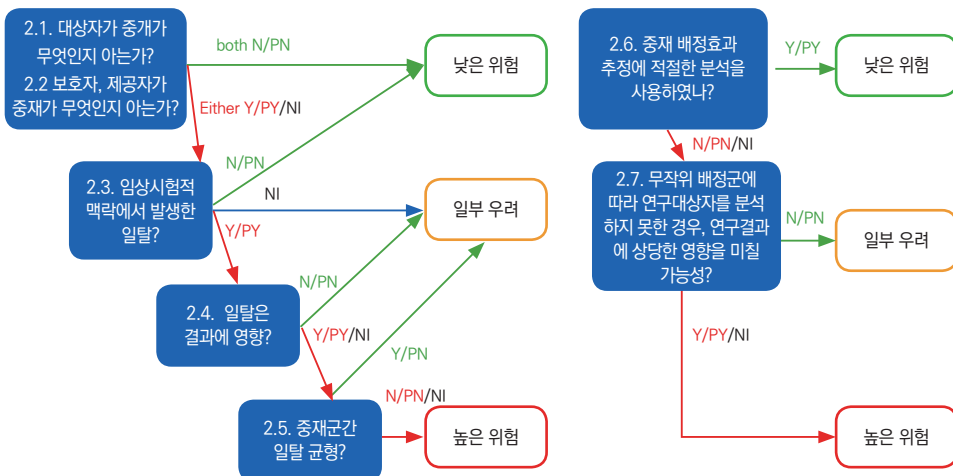
신호질문	응답 옵션
2.1 연구대상자는 임상시험 중에 배정된 중재에 대해 알고 있는가?	Y/PY/PN/N/NI
2.2 간병인과 중재 제공자는 임상시험 중에 연구대상자에 배정된 중재에 대해 알고 있는가?	Y/PY/PN/N/NI
2.3 2.1과 2.2의 Y/PY/NI 인 경우 : 임상시험적 맥락 때문에 의도한 중재에서 이탈이 있었나?	NA/Y/PY/PN/N/NI
2.4. 2.3의 Y/PY 인 경우 : 이러한 이탈로 중재결과에 영향을 미칠 것 같은가?	NA/Y/PY/PN/N/NI
2.5. 2.4의 Y/PY/NI 인 경우 : 의도한 중재에서 이탈이 두 군에서 유사한가?	NA/Y/PY/PN/N/NI
2.6. 중재 배정효과 추정에 적절한 분석을 사용하였나?	Y/PY/PN/N/NI
2.7. 2.6의 N/PN/NI 인 경우 : 무작위 배정군에 따라 연구대상자를 분석하지 못한 경우 (연구결과에) 상당한 영향을 미칠 가능성이 있었는가?	NA/Y/PY/PN/N/NI

비밀리위험 평가	낮음/높음/일부 우려
선택사항: 의도한 중재에서 이탈로 인한 비밀리의 예측 방향은 어떠한가?	적용불가/임상시험 중재선호/비교중재선호/무효점을 향함/무효점에서 멀어짐/예측불가

의도한 중재에서 이탈로 인한 비밀리 위험 판단 알고리즘 (중재 배정 효과) *

Part 1 : 2.1~2.5 문항

Part 2 : 2.6, 2.7 문항



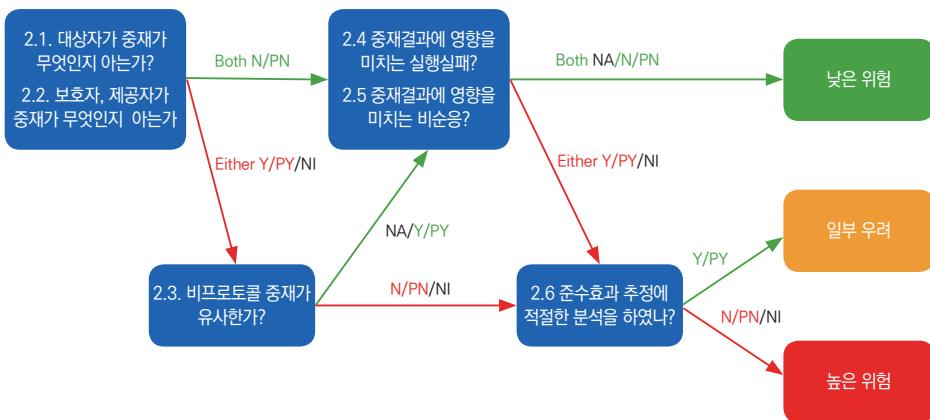


다) 영역 2. 의도한 중재에서 이탈로 인한 비틀림 (중재 준수 효과)

신호질문	응답 옵션
2.1 연구대상자는 임상시험 중에 배정된 중재에 대해 알고 있는가?	Y/PY/PN/N/NI
2.2 보호자와 중재제공자는 임상시험 중에 연구대상자에 대한 배정된 중재에 대해 알고 있는가?	Y/PY/PN/N/NI
2.3 [해당되는 경우] 2.1과 2.2에 대해 Y/PY/NI 인 경우 : 중재군간 중요한 비프로토콜 중재가 유사한가?	NA/Y/PY/PN/N/NI
2.4. [해당되는 경우] 중재결과에 영향을 미칠 수 있는 중재 실행의 실패가 있는가?	NA/Y/PY/PN/N/NI
2.5. [해당되는 경우] 연구대상자의 중재 결과에 영향을 미칠만한 중재에 대한 비준수가 있는가?	NA/Y/PY/PN/N/NI
2.6. 2.3에 대해 N/PN/NI 혹은 2.4나 2.5에 대해 Y/PY 인 경우 : 중재 준수 효과 추정에 적절한 분석법을 사용하였나?	NA/Y/PY/PN/N/NI

비틀림위험 평가	낮음/높음/일부 우려
선택사항: 의도한 중재에서 이탈로 인한 비틀림의 예측 방향은 어떠한가?	적용불가/임상시험 중재선호/비교중재선호/무효점을 향함/무효점에서 멀어짐/예측불가

의도한 중재에서 이탈로 인한 비틀림위험 판단 알고리즘(중재 준수 효과)

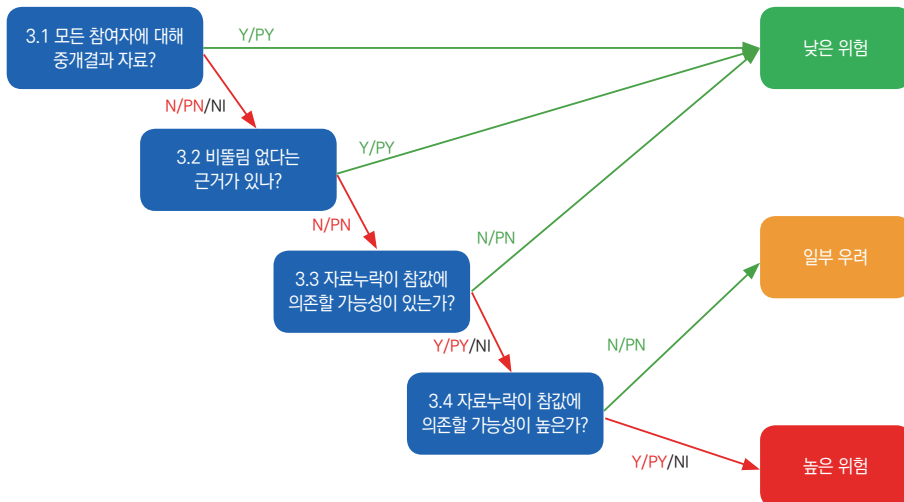


라) 영역 3. 중재결과 자료의 결측으로 인한 비뚤림

신호질문	응답 옵션
3.1 현재 중재결과에 대해 모든 혹은 거의 모든 연구대상자의 자료가 이용가능한가?	Y/PY/PN/N/NI
3.2 3.1에 대해 N/PN/NI인 경우: 결측치로 인해 연구결과의 비뚤림이 없다는 근거가 있는가?	NA/Y/PY/PN/N
3.3 3.2에 대해 N/PN인 경우: 중재 결과의 결측여부가 참값에 의존할 가능성이 있는가?	NA/Y/PY/PN/N/NI
3.4 3.3에 대해 Y/PY/NI: 중재 결과의 결측여부가 참값에 의존할 가능성이 높은가?	NA/Y/PY/PN/N/NI

비뚤림위험 평가	낮음/높음/일부 우려
선택사항: 중재결과 자료의 결측으로 인한 비뚤림의 예측 방향은 어떠한가?	적용불가/실험 중재번호/비교중재번호/무효점을 향함/무효점에서 멀어짐/예측불가

중재결과 자료의 결측으로 인한 비뚤림위험 판단 알고리즘

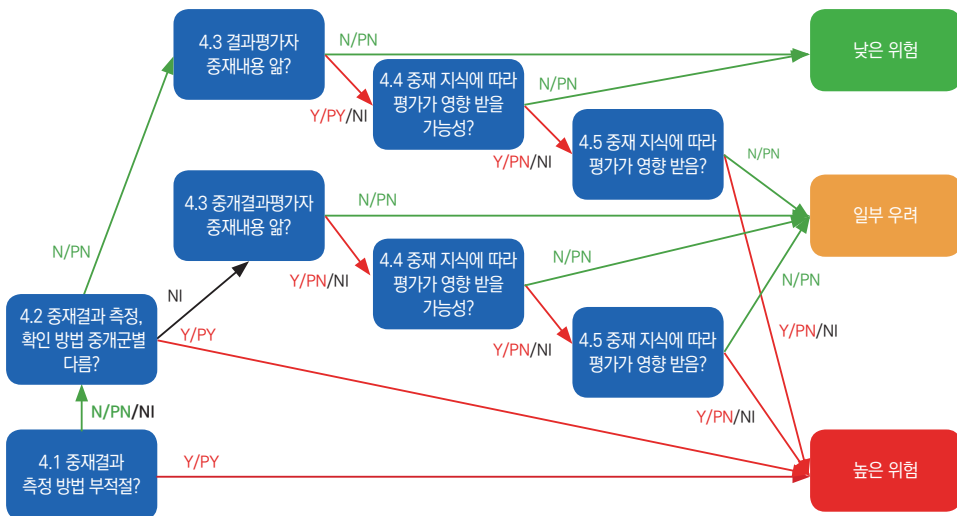


마) 영역 4. 중재결과 측정의 비돌림

신호질문	응답 옵션
4.1 중재결과 측정 방법이 부적절한가?	Y/PY/PN/N/NI
4.2 중재기간 중재결과 측정 혹은 확인 방법이 다른가?	Y/PY/PN/N/NI
4.3 4.1과 4.2에 대해 N/PN/NI 인 경우: 중재결과 평가자는 연구대상자가 받은 중재를 알고 있는가?	NA/Y/PY/PN/N/NI
4.4 4.3에 대해 Y/PY/NI 인 경우: 중재결과에 대한 평가가 중재에 대한 지식에 의해 영향을 받을 가능성이 있는가?	NA/Y/PY/PN/N/NI
4.5 4.4에 대해 Y/PY/NI 인 경우: 중재결과에 대한 평가가 중재에 대한 지식에 영향을 받았을 가능성이 높은가?	NA/Y/PY/PN/N/NI

비돌림위험 평가	낮음/높음/일부 우려
선택사항: 중재결과 측정의 비돌림의 예측 방향은 어떠한가?	적용불가/실험 중재번호/비교중재번호/무효점을 향함/무효점에서 멀어짐/예측불가

결과자료 측정의 비돌림위험 판단 알고리즘

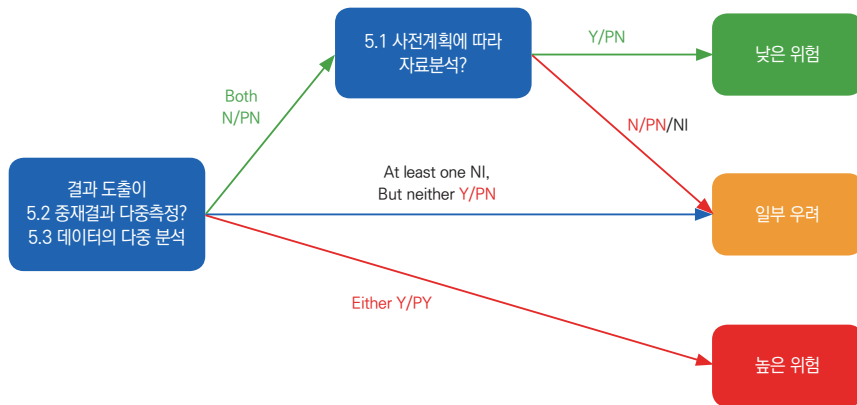


바) 영역 5. 보고된 연구결과 선택의 비뚤림

신호질문	응답 옵션
5.1 이 연구결과를 도출한 자료는 분석할 수 있도록 미리 지정된 분석 계획에 따라, 중재 결과 자료 눈가림이 해제되기 전에 분석되었는가?	Y/PY/PN/N/NI
평가되고 있는 수치형 연구결과는 다음과 같은 연구결과에 기초하여 선택되었을 가능성이 있는가?	
5.2 중재 결과 영역 내에서 여러 개의 적합한 중재 연구결과 측정(예: 척도, 정의, 시점)?	Y/PY/PN/N/NI
5.3 데이터의 다중 적격 분석?	Y/PY/PN/N/NI

비뚤림위험 평가	낮음/높음/일부 우려
선택사항: 보고된 연구결과 선택의 비뚤림의 예측 방향은 어떠한가?	적용불가/실험 중재번호/비교중재번호/무효점을 향함/무효점에서 멀어짐/예측불가

보고된 연구결과 선택의 비뚤림위험 판단 알고리즘



사) 전반적인 비뚤림위험 평가

비뚤림위험 판단	낮은 위험/ 높은 위험/ 일부우려
선택사항: 중재결과에 대한 전반적인 예측 비뚤림 방향은 어떠한가?	적용불가/실험중재번호/비교중재번호/무효점을 향함/무효점에서 멀어짐/예측불가

다. 국문 ROB 2

사전 고려 사항

연구 설계

- 개별 무작위배정 평행군 임상시험
- 군집 무작위배정 평행군 임상시험
- 개별 무작위배정 교차(혹은 기타 매칭방법을 적용한) 임상시험

평가 목적의 관점에서 실험 중재와 비교중재를 기술하십시오

실험 중재 비교중재

비뚤림위험평가를 시행하는 중재결과를 구체적으로 기술하십시오.

평가할 수치형 연구결과를 구체적으로 기술하십시오

분석이 여러번 시행된 경우 어떤 연구결과인지(예 : RR = 1.52 (95% CI 0.83 to 2.77)), 어디에서 얻었는지 명시할 것(표, 그림, 문장 등)

현 결과에 대한 평가 목적은 무엇인가?

- 중재 배정의 효과를 평가(중재 배정, ITT 효과)
- 중재 준수의 효과를 평가(프로토콜 준수, PP 효과)

평가 목적이 중재 준수 효과라면, 다루어야 할 중재 이탈의 종류는 무엇인가(최소한 한 개는 체크해야 함)

- 비프로토콜 중재의 발생
- 중재결과에 영향을 미칠 수 있는 중재 실행 실패
- 연구 대상자가 배정된 중재를 준수하지 않음

비뚤림위험 평가 정보는 어디에서 얻었나? (모두 체크하십시오)

- 학술지 논문
- 임상시험 프로토콜
- 통계분석 계획서(Statistical analysis plan, SAP)
- 비상업적 임상시험 레지스트리 자료(예: ClinicalTrials.gov 자료)
- 회사 소유 임상시험 레지스트리 자료(예: GSK Clinical Study Register record)
- 회색 문헌 (예: 출판되지 않은 학위 논문)
- 학회 초록
- 규제 문서(예: 임상연구 보고서, 약물승인 패키지)
- IRB 지원 서류
- 연구비 지원 데이터베이스 요약(예: NIH RePORTER 혹은 Research Councils UK Gateway to Research)
- 임상시험 연구자와 사적 교신
- 스폰서와 사적 교신

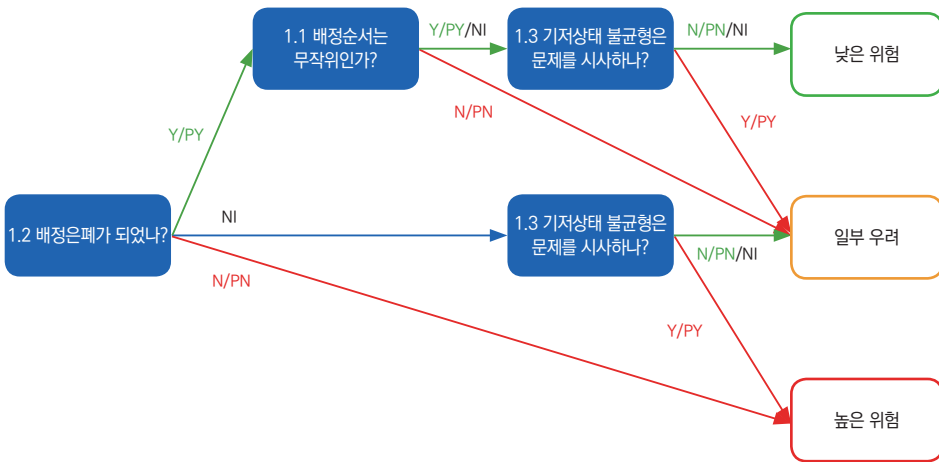
영역1. 무작위과정에서 생기는 비뚤림위험

신호질문	세부 설명	응답 옵션
<p>1.1 배정 순서는 무작위였나?</p>	<ul style="list-style-type: none"> • 순서 생성 과정에서 무작위 구성요소가 사용된 경우 '그렇다(Y)'에 답한다. 예를 들어 컴퓨터 난수 생성, 난수표 사용, 동전 던지기, 카드나 봉투 뽑기, 주사위 던지기, 추첨 등이 해당한다. 최소화는 일반적으로 무작위 요소(적어도 점수가 같을 때)로 구현되므로 최소화를 사용하여 생성되는 배정 순서는 무작위로 간주해야 한다. • 배정 순서를 생성하는 데 무작위 요소를 사용하지 않았거나 순서가 예측 가능한 경우 '아니다(N)'로 응답한다. 예를 들어 교대, 날짜(생년월일 또는 입원일)에 기초한 방법, 환자 의무기록 번호, 임상 의사 또는 연구 대상자에 의한 배정 결정, 중재의 가용성에 기초한 배정 또는 그 밖의 체계적 또는 우발적 방법 등이 해당한다. • 무작위 배정 방법에 대한 유일한 정보가 "연구가 무작위로 진행되었다"라는 문구일 경우 '정보 없음(NI)'에 해당한다. • 어떤 상황에서는 '아마도 아니다(PN)' 또는 '아마도 그렇다(PY)'가 적절한 응답일 수 있다. 예를 들어, 숙련된 임상시험팀에서 시행한 대규모 임상시험 논문이 논문 분량 제한이 엄격하게 적용되는 학술지에 게재되었다면, 무작위 순서 생성에 대한 특정 정보가 없는 경우 '정보 없음'이 아니라 '아마도 그렇다'라고 응답할 수 있다. 또한 유사한 시기에 동일한 연구팀에서 수행한 다른 임상시험이 비무작위 연구로 진행되었다고 되어 있으면 현재의 연구는 유사한 방법을 사용하여 수행되었다고 가정하는 것이 합리적일 수 있다. 	<p>Y/PY/ PN/N/NI</p>
<p>1.2 연구 대상자가 모집되어 중재에 배정될 때까지 배정순서가 은폐되었나?</p>	<ul style="list-style-type: none"> • 임상시험에서 연구대상자들에게 중재를 배정하기 위해 어떤 형태의 원격 또는 중앙 관리 방법을 사용한 경우, 배정 과정은 모집 담당자와 무관하게 외부 조직에서 통제하는 경우(예: 독립적 중앙 약국, 전화 또는 인터넷 기반 무작위화 서비스 제공자) '그렇다'라고 응답한다. 봉투나 약용기가 적절하게 사용된 경우 '그렇다'라고 대답한다. 봉투는 불투명하고 순차적으로 번호가 부여되어야 하며 변조 방지 봉인으로 밀봉되어야 한다. 연구대상자에게 봉투가 비가역적으로 배정된 후에만 개봉되어야 한다. 약물 용기는 순차적으로 번호가 매겨져야 하며 동일한 외관을 가지며, 연구대상자에게 비가역적으로 배정된 후에만 폐기 또는 투여해야 한다. 이 수준의 세부사항은 보고서에 거의 제공되지 않으며, 'PY' 또는 'PN'의 대답을 정당화하기 위한 판단이 필요할 수 있다. 모집담당자 또는 연구대상자가 향후 배정에 대한 지식을 가지고 있다고 의심할 만한 이유가 있는 경우 '아니다'에 응답한다. 	<p>Y/PY/ PN/N/NI</p>

신호질문	세부 설명	응답 옵션
<p>1.3 중재군간 기저상태 차이가 무작위 과정의 문제를 시사하나?</p>	<ul style="list-style-type: none"> • 우연에 의한 차이로 간주되는 차이는 비돌림위험으로 이어지지 않는다는 점에 유의한다. 기존 0.05 임계값에서 '통계적으로 유의함'으로 확인된 적은 차이는 우연에 의한 차이로 간주해야 한다. • 불균형이 뚜렷하지 않거나 관찰된 불균형이 우연에 의한 차이로 간주해야 하는 경우 '아니다'에 답한다. 무작위화 과정의 문제를 시사하는 다음과 같은 불균형이 있는 경우 '그렇다'에 답한다. <ul style="list-style-type: none"> (1) 의도된 배정 비율과 비교하여 중재군 크기 간의 상당한 차이 혹은 (2) 중재군 간 기저상태 특성의 통계적으로 유의한 차이에 대한 우연이라고 할 수 없는 상당한 초과 혹은 (3). 하나 이상의 핵심 예측 요인 또는 예후 변수에 대한 기저 상태 불균형이 우연으로 인해 발생할 가능성이 매우 낮고, 그룹 간 차이가 중재 효과 추정치의 비돌림을 초래할 정도로 큰 경우. 또한 무작위화 과정에 문제가 있었다고 의심하는 다른 이유가 있는 경우 '그렇다'에 답하십시오 : (4) 우연이라고 보기 힘들 정도로 기저상태 특성이 과도하게 유사한 경우 <p>이용 가능한 유용한 기준 정보가 없을 때(예: 최종 분석 참여자의 기준 특성만 보고한 연구) '정보 없음'으로 답한다.</p> <p>이 질문에 대한 응답이 1.1 또는 1.2 질문에 대한 응답에 영향을 미치지 않아야 한다. 예를 들어, 임상시험 결과에 있어서 기저상태 불균형을 가지고 있지만 저자들이 적절한 무작위화 방법을 보고하는 경우, 질문 1.1과 1.2에 대한 응답은 보고된 적절한 방법에 근거하여 답변해야 하며, 불균형에 대한 우려는 질문 1.3에 대한 답변에서 제기되어야 하며, 영역 수준 비돌림위험 판단에 반영되어야 한다.</p> <p>임상시험 시행자는 예측 인자의 불균형을 가져 상태에서 제어함으로써 결함 있는 무작위화를 처리하려는 분석을 수행할 수 있다. 무작위화 과정의 문제로 인한 비돌림위험을 제거하기 위해서는 기저상태에서 불균형이 있는 모든 예측 요인을 알고 측정할 필요가 있을 것이다. 모든 중요한 예측 요인을 알고 측정할 가능성은 낮으므로 그러한 분석은 기껏해야 비돌림의 위험을 줄이는 정도이다. 문헌고찰 저자가 무작위화 실패를 완화하기 위해 기저상태 불균형을 보정하는 임상 시험의 비돌림위험을 평가하고자 하는 경우, ROBINS-I 도구를 사용하여 평가해야 한다.</p>	<p>Y/PY/ PN/N/NI</p>

비뚤림위험 평가	알고리즘 참조	낮음/높음/일부 우려
선택사항: 무작위 배정 과정에서 발생하는 비뚤림의 예측 방향은 어떠한가?	비뚤림의 방향을 예측할 수 있으면, 이를 기술하는 것이 도움이 된다. 방향은 무효점(혹은 무효점의 반대쪽)을 향한, 혹은 실험 중재번호, 비교중재 번호로 기술한다.	적용불가 / 실험 중재번호 / 비교중재번호/무효점을 향한/무효점에서 멀어짐/예측불가

무작위 과정에서 생기는 비뚤림위험 판단 알고리즘



영역 2. 의도한 중재에서 이탈로 인한 비뚤림위험 (중재 배정 효과)

신호질문	세부 설명	응답 옵션
2.1 연구 대상자는 임상시험 도중 배정된 중재에 대해 알고 있는가?	<ul style="list-style-type: none"> 연구대상자가 배정된 중재를 알고 있다면, 건강 관련 행동이 중재군에 따라 다를 가능성이 있다. 연구대상자에 대한 눈가림은 위약이나 가짜 중재를 통해, 그러한 차이를 방지할 수 있다. 연구대상자가 중재 중 하나에 특정한 것으로 알려진 부작용이나 독성을 경험한 경우, 'Y' 또는 'PY'라고 응답한다. 	<p>Y/PY/ PN/N/NI</p>
2.2 중재 제공자는 임상시험 도중 연구대상자에 배정된 중재에 대해 알고 있는가?	<ul style="list-style-type: none"> 중재 제공자가 배정된 중재를 알고 있다면, 그 중재를 시행했는지 비프로토콜 중재가 발생했는지 여부가 중재군 간에 다를 수 있다. 눈가림은 그러한 차이를 막을 수 있다. 중재제공자가 중재 중 하나에 특정한 것으로 알고 있는 부작용이나 독성을 연구 대상자들이 경험한 경우, 'Y' 또는 'PY'라고 대답한다. 무작위 배정 은폐가 되지 않는 경우, 중재제공자는 임상시험 중에 연구대상자들에 배정된 중재가 무엇인지 알고 있었을 가능성이 높다. 	<p>Y/PY/ PN/N/NI</p>

신호질문	세부 설명	응답 옵션
<p>2.3 2.1과 2.2이 Y/PY//NI 인 경우 : 임상시험적 맥락 때문에 의도한 중재에서 이탈이 있었나?</p>	<ul style="list-style-type: none"> • 중재 배정 효과의 경우, 시험 프로토콜과 일치하지 않는 배정 중재의 변경이 임상시험적인 맥락 때문에 발생하여 결과에 영향을 미칠 가능성이 있는 경우에만 이 영역의 비독립이 발생한다. 예를 들어, 동의서를 받는 과정에서 대조 중재에 배정되면 불운한 것이라고 느끼게 되어, 실험 중재나 더 나은 예후를 가지는 중재에 배정되도록 노력할 수 있다. • 임상시험 맥락으로 인해 임상시험이 진행되지 않은 상황과 비교했을 때, 프로토콜에 명시된 것 이외의 추가 중재로 이어졌거나, 의도한 대로 중재를 수행하지 못했거나, 연구대상자가 지정된 중재에 대한 비준수로 이어졌다는 증거가 있는 경우, 'Y' 또는 'PY'로 응답한다. • 시험 프로토콜과 일치하는 중재 변경에 대해서는 'N' 또는 'PN'으로 대답한다. 예를 들어 급성 독성 때문에 약물 중재가 중단되거나 의도된 중재 중에 발생한 중재 결과를 처리하는 목적으로 추가 중재를 사용하는 것과 같은 경우이다. 임상시험 맥락과 관련이 없는 경우, 중재에 대한 비준수에 대해 'N' 또는 'PN'로 응답한다. • 연구대상자가 중재 중 하나에 특이한 부작용이나 독성이 생겨서 눈가림이 깨진 경우, 시험 프로토콜과 일관되지 않고 임상시험적인 맥락 때문에 발생한 중재 변경인 경우에만 'Y' 또는 'PY'라고 응답한다. • '정보 없음'이라는 응답은 임상시험 연구자가 임상시험적인 맥락 때문에 이탈이 발생했는지 여부를 항상 보고하지 않는 경우에 부여한다. 그러나 그러한 이탈이 발생했을 가능성이 있는 경우 응답은 'PY'여야 한다. 	<p>NA/Y/PY/ PN/N/NI</p>
<p>2.4. 2.3이 Y/PY 인 경우 : 이러한 이탈로 중재결과에 영향을 미칠 것 같은가?</p>	<ul style="list-style-type: none"> • 임상시험 프로토콜과 일관되지 않고 임상시험 맥락 때문에 발생한 배정된 중재의 변경은 중재 결과에 영향을 미치는 경우에는 중요할 것이고 그렇지 않은 경우에는 중요하지 않을 것이다. 	<p>NA/Y/PY/ PN/N/NI</p>
<p>2.5. 2.4이 Y/PY/NI 인 경우 : 의도한 중재에서 이탈이 두 군에서 유사한가?</p>	<ul style="list-style-type: none"> • 시험 프로토콜과 일관되지 않고 임상시험 맥락 때문에 발생한 배정된 중재의 변경은 중재군 간에 균형이 맞지 않는 경우 비독립을 초래할 가능성이 더 크다. 	<p>NA/Y/PY/ PN/N/NI</p>

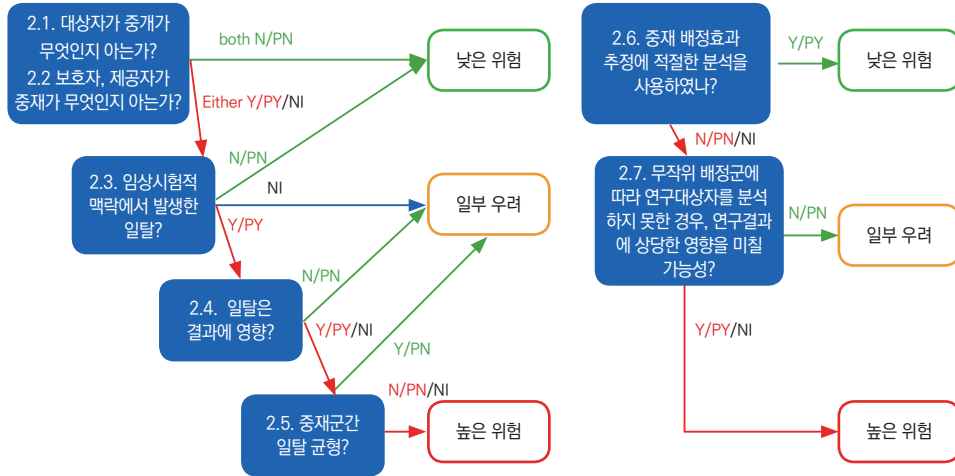
신호질문	세부 설명	응답 옵션
2.6. 중재 배정효과 추정에 적절한 분석을 사용하였나?	<ul style="list-style-type: none"> 중재결과 자료가 결측된 연구대상자를 제외한 배정된 대로 분석(ITT) 과 수정된 배정된 대로 분석(mITT) 모두 적절한 것으로 간주해야 한다. naïve 'per-protocol' 분석 (배정된 중재를 받지 않은 연구대상자는 제외)과 치료한대로 분석(배정된 중재가 아닌, 받은 중재에 따라 연구대상자가 그룹화되는 경우) 모두 부적절한 것으로 간주해야 한다. 적격 연구대상자를 제외한 분석 역시 부적절하다고 보아야 하지만, 부적격 연구대상자의 무작위 배정 후 배제(무작위 배정 이후까지 자격이 확인되지 않았고, 중재군 배정에 의해 영향을 받을 수 없었을 때)는 적절한 것으로 간주될 수 있다. 	<p style="text-align: center;">Y/PY/ PN/N/NI</p>
2.7. 2.6이 N/PN/NI 인 경우 : 무작위 배정군에 따라 연구대상자를 분석하지 못한 경우 (연구 결과에) 상당한 영향을 미칠 가능성이 있었는가?	<ul style="list-style-type: none"> 이 질문은 잘못된 중재군에서 분석되었거나 분석에서 제외된 연구대상자의 수가 결과에 상당한 영향을 미칠 수 있을 정도로 충분하지 여부를 다룬다. 정확한 규칙을 지정할 수 없다. 연구대상자의 5% 미만이 잘못된 그룹에서 분석되거나 배제되었더라도, 중재 결과가 드물거나 예측 요인과 강하게 관련되는 경우라면 상당한 영향이 있을 수 있다. 	<p style="text-align: center;">NA/Y/PY/ PN/N/NI</p>

비뚤림위험 평가	알고리즘 참조	낮음/높음/일부 우려
선택사항: 무작위 배정 과정에서 발생하는 비뚤림의 예측 방향은 어떠한가?	비뚤림의 방향을 예측할 수 있으면, 이를 기술하는 것이 도움이 된다. 방향은 무효점(혹은 무효점의 반대쪽)을 향함, 혹은 실험 중재번호, 비교중재 번호로 기술한다.	적용불가 / 실험 중재번호 / 비교중재번호/무효점을 향함/무효점에서 멀어짐/예측불가

의도한 중재에서 이탈로 인한 비밀리위험 판단 알고리즘(중재 배정 효과)

Part 1 : 문항 2.1에서 2.5까지

Part 2 : 문항 2.6, 2.7



PART 2

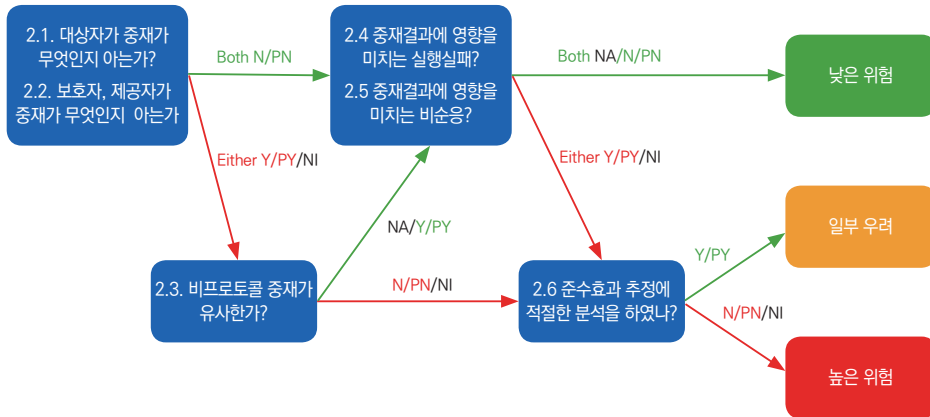
영역 2. 의도한 중재에서 이탈로 인한 비뚤림위험 (중재 준수 효과)

신호질문	세부 설명	응답 옵션
2.1 연구 대상자는 임상시험 도중 배정된 중재에 대해 알고 있는가?	<ul style="list-style-type: none"> 연구대상자가 배정된 중재를 알고 있다면, 건강 관련 행동이 중재군에 따라 다를 가능성이 크다. 연구대상자에 대한 눈가림은 위약이나 가짜 중재를 통해, 그러한 차이를 방지할 수 있다. 연구대상자가 중재 중 하나에 특정한 것으로 알려진 부작용이나 독성을 경험한 경우, 'Y' 또는 'PY'라고 응답한다. 	<p>Y/PY/ PN/N/NI</p>
2.2 중재 제공자는 임상시험 도중 연구 대상자에 대한 배정된 중재에 대해 알고 있는가?	<ul style="list-style-type: none"> 중재 제공자가 배정된 중재를 알고 있다면, 그 중재를 시행했는지 비프로토콜 중재를 시행했는지 여부가 중재군 간에 다를 수 있다. 눈가림은 그러한 차이를 막을 수 있다. 중재제공자가 중재 중 하나에 특정한 것으로 알고 있는 부작용이나 독성을 연구 대상자들이 경험한 경우, 'Y' 또는 'PY'라고 응답한다. 무작위 배정 은폐가 되지 않는 경우, 중재제공자는 임상시험 중에 연구대상자들에 배정된 중재가 무엇인지 알고 있었을 가능성이 크다. 	<p>Y/PY/ PN/N/NI</p>
2.3 [해당되는 경우] 2.1과 2.2에 대해 Y/PY//NI 인 경우 : 중재군 간 중요한 비프로토콜 중재가 유사한가?	<ul style="list-style-type: none"> 이 질문은 사전 고려사항에서 평가 과정에 중재군 간의 중요한 비프로토콜 개입의 불균형을 다룰 것이라고 표시한 경우에만 하도록 한다. 중요한 비프로토콜 중재는 (1) 시험 프로토콜에 부합하지 않는 추가 중재 또는 노출, (2) 연구대상자가 배정된 중재와 같이 혹은 이후에 받을 수 있는 추가 중재 또는 노출, (3) 중재결과에 대해 예측요인인 추가 중재 또는 노출이다. 중재군간에 중재에 불균형이 있을 경우 비뚤림위험이 더 높아질 것이다. 	<p>NA/Y/PY/ PN/N/NI</p>
2.4. [해당되는 경우] 중재결과에 영향을 미칠 수 있는 중재 실행의 실패가 있는가?	<ul style="list-style-type: none"> 이 질문은 사전 고려사항에서 평가과정에 중재 결과에 영향을 미칠 수 있는 중재의 실행 실패'를 다룰 것이라고 표시한 경우에만 하도록 한다. 예를 들어, 치료를 제공하는 보건의료인에 의해 의도한 대로 중재가 실행되지 않은 경우 비뚤림위험이 더 높아질 것이다. 대부분의 연구대상자가 중재 실행에 성공했다면 'N' 또는 'PN'라고 응답한다. 	<p>NA/Y/PY/ PN/N/NI</p>
2.5. [해당되는 경우] 연구 대상자의 중재 결과에 영향을 미칠만한 중재에 대한 비준수가 있는가?	<ul style="list-style-type: none"> 이 질문은 사전 고려사항에서 평가 과정에 '연구대상자의 결과에 영향을 미칠 수 있는 비준수를 다룰 것이라고 표시한 경우에만 하도록 한다. 비준수에는 지속적인 중재에 대한 불안정한 이행, 중재 중단, 대조군 중재로 교차 이동, 다른 활성 중재로 전환 등이 포함된다. 추적관찰 기간 동안 배정된 중재를 계속 받고 있는 연구 대상자 비율에 대해 이용할 수 있는 정보를 고려하고, 비준수 비율이 문제를 제기할 정도로 높으면 'Y' 또는 'PY'라고 응답한다. 불완전한 준수가 불가능한 1회 투여 중재 연구나 모든 혹은 대부분의 연구 대상자가 배정된 중재를 받은 경우 'N'라고 응답한다. 	<p>NA/Y/PY/ PN/N/NI</p>

신호질문	세부 설명	응답 옵션
<p>2.6. 2.3에 대해 N/PN/NI 혹은 2.4나 2.5에 대해 Y/PY 인 경우 : 중재 준수 효과 추정에 적절한 분석법을 사용하였나?</p>	<ul style="list-style-type: none"> • naïve 'per-protocol' 분석 (배정된 중재를 받지 않은 연구대상자는 제외)과 치료한대로 분석(배정된 중재가 아닌, 받은 중재에 따라 연구대상자가 그룹화되는 경우) 모두 중재 준수 효과를 추정하는 데 적합하지 않다. 그러나 무작위 임상 시험 자료를 사용하여 중재에 따르는 효과에 대한 편향 없는 추정치를 도출할 수 있다. 적절한 방법의 예로는 (1) 기저 상태에서 단일 중재로 시행되고 준수가 전무 혹은 모두인 경우 표준 치료와 비교되는 임상 시험에서 중재의 영향을 추정하기 위한 도구 변수분석(instrumental variable analyses), (2) 지속적인 치료 전략에 대한 임상시험에서 중도절단된 연구 대상자에 대해 보정해 주는 역확률가중치(inverse probability weighting)가 그것이다. 이러한 방법은 강력한 가정에 의존할 수밖에 없으므로 응답을 'Y' 혹은 'PY'로 하려면 방법 자체가 적절해야 하며 정당화되어야 한다. 논문에서 이러한 방법을 사용하였다고 보고하지만 의도한 중재부터 이탈된 사람에 대한 정보가 없는 경우가 있다. 이런 경우 분석이 적절했다고 판단하기는 매우 어렵다. • 중요한 비프로토콜 중재가 한 중재군의 모든 연구 대상자에게 시행된 경우, 보정으로 이를 극복할 수 없다. • 중재 준수 효과를 추정하는 데 적절하지 않을 분석 전략의 몇 가지 예는 (i) 배정된 대로 (ITT) 분석, (ii) 시행한대로 분석(PP), (iii) 치료한 대로 분석 (iv) 받은 치료에 의한 분석('analysis by treatment received') 등이다. 	<p>NA/Y/PY/ PN/N/NI</p>

비뚫림위험 평가	알고리즘 참조	낮음/높음/일부 우려
<p>선택사항: 무작위 배정 과정에서 발생하는 비뚫림의 예측 방향은 어떠한가?</p>	<p>비뚫림의 방향을 예측할 수 있으면, 이를 기술하는 것이 도움이 된다. 방향은 무효점(혹은 무효점의 반대쪽)을 향함, 혹은 실험 중재선호, 비교중재 선호로 기술한다.</p>	<p>적용불가 / 실험 중재선호 / 비교중재선호/무효점을 향함/무효점에서 멀어짐/예측불가</p>

의도한 중재에서 이탈로 인한 비뚤림위험 판단 알고리즘(중재 준수 효과)



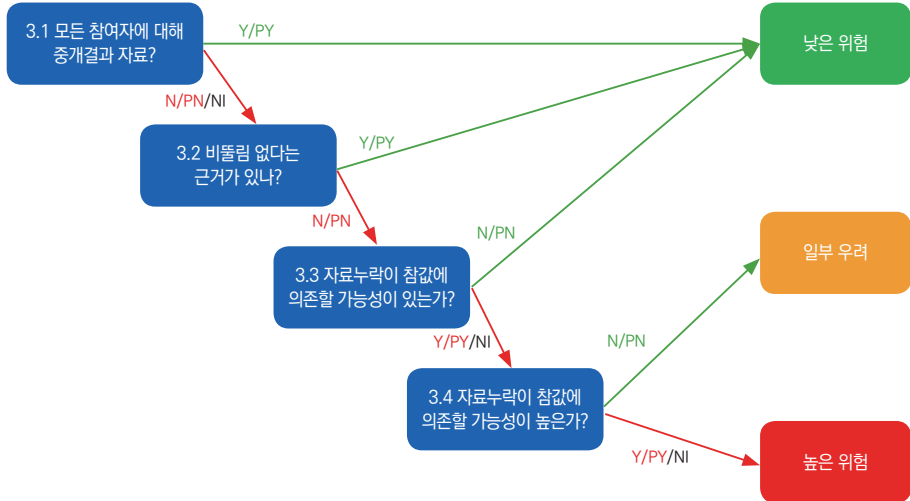
영역 3 중재결과 자료의 결측으로 인한 비뚤림위험

신호질문	세부 설명	응답 옵션
3.1 현재 중재결과에 대해 모든 혹은 거의 모든 연구 대상자의 자료가 이용가능한가?	<ul style="list-style-type: none"> • 중재배정 효과 분석(ITT effect analysis)을 위해서는 무작위 임상시험 연구 대상자 전부가 필요하다. • "거의 모두"는 결측치가 있는 연구대상자의 수가 충분히 적기 때문에 결과가 어떨든 간에 중재의 추정 효과에 중요한 차이가 없는 것으로 해석해야 한다. • 연속변수 중재결과의 경우, 연구대상자의 95%에서 자료를 이용할 수 있으면 충분하다고 할 수 있을 것이다. 이분변수 중재 결과의 경우, 요구되는 비율은 사건의 위험과 직접적으로 연관된다. 관찰된 사건 수가 결측치가 있는 연구 대상자 수보다 훨씬 많은 경우, 비뚤림은 매우 적을 것이다. • 임상시험 보고서에 결측치의 범위에 대한 정보를 제공하지 않는 경우에만 '정보 없음'으로 응답한다. 이러한 상황에서는 대개 '결측치로 인한 비뚤림위험'을 높다고 판단하게 될 것이다. • 결측치 대체(imputed)된 데이터는 결측치로 간주되어야 하며, 이 질문의 맥락에서는 '중재결과 자료'로 간주되지 않아야 한다. 	Y/PY/PN/N/NI
3.2 3.1에 대해 N/PN/NI인 경우: 결측치로 인해 연구 결과의 비뚤림이 없다는 근거가 있는가?	<ul style="list-style-type: none"> • 결측치로 인해 결과가 편향되지 않았다는 증거는 (1) 비뚤림을 교정하는 분석 방법, 또는 (2) 결측여부와 그 참값 사이의 관계에 대한 그럴듯한 가정 하에 연구결과가 거의 변하지 않는다는 것을 보여주는 민감도 분석에서 비롯될 수 있다. 그러나 최종관찰 이월(last-observation-carried-forward)과 같은 방법 또는 중재군에만 기초한 복수의 결측치 대체는 결측치로 인한 비뚤림위험을 교정하는 것으로 가정해서는 안 된다. 	NA/Y/PY/ PN/N

신호질문	세부 설명	응답 옵션
<p>3.3 3.2에 대해 N/PN인 경우: 중재 결과의 결측여부가 참값에 의존할 가능성이 있는가?</p>	<ul style="list-style-type: none"> 추적관찰 실패 또는 연구 참여 철회는 연구대상자의 건강 상태와 관련될 수 있다면, 중재 결과의 결측여부는 참값에 의해 영향을 받았을 수 있다. 그러나 모든 결측치가 중재 결과와 무관하다는 것이 문서화 되어 있으면, 결측치로 인한 비틀림위험은 낮을 것이다(예: 측정 장치의 고장 또는 일상적인 데이터 수집 중단). 사건발생시간 분석에서, 예를 들어 연구 참여 철회로 인해 추적관찰이 중단된 경우는 일부 추적 관찰 자료가 분석에 포함되더라도 결측치로 간주해야 한다. 이러한 연구 대상자는 CONSORT 흐름도에는 분석에 포함된 것으로 표시된다. 	<p>NA/Y/PY/ PN/N/NI</p>
<p>3.4 3.3에 대해 Y/PY/NI: 중재 결과의 결측여부가 참값에 의존할 가능성이 높은가?</p>	<ul style="list-style-type: none"> 이 질문은 (i) 결측여부가 참값에 좌우될 가능성이 있는 상황('일부 우려'로 평가됨)과 (ii) 자료 누락이 참값에 좌우될 가능성이 높은 상황('비틀림 위험이 높음으로 평가됨')를 구별한다. '예'라고 대답하는 다섯 가지 이유는 다음과 같다. <ol style="list-style-type: none"> 중재군간 결측치 비율의 차이. 실험 중재와 대조 중재가 중재결과에 미치는 영향이 차이가 있고, 결측여부가 참값에 영향을 받으면, 결측치의 비율이 중재군간에 다를 수 있다. 이러한 차이는 결측치로 인한 비틀림위험을 시사한다. 왜냐하면 시험 결과는 참값에 따른 중재결과 누락에 민감할 것이기 때문이다. 사건발생시간 분석의 경우, 중도절단 비율(추적관찰 실패)이 중재군마다 다르다는 것이 유사한 점이다. 결측치 이유로 보고된 내용에서 결측여부가 참값에 의존한다는 증거를 제공한다. 결측치 이유로 보고된 내용이 중재군마다 다르다. 임상시험의 정황으로 보아 결측여부는 참값에 의존할 가능성이 높다. 예를 들어, 정신분열증을 치료하기 위한 중재 임상시험에서 증상이 지속되는 경우 탈락이 더 많다는 사실은 널리 알려져 있다. 사건발생시간 분석에서, 연구대상자가 약물 독성 때문에 또는 암 임상 시험에서 2차 화학 요법으로 전환할 때, 배정된 중재를 중지하거나 변경하는 경우 중단이 발생하는 경우. 결측여부와 참값 사이의 관계를 설명하는 것이 연구대상자 특성인 경우 '아니오'에 답한다. 	<p>NA/Y/PY/ PN/N/NI</p>

비틀림위험 평가	알고리즘 참조	낮음/높음/일부 우려
<p>선택사항: 무작위 배정 과정에서 발생하는 비틀림의 예측 방향은 어떠한가?</p>	<p>비틀림의 방향을 예측할 수 있으면, 이를 기술하는 것이 도움이 된다. 방향은 무효점(혹은 무효점의 반대쪽)을 향한, 혹은 실험 중재번호, 비교중재 번호로 기술한다.</p>	<p>적용불가 / 실험 중재번호 / 비교중재번호/무효점을 향한/무효점에서 멀어짐/적용불가</p>

중재결과 자료의 결측으로 인한 비뚤림위험 판단 알고리즘



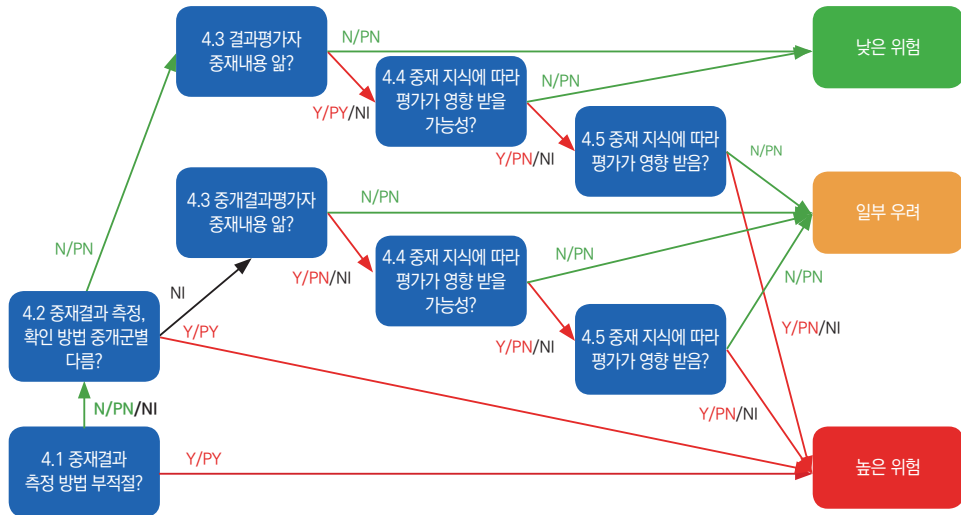
영역 4 중재결과 측정의 비뚤림위험

신호질문	세부 설명	응답 옵션
4.1 중재결과 측정 방법이 부적절 한가?	<ul style="list-style-type: none"> 이 질문은 평가하고자 하는 중재 결과에 적합하지 않은 결과 측정 방법(데이터 수집)을 확인하는 것이 목적이다. 이 질문은 평가되는 중재 결과의 선택이 합리적이었는지를 평가하기 위한 것이 아니다(예: 관심 주요 결과에 대한 대리결과 또는 중간결과). 대부분의 경우 사전 지정된 중재 결과에 대해 이 질문에 대한 응답은 'N' 또는 'PN'이 될 것이다. 예를 들어 다음과 같은 이유로 결과를 측정하는 방법이 적절하지 않은 경우 'Y' 또는 'PY'로 응답한다. <p>(1) 그럴듯한 중재 효과에 민감할 것 같지 않음(예: 중재 결과 값의 중요한 범위가 측정 방법을 사용하여 감지 가능한 수준을 벗어나는 경우) (2) 측정기기의 타당도가 불량하다는 것이 입증된 경우</p>	<p>Y/PY/ PN/N/NI</p>
4.2 중재군간 중재결과 측정 혹은 확인 방법이 다른가?	<ul style="list-style-type: none"> 비교 가능한 중재결과 측정 방법(데이터 수집)이란 유사한 시점에 동일한 측정 방법과 임계치로 측정하는 것이다. 중재군 간의 차이가 발생하는 경우는 중재결과 자료의 수동적 수집이라는 맥락에서 '진단적 결과확인 비뚤림' 때문이거나, 중재에 의료 제공자를 추가로 방문하는 것이 포함되어서 중재결과 사건을 확인할 수 있는 기회가 추가되는 경우이다. 	<p>Y/PY/ PN/N/NI</p>

신호질문	세부 설명	응답 옵션
4.3 4.1과 4.2에 대해 N/PN/NI 인 경우: 중재 결과 평가자는 연구대상자가 받은 중재를 알고 있는가?	<ul style="list-style-type: none"> 중재 결과 평가자가 중재 상태에 눈가림되어 있으면 'N'로 응답한다. 연구대상자 보고 결과의 경우, 중재 결과 평가자는 연구대상자이다. 	<p>NA/Y/PY/ PN/N/NI</p>
4.4 4.3에 대해 Y/PY/NI 인 경우: 중재결과에 대한 평가가 중재에 대한 지식에 의해 영향을 받을 가능성이 있는가?	<ul style="list-style-type: none"> 배정된 중재에 대한 지식은 연구대상자 보고 결과(통증 수준 등), 일부 판단이 포함된 관찰자 보고 결과 및 중재 제공자의 의사결정 결과에 영향을 미칠 수 있다. 판단을 수반하지 않는 관찰자가 보고한 중재 결과(예를 들어, 모든 원인 사망률과 같은)에 영향을 미칠 가능성은 낮다. 	<p>NA/Y/PY/ PN/N/NI</p>
4.5 4.4에 대해 Y/PY/NI 인 경우: 중재결과에 대한 평가가 중재에 대한 지식에 영향을 받았을 가능성이 높은가?	<ul style="list-style-type: none"> 이 질문은 (i) 중재 상태에 대한 지식이 결과 평가에 영향을 미칠 수 있었지만 그렇게 했다고 믿을 이유는 없는 상황('일부 우려'로 평가)과 (ii) 중재 상태에 대한 지식이 결과 평가에 영향을 미칠 가능성이 있는 상황('높음'으로 평가됨)을 구별한다. 중재가 유익할거라는 혹은 유해할 거라는 믿음이 강할 때, 중재 결과가 받은 중재에 대한 지식에 영향을 받았을 가능성이 더 높다. 동종요법 임상시험에서 환자 보고증상이나, 중재를 제공한 물리치료사가 기능회복에 대한 평가를 시행하는 경우 등이 예가 될 수 있다. 	<p>NA/Y/PY/ PN/N/NI</p>

비틀림위험 평가	알고리즘 참조	낮음/높음/일부 우려
선택사항: 무작위 배정 과정에서 발생하는 비틀림의 예측 방향은 어떠한가?	비틀림의 방향을 예측할 수 있으면, 이를 기술하는 것이 도움이 된다. 방향은 무효점(혹은 무효점의 반대쪽)을 향함, 혹은 실험 중재선호, 비교중재 선호로 기술한다.	적용불가 / 실험 중재선호 / 비교중재선호/무효점을 향함/무효점에서 멀어짐/예측불가

중재결과 측정의 비뚤림위험 판단 알고리즘



영역 5. 보고된 연구결과 선택의 비뚤림위험

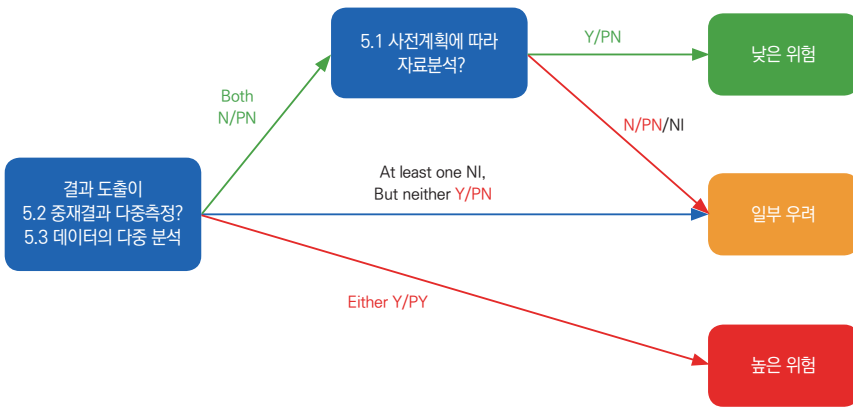
신호질문	세부 설명	응답 옵션
5.1 이 연구 결과를 도출한 자료는 분석할 수 있도록 미리 지정된 분석 계획에 따라 중재 결과 자료 눈가림이 해제되기 전에 분석되었는가?	<ul style="list-style-type: none"> • 사전에 지정된 연구자의 의도를 충분히 정도로 상세히 이용할 수 있는 경우, 계획된 중재 결과 측정과 보고서에 제시된 것을 비교할 수 있다. 보고된 결과의 선택 가능성을 피하기 위해, 분석 의도가 최종화되는 시점이 자료의 눈가림이 해제되는 시점보다 선행되어야 한다. • 자료의 눈가림이 풀리기 전에 분석계획이 변화한 경우 혹은 그러한 변화가 결과와 무관한 경우 (예를 들어, 고장난 기계로 인해 데이터 수집이 불가능함) 보고 결과의 선택에 대한 비뚤림 우려가 제기되지 않는다. 	<p>Y/PY/PN/N /NI</p>
평가되고 있는 수치형 연구결과는 다음과 같은 연구 결과에 기초하여 선택되었을 가능성이 있는가?		

신호질문	세부 설명	응답 옵션
<p>5.2 중재 결과 영역 내에서 여러 개의 적합한 중재 연구 결과 측정(예: 척도, 정의, 시점)</p>	<ul style="list-style-type: none"> 특정 중재 결과 영역(즉, 실제 상태 또는 관심 종결점)은 여러 가지 방법으로 측정될 수 있다. 예를 들어, 통증 영역은 여러 척도(예: 시간적 아날로그 척도 및 맥길 통증 설문도구)를 사용하여 각각 여러 시점(예: 치료 후 3, 6, 12주)에서 측정할 수 있다. 복수의 측정이 이루어졌지만 결과(예: 통계적 유의성)에 근거하여 1개 또는 하나의 하위세트만 보고되는 경우, 비폭력위험은 높을 것이다. RoB 2 도구 사용자가 고려하기에 적절한 중재결과만 주의를 기울이면 된다. 예를 들어 특정 측정 척도를 사용한 결과만 메타 분석(예: 해밀턴 우울증 평가 척도)에 포함할 수 있고, 임상시험에 이 중재결과가 보고된 경우, 다른 측정 척도에서 나온 결과(예: Beck depression Inventory)보다 우선적으로 이 결과가 보고되었다 하더라도(통계적 의미성 등에 의하여 해당하여도) 선택 문제는 없을 것이다. 다음과 같은 경우 'Y' 또는 'PY'로 응답한다. 한 영역을 복수의 적격 방법으로 측정했다는 명확한 증거(보통 임상 시험 프로토콜이나 통계 분석 계획서의 검토를 통해)가 있지만, 여러 측정 결과 중 하나 또는 일부 subset만 완전히 보고되고 (사유 없이), 최종 보고 결과는 그 결과(통계적 유의성 등)에 근거하여 선택되었을 가능성이 높다. 결과에 기초하여 선택하는 것은 연구 결과가 뉴스 가치가 있고, 출판할 가치가 있거나, 사전 가설을 확인하기에 충분히 주목할 만한 것이어야 한다는 열망에서 비롯될 수 있다. 예를 들어, 임상시험 중재가 유익하다는 선입견을 가지고 있거나 이해상충이 있는 임상시험 연구자들은 임상시험 중재에 유리한 결과 측정을 선택적으로 보고하는 경향이 있을 수 있다. 다음과 같은 경우 'N' 또는 'PN'으로 응답한다. (일반적으로 시험 프로토콜 또는 통계 분석 계획의 조사를 통해) 결과 영역에 대해 모든 적격 보고 결과가 의도된 모든 중재 결과 측정과 일치한다는 명확한 증거가 있다. 아니면 중재 결과 영역을 측정할 수 있는 방법은 단 한 가지뿐이다(여러번 측정해서 하나를 선택할 기회가 없는 경우). 아니면 동일한 임상시험에 대한 서로 다른 보고서 간에 중재 결과 측정들이 일관되지 않지만, 임상시험 시행자들이 불일치의 이유를 제공했고, 결과의 성격과 관련이 없다. 다음과 같은 경우 '정보 없음'으로 응답한다 분석 의도를 이용할 수 없거나, 분석 의도를 충분히 상세히 보고하지 않아 평가할 수 없고 중재 결과 영역을 측정할 수 있었던 방법이 두 개 이상 있다. 	<p>Y/PY/PN/N /NI</p>

신호질문	세부 설명	응답 옵션
<p>5.3 데이터의 다중 적격 분석</p>	<p>특정 중재 결과는 여러 가지 방법으로 분석할 수 있다. 예를 들어, 비보정 vs 보정 모형; 최종값 vs 기저 상태에서부터 변화량 vs 공분산 분석; 변수 변환; 복합 결과에 대한 다른 정의(예: '주요 유해사건'); 다양한 역치를 가진 연속형 변수를 범주형 변수로 변환; 다양한 공변량 집합에 의한 보정; 결측치를 처리하기 위한 다른 전략 등이 그것이다. 복수의 방법을 적용하면 특정 결과 측정에 대한 복수의 효과 추정치가 생성된다. 복수의 추정치가 생성 되지만 결과(예: 통계적 유의성)에 근거하여 1개 결과 또는 하나의 하위 세트만 보고되는 경우, 최종 보고 결과는 비뮴림위험이 높다. RoB 2 도구 사용자가 적격으로 고려할 수 있는 분석으로 제한하도록 주의를 기울여야 한다. 예를 들어, 중재 시행 후 측정치만 메타 분석(예: 무작위배정 후 12주)에 포함하는 것이 적격이고, 이 결과가 보고된 경우, 기저치로부터 변화량에 우선하여 이 결과에 기초해서 보고하였다고 하여도 선택 문제는 없을 것이다.</p> <p>다음과 같은 경우 'Y' 또는 'PY'로 응답한다 (보통 시험 프로토콜이나 통계 분석 계획의 조사를 통해) 측정이 복수의 적격 방법으로 분석되었다는 명확한 증거가 있지만, 분석의 한 부분 또는 일부에 대한 데이터는 충분히 보고되고, 최종 보고 결과는 하나만 혹은 하나의 하부세트만 보고한 경우(이에 대한 정당화 없이), 최종 보고 결과는 결과에 기초해서 선택되었을 가능성이 높은 경우이다. 결과에 기초하여 선택하는 것은 연구 결과가 뉴스 가치가 있고, 출판할 가치가 있거나, 사전 가설을 증명하기에 충분히 주목할 만한 것이어야 한다는 열망에서 비롯될 수 있다. 예를 들어, 임상시험 중재가 유익하다는 선입견을 가지고 있거나 이해상충이 있는 임상시험 연구자들은 임상시험 중재에 유리한 결과 측정을 선택적으로 보고하는 경향이 있을 수 있다.</p> <p>다음과 같은 경우 'N' 또는 'PN'으로 응답한다. 결과 측정에 대한 모든 적격 보고 결과가 의도된 분석과 동일하다는 명확한 증거가 있다(일반적으로 임상 시험 프로토콜 또는 통계 분석 계획 조사를 통해). 아니면 결과 측정을 분석할 수 있는 방법은 단 한 가지뿐이다(다중 분석에서 선택할 기회가 없는 경우). 아니면 동일한 임상시험에 대한 서로 다른 보고서 간에 중재 결과 측정들이 일관되지 않지만, 임상시험 시행자들이 불일치의 이유를 제공했고, 결과의 성격과 관련이 없다.</p> <p>다음과 같은 경우 '정보 없음'으로 응답한다 분석 의도를 이용할 수 없거나, 분석 의도를 충분히 상세히 보고하지 않아 평가할 수 없고 중재 결과 영역을 측정할 수 있었던 방법이 두 개 이상 있다.</p>	<p>Y/PY/PN/N /NI</p>

비뚫림위험 평가	알고리즘 참조	낮음/높음/일부 우려
선택사항: 무작위 배정 과정에서 발생하는 비뚫림의 예측 방향은 어떠한가?	비뚫림의 방향을 예측할 수 있으면, 이를 기술하는 것이 도움이 된다. 방향은 무효점(혹은 무효점의 반대쪽)을 향한, 혹은 실험 중재선호, 비교중재 선호로 기술한다.	적용불가 / 실험 중재선호 / 비교중재선호/무효점을 향한/무효점에서 멀어짐/예측불가

보고된 연구결과 선택의 비뚫림위험 판단 알고리즘



전반적 비뚫림위험

전반적인 비뚫림위험 기준	기준
낮은 비뚫림위험	본 연구는 현 결과의 모든 영역에 대해 비뚫림위험이 낮은 것으로 판단된다.
일부 우려	본 연구는 현 결과의 적어도 한 영역에서 일부 우려가 있으나 어떠한 영역에도 비뚫림위험이 높지 않다.
높은 비뚫림위험	본 연구는 현 결과의 적어도 한 영역에서 비뚫림위험이 높은 것으로 판단된다. 혹은 현 연구는 여러 영역에서 일부 우려가 있어서 결과의 신뢰도를 상당히 떨어뜨리는 것으로 판단된다.

라. 국문 ROB 2 사용자 매뉴얼

1) 사전 고려사항

- 비뿔림위험 평가를 완료하기 전에, 평가 대상의 설계, 평가할 중재 결과(평가되는 특정 연구결과뿐만 아니라), 관심 효과가 배정효과인지 중재 준수효과 같은 평가의 중요한 특성을 문서화하는 것이 도움이 된다.
- 문헌고찰 저자는 평가를 완료하는 데 사용되는 출처를 문서화해야 한다(실무적으로 가능한 많은 출처를 사용해야 한다).
- RoB 2 표준 템플릿에는 이러한 세부 사항을 기록하고(박스 #, 어떤 중재가 '임상시험 중재' 인지, 어떤 중재가 비교 중재인지를 명확하게 하게 기술해야 한다.

연구설계

개별 무작위배정 평행군 임상시험

군집 무작위배정 평행군 임상시험

개별 무작위배정 교차(혹은 other matched) 임상시험

평가 목적의 관점에서 중재와 비교중재의 정의는?

중재 비교중재

비뿔림위험 평가를 시행하는 중재결과는 구체적으로 무엇인가?

평가할 수치형 연구결과는 구체적으로 무엇인가?

여러 분석이 시행된 경우 수치형 연구결과가 구체적으로 무엇인지(예 : RR = 1.52 (95% CI 0.83 to 2.77)), 평가 결과를 어디에서 얻었는지 명시할 것(표, 그림, 문장 등)

현 결과에 대한 평가 목적은 무엇인가?

중재 배정 효과 평가(ITT 효과)

중재 준수 효과 평가(PP 효과)

평가 목적이 중재 준수 효과라면, 다루어야 할 중재 이탈의 종류는? (최소한 한 개는 체크)

- 비프로토콜 중재의 발생
- 중재결과에 영향을 미칠 수 있는 중재 실행 실패
- 연구대상자가 배정된 중재 비준수

비돌림위험 평가 정보는 어디에서 얻었나? (모두 체크하시오)

- 학술지 논문 임상시험 프로토콜
- 통계분석 계획(Statistical analysis plan, SAP)
- 비상업적 임상시험 등록처 자료(예: ClinicalTrials.gov 자료)
- 회사 소유 임상시험 등록처 자료(예: GSK Clinical Study Register record)
- 회색 문헌 (예: 출판되지 않은 학위 논문) 학회 초록
- 규제 문서(예: Clinical Study Report, Drug Approval Package)
- IRB 지원 서류 연구비 지원 자료베이스 요약(예: NIH RePORTER 등)
- 임상시험 시행자와 사적 교신 스폰서와 사적 교신

2) 사전 고려사항 평가의 예

계절성 인플루엔자 감염환자에서 경구용 오셀타미비르와 정맥용 페라미비르 중재의 무작위배정 이중눈가림 연구(Kohno et al., 2011)¹⁾에서 적용결과는 다음과 같다.

1) Kohno S, Yen MY, Cheong HJ, Hirotsu N, Ishida T, Kadota J, Mizuguchi M, Kida H, Shimada J; S-021812 Clinical Study Group. Phase III randomized, double-blind study comparing single-dose intravenous peramivir with oral oseltamivir in patients with seasonal influenza virus infection. Antimicrob Agents Chemother. 2011 Nov;55(11):5267-76. doi: 10.1128/AAC.00360-11.

연구설계

- 개별 무작위배정 평행군 임상시험
- 군집 무작위배정 평행군 임상시험
- 개별 무작위배정 교차(혹은 other matched) 임상시험

평가 목적의 관점에서 중재와 비교중재의 정의는?

중재 Single-Dose Intravenous Peramivir 비교중재 Oral Oseltamivir

비뿔림위험 평가를 시행하는 중재결과는 구체적으로 무엇인가?

the time to alleviation of influenza symptoms

평가할 수치형 연구결과는 구체적으로 무엇인가?

여러 분석이 시행된 경우 수치형 연구결과가 구체적으로 무엇인지(예 : RR = 1.52 (95% CI 0.83 to 2.77)), 평가 결과를 어디에서 얻었는지 명시할 것(표, 그림, 문장 등)

TABLE 3. Time to alleviation of symptoms for the intent-to-treat infected population Peramivir 300 mg group HR= 0.946 (0.793, 1.129)
Peramivir 600 mg group HR= 0.970 (0.814, 1.157)
Oseltamivir group Reference

현 결과에 대한 평가 목적은 무엇인가?

- 중재 배정 효과 평가(ITT 효과)
- 중재 준수 효과 평가(PP 효과)

평가 목적이 중재 준수 효과라면, 다루어야 할 중재 이탈의 종류는? (최소한 한 개는 체크)

- 비프로토콜 중재의 발생
- 중재결과에 영향을 미칠 수 있는 중재 실행 실패
- 연구대상자가 배정된 중재 비준수

비뿔림위험 평가 정보는 어디에서 얻었나? (모두 체크하시오)

- 학술지 논문 임상시험 프로토콜
- 통계분석 계획(Statistical analysis plan, SAP)
- 비상업적 임상시험 등록처 자료(예: ClinicalTrials.gov 자료)
- 회사 소유 임상시험 등록처 자료(예: GSK Clinical Study Register record)
- 회색 문헌 (예: 출판되지 않은 학위 논문) 학회 초록
- 규제 문서(예: Clinical Study Report, Drug Approval Package)
- IRB 지원 서류 연구비 지원 자료베이스 요약(예: NIH RePORTER 등)
- 임상시험 시행자와 사적 교신 스폰서와 사적 교신

3) 영역별 평가 매뉴얼

아) 영역1. 무작위과정에서 생기는 비틀림

(1) 일러두기

- 연구대상자를 무작위로 배정하기 위해 연구대상자가 중재에 배정되는 방법을 지정하는 배정 순서를 우연의 요소를 포함하는 프로세스에 기반 하여 생성되도록 하는 것을 **배정 순서 생성**이라고 하며, 모집 확정 후까지 연구대상자나 임상시험 담당자가 향후 배정을 알지 못하도록 조치를 취하는 것을 **배정 순서 은폐**라고 한다.
- 순서 생성에 대한 접근 방법 중 블록 무작위 배정(제한적 무작위 배정)을 사용하는 경우 각 블록 내에서 무작위 순서로 배정되어야 하는데 그 이유는 블록 크기가 임상시험 담당자에게 알려지면 각 블록 내의 마지막 배정을 항상 예측할 수 있기 때문이다.
- **최소화**법은 계층화와 제한된 무작위 배정을 모두 포함하는 특정 예측 인자와 관련하여 중재군을 밀접하게 유사하게 만드는 데 사용될 수 있다. 최소화에는 일반적으로 무작위 요소(최소 알고리즘에 포함된 예측 요인과 관련하여 균이 균형을 이루도록 연구대상자를 참여시킴)를 포함한다.
- 배정은폐에 실패하여 연구대상자가 부적절한 것으로 생각하는 중재에 배정되면 참여를 거부하거나 원하는 배정이 될 때까지 임상시험 진입을 늦춤으로써 '적절한' 중재를 받도록 조작할 수 있다. 향후 배정을 예상하여 배정 은폐의 실패로 이어질 수 있는 방법에는 다음과 같은 것이 있다.
 - (1) 교대, 생일 또는 내원일과 같은 결정론적 배정 규칙을 알게 되는 것
 - (2) 배정 순서에 대한 지식(예: 무작위 배정 순서가 게시판에 공개적으로 게시된 경우)
 - (3) 이전 배정을 기준으로 배정을 성공적으로 예측할 수 있는 것
- 배정순서 은폐에 대해 판단하는 최소 기준, 확장기준은 다음과 같다.

최소 기준	확장 기준
중앙 무작위법	중앙의 무작위 배정 센터는 환자 모집 센터에서 멀리 떨어져 있었다. 연구대상자의 세부 사항은 전화(상호작용하는 음성 응답 시스템 포함), 이메일 또는 상호작용하는 온라인 시스템으로 제공되었고, 연구대상자가 최종 등록될 때까지 배정 순서를 숨겼다.
연속번호가 부여된 약용기	독립적인 약사가 준비한 약용기에 연속 번호를 매기고 순차적으로 개봉했다. 약용기는 외관이 같았고, 조작이 불가능했으며(tamper-proof) 무게도 같았다.
연속번호가 부여된 밀봉, 불투명 봉투	봉투에 연구대상자 세부사항을 적은 후에만 봉투에 연속 번호를 부여하고 순차적으로 개봉했다. 봉투 안에 압력을 감지하는 또는 탄소 용지가 연구대상자의 세부 정보를 배정 카드로 전송했다. 봉투 안의 판지나 알루미늄 포일이 광택한 빛에 비칠 수 없게 만들었다. 봉투는 변조 방지 보안 테이프로 봉인되었다.

- 기저상태 불균형 평가는 무작위 배정된 모든 연구대상자 자료에 기초해야 한다. 기저 상태 자료가 임상시험을 완료한 연구대상자(또는 무작위 배정된 연구대상자 중 일부 다른 하위 집합)에게만 제공되는 경우, 기저상태 불균형을 평가하기가 더 어려우며, 자료누락의 비율과 이유를 고려할 필요가 있다.

- 우연에 의한 불균형은 체계적인 비뚤림의 원천이 아니며, RoB 2 도구는 **우연으로 인해 발생한 기저 상태 변수의 불균형을 확인하는 것을 목표로 하지 않는다**. 기존 0.05 임계값에서 '통계적으로 유의함'으로 확인된 약간의 차이는 대개 우연의 결과로 간주해야 한다.
- 다음과 같은 경우 무작위 과정에 문제가 있었음을 암시하는 기저상태 불균형으로 본다.
 - (1) 의도된 배정 비율과 비교하여 중재군 크기 간의 상당한 차이
 - (2) 중재군 간 기저상태 특성 차이가 통계적으로 유의한 차이가 나고 이것이 우연으로 예상되는 범위를 벗어난 경우
 - (3) 확률로 인해 발생할 가능성이 낮은 주요 예측 요인 또는 결과 변수의 기저상태 불균형
 - (4) 우연으로 설명할 수 없는 기저상태 특성의 과도한 유사성

(2) 무작위 배정 과정에서 발생하는 비뚤림위험 평가 가이드

신호질문	세부 설명	응답 옵션
1.1 배정 순서는 무작위였나?	<ul style="list-style-type: none"> · '그렇다(Y)' : 순서 생성 과정에서 무작위 구성요소가 사용된 경우 (예를 들어 컴퓨터 난수 생성, 난수표 사용, 동전 던지기, 카드나 봉투 뽑기, 주사위 던지기, 추첨 등). 최소화법은 일반적으로 무작위 요소(적어도 점수가 같을 때)로 구현되므로 무작위로 간주. · '아니다(N)' : 배정 순서를 생성하는 데 무작위 요소를 사용하지 않았거나 순서가 예측 가능한 경우 (예를 들어 교대, 날짜(생년 월일 또는 입원 일)에 기초한 방법, 환자 의무기록 번호, 임상사 사 또는 연구대상자에 의한 배정 결정, 중재의 가용성에 기초한 배정 또는 그 밖의 체계적 또는 우발적 방법 등) · '정보 없음(NI)' : 무작위 배정 방법에 대한 유일한 정보가 "연구가 무작위로 진행되었다"라는 문구일 경우 · '아마도 아니다(PN)' 또는 '아마도 그렇다(PY)' : '상황에 대한 판단에 따라(예를 들어, 숙련된 임상시험팀에서 시행한 대규모 임상시험 논문이 논문 분량 제한이 엄격하게 적용되는 학술지에 게재되었다면, 무작위 순서 생성에 대한 특정 정보가 없는 경우 '정보 없음'이 아니라 '아마도 그렇다'라고 응답, 유사한 시기에 동일한 연구팀에서 수행한 다른 임상시험이 비무작위 연구로 진 행되었다고 되어 있으면 현재의 연구는 유사한 방법을 사용하여 수행되었다고 가정하는 것이 합리적일 수 있음 	<p>Y/PY/ PN/N/NI</p>

예시

적절한 무작위 순서생성 예 :

“우리는 단순 무작위법으로, 동일한 배정 비율 사용하고, 난수표를 써서 두 개의 비교군을 생성했다.”

“무작위 순서생성 예: “블록 무작위법을 사용하여 두 개의 비교군에 대한 배정 목록을 구성했다. 컴퓨터 무작위 번호 발생기를 이용해 블록 크기가 8이고 배정 비율이 같은 임의의 permuted blocks을 사용하였다”

부적절한 순서 생성 예 :

“환자들은 1차 거주자 성씨의 첫 글자에 의해 무작위로 배정되었다.”

“짜수날 날짜에 태어난 사람은 실내 공기(룸 에어군)로 배정했고, 흄수 날짜에 태어난 사람은 100% 산소군(산소군)에 배정하였다.”

신호질문	세부 설명	응답 옵션
<p>1.2 연구대상자가 모집되어 증재에 배정될 때까지 배정순서가 은폐되었나?</p>	<ul style="list-style-type: none"> • Y : 1) 임상시험에서 연구대상자들에게 증재를 배정하기 위해 원격 또는 중앙 관리 방법을 사용한 경우 혹은 배정 과정은 모집 담당자와 무관하게 외부 조직에서 통제하는 경우(예: 독립적 중앙 약국, 전화 또는 인터넷 기반 무작위화 서비스 제공자), 2) 봉투나 약용기가 적절하게 사용된 경우(봉투는 불투명하고 순차적으로 번호가 부여되어야 하며 변조 방지 봉인으로 밀봉. 또한 연구대상자에게 봉투가 비가역적으로 배정된 후에만 개봉되어야 함. 약물 용기는 순차적으로 번호가 매겨져야 하며 동일한 외관을 가지며, 연구대상자에게 비가역적으로 배정된 후에만 폐기 또는 투여 • N : 모집담당자 또는 연구대상자가 향후 배정에 대한 지식을 가지고 있다고 의심할 만한 이유가 있는 경우 • 'PY' 또는 'PN' : 이 수준의 세부사항은 보고서에 거의 제공되지 않으며, 답변의 정당화를 위해서는 판단이 필요한 경우 	<p>Y/PY/ PN/N/NI</p>

예시

적절한 배정순서 은폐의 예

"...연속 번호 부여, 밀봉, 불투명한 봉투에 의해 은폐되었고, 두 센터 병원 약사가 보관하고 있다."

"포함기준의 정확성에 대한 전화 검증을 받고 중앙에서 처방을 배정하였다..."

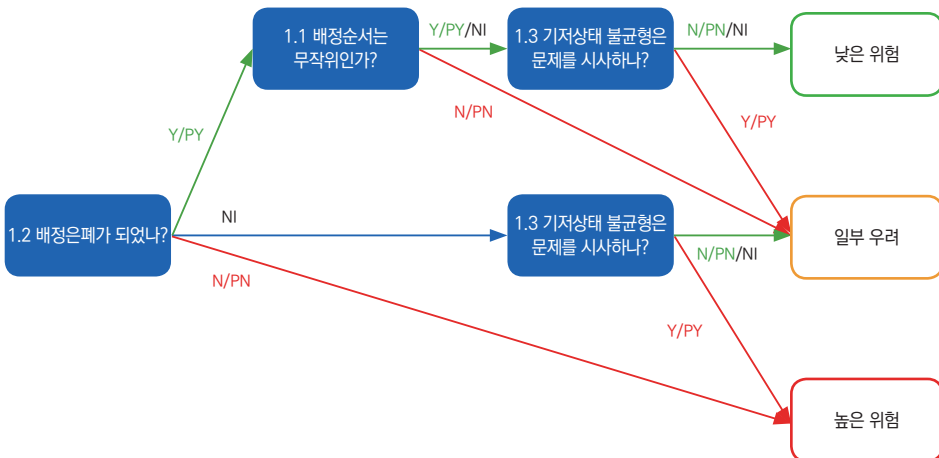
"글렌필드 병원 약국은 무작위 배정을 하고, 약제를 배포하고, 시험코드를 보유하고, 이것은 연구 후에 공개되었다."

<p>1.3 증재기간 기저상태 차이가 무작위 과정의 문제를 시사하나?</p>	<p>주요 포인트</p> <ol style="list-style-type: none"> (1) 우연에 의한 차이로 간주되는 차이는 비돌림위험으로 이어지지 않는다는 점에 유의한다. 기준 0.05 임계값에서 '통계적으로 유의함'으로 확인된 적은 차이는 우연에 의한 차이로 간주해야 함 (2) 이 질문에 대한 응답이 1.1 또는 1.2 질문에 대한 응답에 영향을 미치지 않아야 함. 예를 들어, 임상시험 결과에 있어서 기저상태 불균형을 가지고 있지만 저자들이 적절한 무작위화 방법을 보고하는 경우, 질문 1.1과 1.2에 대한 응답은 보고된 적절한 방법에 근거하여 답변해야 하며, 불균형에 대한 우려는 질문 1.3에 대한 답변에서 제기되어야 하며, 영역 수준 비돌림위험 판단에 반영되어야 함 (3) 임상시험 시행자는 예측 인자의 불균형을 기저 상태에서 통제함으로써 결함있는 무작위화를 처리하려는 분석을 수행할 수 있음. 무작위화 과정의 문제로 인한 비돌림위험을 제거하기 위해서는 기저상태에서 불균형이 있는 모든 예측 요인을 알고 측정할 필요가 있을 것임 (4) 모든 중요한 예측 요인을 알고 측정할 가능성은 낮으므로 그러한 분석은 기껏해야 비돌림의 위험을 줄이는 정도임. 문헌고찰 저자가 무작위화 실패를 완화하기 위해 기저상태 불균형을 보정하는 임상 시험의 비돌림위험을 평가하고자 하는 경우, ROBINS-I 도구를 사용하여 평가해야 함 	<p>Y/PY/ PN/N/NI</p>
--	---	--------------------------

신호질문	세부 설명	응답 옵션
	<ul style="list-style-type: none"> • Y : 무작위화 과정의 문제를 시사하는 다음과 같은 불균형이 있는 경우 (1) 의도된 배정 비율과 비교하여 중재군 크기 간의 상당한 차이 혹은 (2) 중재군 간 기저상태 특성의 통계적으로 유의한 차이에 대한 우연이라고 할 수 없는 상당한 초과 혹은 (3) 하나 이상의 핵심 예측 요인 또는 예측 변수에 대한 기저 상태 불균형이 우연으로 인해 발생할 가능성이 매우 낮고, 군간 차이가 중재 효과 추정의 비뿔림을 초래할 정도로 큰 경우. 또한 무작위화 과정에 문제가 있었다고 의심하는 다른 이유가 있는 경우, 또는 (4) 우연이라고 보기 힘들 정도로 기저상태 특성이 과도하게 유사한 경우 • N : 불균형이 뚜렷하지 않거나 관찰된 불균형이 우연에 의한 차이로 간주해야 하는 경우 • NI : 이용 가능한 유용한 기준 정보가 없을 때(예: 최종 분석 참여자의 기준 특성만 보고한 연구) 	

비뿔림위험 평가	알고리즘 참조	낮음/높음/일부 우려
선택사항: 무작위 배정 과정에서 발생하는 비뿔림의 예측 방향은 어떠한가?	비뿔림의 방향을 예측할 수 있으면, 이를 기술하는 것이 도움이 된다. 방향은 무효점(혹은 무효점의 반대쪽)을 향한, 혹은 실험 중재선호, 비교중재선호로 기술한다.	적용불가/실험 중재선호/비교중재선호/무효점을 향한/무효점에서 멀어짐/예측불가

(3) 무작위 배정 과정에서 발생하는 비뿔림위험 판단을 위한 알고리즘



(4) 무작위 배정 과정에서 발생하는 비뚫림위험 판단 예시

계절성 인플루엔자 감염환자에서 경구용 오셀타미비르와 정맥용 페라미비르 중재의 무작위 배정 이중눈가림 연구(Kohno et al. 2011)²⁾에서 적용결과는 다음과 같다.

신호질문	서술	응답
1.1 배정 순서는 무작위였나?	Study design. Our study was a multicenter, double-blind, randomized, controlled study with dynamic allocation using the minimization method....	<u>Y</u>
1.2 연구대상자가 모집되어 중재에 배정될 때까지 배정순서가 은폐되었나?	No information about allocation concealment.	NI
1.3 중재군간 기저상태 차이가 무작위 과정의 문제를 시사하나?	The three treatment groups did not differ significantly with respect to any baseline characteristics (Table 1).	<u>N</u>
비뚫림위험 평가		일부우려

자) 영역 2. 의도한 중재에서 이탈로 인한 비뚫림

(1) 일러두기

- 이 영역은 의도한 중재에서 이탈이 있을 때 발생하는 비뚫림에 대한 것이다. 이러한 이탈은 임상임상시험 프로토콜과 일치하지 않는 추가 중재, 의도한 프로토콜 중재 실행 실패, 임상시험 연구대상자가 배정된 중재 비준수 등이 포함된다. 의도한 중재에서 이탈로 인한 비뚫림은 실행 비뚫림의 일종이다.
- 임상임상시험 프로토콜과 일치하는 중재에 대한 변경은 비뚫림을 유발하지 않으며, 의도한 중재에서 이탈로 간주되지 않아야 한다.
- 비프로토콜 중재는 임상시험 연구대상자가 임상시험 추적관찰 기간 동안 받을 수 있고 관습 결과에 영향을 줄 수 있는 비프로토콜 중재는 추정된 중재 효과에 비뚫림을 유발할 수 있다. 가능한 경우, 문헌고찰 저자는 잠재적인 비프로토콜 중재를 사전에 지정해야한다(문헌고찰 프로토콜 작성 단계에서). 그것은 문헌고찰팀 구성원의 전문 지식, 문헌의 초기(주제범위) 문헌고찰 및 보건의료 전문가와의 토론을 통해 확인 될 수 있다.

2) Kohno S, Yen MY, Cheong HJ, Hirotsu N, Ishida T, Kadota J, Mizuguchi M, Kida H, Shimada J; S-021812 Clinical Study Group. Phase III randomized, double-blind study comparing single-dose intravenous peramivir with oral oseltamivir in patients with seasonal influenza virus infection. Antimicrob Agents Chemother. 2011 Nov;55(11):5267-76. doi: 10.1128/AAC.00360-11.

· **관심 효과의 역할:** 이 영역에 대한 평가는 문헌고찰 저자가 관심을 갖는 중재 효과에 따라 달라진다. 이러한 효과는 일부 환자가 배정된 중재를 받지 않거나 기저시점 후에 배정된 중재에서 이탈이 생기면 달라진다.

- (1) 기저상태에서 중재에 대한 배정의 효과 (추적관찰 기간 동안 중재를 받거나 준수했는지 여부에 관계없이, 때로는 '중재 배정 효과'라고도 함) 또는
- (2) 임상시험 프로토콜에 명시된 대로 중재를 준수한 효과 ('중재 준수 효과'라고도 함)

· **중재 배정 효과에서 비뿔림:**

일부 임상 시험, 특히 눈가림이 되지 시험에서는 임상시험 모집 및 참가 활동으로 인해 비프로토콜 중재가 발생할 수 있다. 중재 제공자가 비 프로토콜 중재를 시행하거나 프로토콜 중재를 시행하지 않음으로써 임상 시험을 통한 중재 비교의 과학성이 훼손 할 수도 있다. 이것은 무의식적 과정 (예를 들어, 임상적 동등성이 명확하지 않아 한 그룹에서 비 프로토콜 중재를 시행하고 다른 군에는 프로토콜 중재를 제공하는 등)일수도 있고 무의식적 과정 (예 : 이해 상충으로 인한)일 수도 있다

우리는 연구대상자 모집 및 참여 활동의 영향과 임상시험 인력이 중재를 전달하는 과정에서 나타나는 무의식적 또는 의식적 과정의 영향을 임상시험적 맥락이라는 용어로 표현한다. RoB 2에서 중재 배정의 효과와 관련하여 다루는 의도한 중재에서 이탈은 다음과 같다.

- (1) 임상시험 맥락에서 발생했다.
- (2) 임상시험 프로토콜과 일치하지 않았다.
- (3) 중재결과에 영향을 미친다.

예를 들어, 눈가림되지 않은 연구에서 동의를 확보하는 과정은 추후 비교 그룹에 배정된 연구대상자가 운이 좋지 않다고 느껴 임상시험 중재 또는 예후를 향상시킬 수 있는 다른 중재를 받으려고 할 수 있다. 강력한 모집 또는 참여 효과가 있는 임상시험의 효과 추정치는 그런 효과가 없는 환경에는 적용되지 않을 수 있기 때문에 RoB 2에서 평가된다.

비프로토콜 중재는 중재의 안전성 (임상적 동등성에 못미침)에 대한 외과 의사의 신념으로 인해 일어나며, 중재에 대한 배정 중재 추정 효과는 비뿔림위험이 높아진다. 하지만 환자가 배정 약물을 복용 중단 하는 것과 같은 임상시험 맥락 때문에 발생한 이탈은 중재 배정 효과에 비뿔림을 초래하지 않는다.

· **중재 준수 효과의 비뿔림**

임상시험 프로토콜과 일치하지 않고 중재 결과에 영향을 미치는 의도한 중재에서 이탈은 임상시험 맥락에서 발생했는지 여부에 관계없이 중재 준수 효과에 관련하여 다루어져야 한다. 의도한 중재에서 이탈한 것을 보정할 수없는 경우가 있다. 그러므로, 임상 임상시험 프로토콜에 명시된 대로 중재 순응효과를 다룰 때 문헌고찰 저자는 사전 고려 사항 (섹션 3 참조)에서 의도한 중재 (임상임상시험 프로토콜에서 출발)에서 어떤 유형의 이탈이 조사되는지를 명시하여야 한다. 그러한 것에는, (1) 중재 결과에 영향을 줄 수 있는 비프로토콜 중재의 발생; (2) 중재 결과에 영향을 줄 수 있는 프로토콜 중재 실패; 그리고 (3) 시험 연구대상자가 배정된 중재 비준수. 중재 준수에서 이탈이 있는 경우 문헌고찰 저자는 적절한 통계적 방법을 사용하여 효과를 보정했는지 여부를 고려해야 한다.

· **눈가림의 역할**

연구대상자, 중재 제공자가 제공하는 중재를 인지하지 못하도록 하는 것을 눈가림이라고 한다. 일부 영역(예: 안과)에서는 '마스킹'이라는 용어가 선호된다. 성공적인 경우, 눈가림은 중재 배정에 대한 지식 때문에 오염(특정 중재에 배정된 연구대상자가 나머지 다른 중재에 적용을 받는 것)이 늘어나는 것을 막고, 프로토콜에 없는 중재를 받거나 임상시험 연구대상자가 배정된 중재를 받지 않는 것을 방지할 수 있다.

눈가림이 시행되지 않으면, 임상시험적 맥락과 관련된 의도한 중재에서 이탈이 발생시키거나, 중재준수 효과를 추정하기 위해 중재군간 이탈의 정도가 달라진다면 비뚫림을 야기할 수 있다. 예를 들어 수술을 비수술적 중재와 비교하는 임상시험에서, 임상시험 연구대상자에 대한 눈가림은 어렵거나 일부 상황에서는 불가능하다. 눈가림 되지 않은(‘공개된’) 연구는 프로토콜에 없는 중재의 시행을 막기 위한 엄격한 기준에 따라 환자를 치료하는 것과 같이 의도한 중재에서 이탈을 피하기 위해 다른 조치를 취할 수 있다. 연구 출판물들은 자주 “이중 눈가림”과 같은 광범위한 용어로 눈가림을 기술한다. 그러나 이 용어는 누가 눈가림 되었는지 알기 어렵다.

(2) 의도한 중재에서 이탈로 인한 비뚫림위험 평가 가이드(중재 배정 효과)

신호질문	세부 설명	응답 옵션
2.1 연구대상자는 임상시험 중에 배정된 중재에 대해 알고 있는가?	<p>연구대상자가 배정된 중재를 알고 있다면, 건강 관련 행동이 중재군에 따라 다를 가능성이 있다. 연구대상자에 대한 눈가림은 위약이나 가짜 중재를 통해, 그러한 차이를 방지할 수 있다.</p> <p>· Y/PY: 연구대상자가 중재 중 하나에 특정한 것으로 알려진 부작용이나 독성을 경험한 경우</p>	<p>Y/PY/ PN/N/NI</p>
2.2 보호자와 중재 제공자는 임상시험 중에 연구대상자에 배정된 중재에 대해 알고 있는가?	<p>중재 제공자가 배정된 중재를 알고 있다면, 그 중재를 실행했는지, 비프로토콜 중재가 발생했는지 여부가 중재군 간에 다를 수 있다. 눈가림은 그러한 차이를 막을 수 있다.</p> <p>· Y/PY: 중재제공자가 중재 중 하나에 특정한 것으로 알고 있는 부작용이나 독성을 연구대상자들이 경험한 경우, 무작위 배정 은폐가 되지 않는 경우, 중재제공자는 임상시험 중에 연구대상자들에 배정된 중재가 무엇인지 알고 있었을 가능성이 높다.</p>	<p>Y/PY/ PN/N/NI</p>
2.3 2.1과 2.2의 Y/PY//NI 인 경우 : 임상시험적 맥락 때문에 의도한 중재에서 이탈이 있었나?	<p>중재 배정 효과의 경우, 임상시험 프로토콜과 일치하지 않는 배정 중재의 변경이 임상시험적인 맥락 때문에 발생하여 결과에 영향을 미칠 가능성이 있는 경우에만 이 영역의 비뚫림이 발생한다. 예를 들어, 동의를 받는 과정에서 대조 중재에 배정되면 불운한 것이라고 느끼게 되어, 임상시험 중재나 더 나은 예후를 가지는 중재에 배정되도록 노력할 수 있다.</p> <p>· Y/PY: 임상시험적 맥락으로 인해 임상시험이 진행되지 않은 상황과 비교했을 때, 프로토콜에 명시된 것 이외의 추가 중재가 제공되었거나, 중재제공자가 의도한 대로 중재 실행을 실패하였거나, 연구대상자가 배정된 중재에 대해 비준수했다는 근거가 있는 경우, 연구대상자가 중재 중 하나에 특이한 부작용이나 독성이 생겨서 눈가림이 깨진 경우, 시험 프로토콜과 일관되지 않고 임상시험적인 맥락 때문에 발생한 중재 변경인 경우</p> <p>· PY: 임상시험 연구자가 임상시험적 맥락 때문에 이탈이 발생했을 가능성이 있는 경우</p>	<p>NA/Y/PY/ PN/N/NI</p>

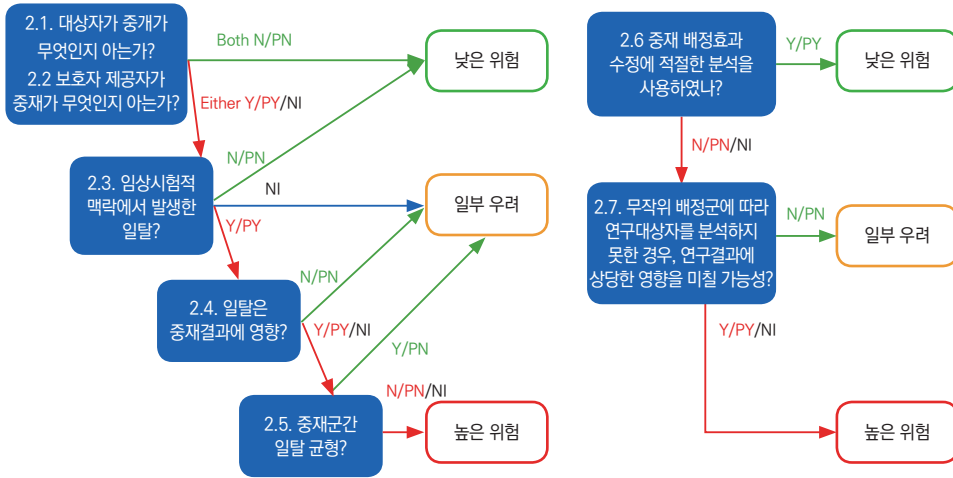
신호질문	세부 설명	응답 옵션
	<ul style="list-style-type: none"> · N/PN: 임상시험 프로토콜과 일치하는 중재 변경 예를 들어 급성 독성 때문에 약물 중재가 중단되거나 의도된 중재 중에 발생한 중재 결과를 처리하는 목적으로 추가 중재를 사용하는 것과 같은 임상시험적 맥락과 관련이 없는 경우 · NI: 임상시험 연구자가 임상시험적 맥락 때문에 이탤이 발생했는지 여부를 보고하지 않는 경우 	
2.4. 2.30이 Y/PY 인 경우 : 이러한 이탤로 중재결과에 영향을 미칠 것 같은가?	임상시험 프로토콜과 일관되지 않고 임상시험적 맥락 때문에 발생한 배정된 중재의 변경은 중재 결과에 영향을 미치는 경우에는 중요할 것이고 그렇지 않은 경우에는 중요하지 않을 것이다.	NA/Y/PY/ PN/N/NI
2.5. 2.40이 Y/PY/NI 인 경우 : 의도한 중재에서 이탤이 두 군에서 유사한가?	임상시험 프로토콜과 일관되지 않고 임상시험 맥락 때문에 발생한 배정된 중재의 변경은 중재군 간에 균형이 맞지 않는 경우 비뿔림을 초래할 가능성이 더 크다.	NA/Y/PY/ PN/N/NI
2.6. 중재 배정효과 추정에 적절한 분석을 사용하였나?	<ul style="list-style-type: none"> · Y/PY: 중재결과 자료가 결측된 연구대상자를 제외한 '배정된 대로 분석(ITT)'과 '수정된 ITT(mITT)분석', 부적격 연구대상자의 무작위 배정 후 배제(무작위 배정 이후까지 자격이 확인되지 않았고, 중재군 배정에 의해 영향을 받을 수 없었을 때) · N/PN: '계획서 순응군 분석'(naive 'per-protocol', 배정된 중재를 받지 않은 연구대상자는 제외)과 '치료한대로 분석(as-treated)'(배정된 중재가 아닌, 받은 중재에 따라 연구대상자가 그룹화되는 경우) 	Y/PY/PN/N/ NI
2.7. 2.60이 N/PN/NI 인 경우 : 무작위 배정군에 따라 연구대상자를 분석하지 못한 경우 (연구 결과에) 상당한 영향을 미칠 가능성이 있었는가?	이 질문은 잘못된 중재군에서 분석되었거나 분석에서 제외된 연구대상자의 수가 결과에 상당한 영향을 미칠 수 있을 정도로 충분하지 여부를 다룬다. 정확한 규칙을 지정할 수 없다. 연구대상자의 5% 미만이 잘못된 그룹에서 분석되거나 배제되었더라도, 중재 결과가 드물거나 예측 요인과 강하게 관련되는 경우라면 상당한 영향이 있을 수 있다.	NA/Y/PY/ PN/N/NI

비뿔림위험 평가	알고리즘 참조	낮음/높음/일부 우려
선택사항: 의도한 중재에서 이탤로 인한 비뿔림의 예측 방향은 어떠한가?	비뿔림의 방향을 예측할 수 있으면, 이를 기술하는 것이 도움이 된다. 방향은 무효점(혹은 무효점의 반대쪽)을 향한, 혹은 임상시험 중재선호, 비교중재 선호로 기술한다.	적용불가/임상시험 중재선호/비교 중재선호/무효점을 향한/무효점에 서 멀어짐/예측불가

(3) 의도한 중재에서 이탈로 인한 비밀취급위험 판단 알고리즘 (중재 배정 효과)

Part 1: 2.1~2.5 문항

Part 2: 2.6 & 2.7 문항



PART 2

(4) 의도한 중재에서 이탈로 인한 비뚤림위험 판단 예시 (중재배정효과)

신호질문	서술	응답
2.1 연구대상자는 임상시험 중에 배정된 중재에 대해 알고 있는가?	<u>Blinding was maintained by the double-dummy technique using two placebos identical to peramivir and oseltamivir.</u>	<u>N</u>
2.2 보호자와 중재 제공자는 임상시험 중에 연구 대상자에 배정된 중재에 대해 알고 있는가?		<u>N</u>
2.3 2.1과 2.2이 Y/PY/NI 인 경우 : 임상시험적 맥락 때문에 의도한 중재에서 이탈이 있었나?		NA
2.4. 2.3이 Y/PY 인 경우 : 이러한 이탈로 중재 결과에 영향을 미칠 것 같은가?		NA
2.5. 2.4이 Y/PY/NI 인 경우 : 의도한 중재에서 이탈이 두 군에서 유사한가?		NA
2.6. 중재 배정효과 추정에 적절한 분석을 사용하였나?	FIG. 1. Study profiles. *, One patient assigned to the 600-mgperamivir group received 300 mg peramivir. <u>This patient's data were included in the analyses according to the treatment actually administered.</u> Given the nature of the noninferiority study, <u>patients who were not treated as assigned were included in the analyses according to the actual treatment received.....</u> → 'As treated' analysis	PN
2.7. 2.6이 N/PN/NI 인 경우 : 무작위 배정군에 따라 연구대상자를 분석하지 못한 경우 (연구결과에) 상당한 영향을 미칠 가능성이 있었는가?	The study reported the reason for drop out and the drop out rate were relatively low (4 of the 1,093 participants).	<u>PN</u>
비뚤림위험 평가		일부우려

(5) 의도한 중재에서 이탈로 인한 비뚤림위험 평가 가이드 (중재 준수 효과)

신호질문	세부 설명	응답 옵션
2.1 연구대상자는 임상 시험 중에 배정된 중재에 대해 알고 있는가?	연구대상자가 배정된 중재를 알고 있다면, 건강 관련 행동이 중재군에 따라 다를 가능성이 크다. 연구대상자에 대한 눈가림은 위약이나 가짜 중재를 통해, 그러한 차이를 방지할 수 있다. · Y/PY: 연구대상자가 중재 중 하나에 특정한 것으로 알려진 부작용이나 독성을 경험한 경우	<u>Y/PY/</u> <u>PN/N/NI</u>

신호질문	세부 설명	응답 옵션
2.2 중재 제공자는 임상 시험 중에 연구대상자에 대한 배정된 중재에 대해 알고 있는가?	<p>중재 제공자가 배정된 중재를 알고 있다면, 그 중재를 시행했는지 비프로토콜 중재를 시행했는지 여부가 중재군 간에 다를 수 있다. 눈가림은 그러한 차이를 막을 수 있다.</p> <ul style="list-style-type: none"> · Y/PY: 중재제공자가 중재 중 하나에 특정한 것으로 알고 있는 부작용이나 독성을 연구대상자들이 경험한 경우 무작위 배정 은폐가 되지 않는 경우, 중재제공자는 임상시험 중에 연구대상자들에 배정된 중재가 무엇인지 알고 있었을 가능성이 크다. 	<p style="text-align: center;">Y/PY/ PN/N/NI</p>
2.3 [해당되는 경우] 2.1 과 2.2에 대해 Y/PY/NI 인 경우 : 중재군간 중요한 비프로토콜 중재가 유사한가?	<p>이 질문은 사전 고려사항에서 평가 과정에 중재군 간의 중요한 비프로토콜 개입의 불균형을 다룰 것이라고 표시한 경우에만 하도록 한다. 중요한 비프로토콜 중재는 (1) 임상시험 프로토콜과 일치하지 않는 추가 중재 또는 노출, (2) 연구대상자가 배정된 중재와 같이 혹은 그 이후에 받을 수 있는 추가 중재 또는 노출, (3) 중재결과에 대해 예후요인인 추가 중재 또는 노출이다. 중재군간에 중재에 불균형이 있을 경우 비뚫림위험이 더 높아질 것이다.</p>	<p style="text-align: center;">NA/Y/PY/ PN/N/NI</p>

예시 : 무작위 배정된대로 치료 받지 않은 상당수의 연구대상자

요추 추간판 탈출증 수술의 효과를 결정하기 위해, SPORT 임상시험 (Spine Patient Outcomes Research Trial) 은 요추 추간판 탈출증을 가진 환자를 무작위로 배정하여 수술 치료(절제술) 또는 비수술 치료(진통제, 교육, 물리 치료 및 침술 등을 포함한 다양한 중재 포함)를 받았다.¹²⁾ ITT 분석에서는 무작위 배정 후 2년 동안 1차 결과 (SF -36 신체 통증 및 신체 기능 척도)에서 중재군간에 차이가 있다는 증거를 발견하지 못했다. 그러나 외과 치료에 배정된 환자의 60% 만이 수술을 받았고 비수술군의 45%가 외과적으로 치료되었다. 저자는 '치료된 대로' 분석을 수행하여 수술의 이점을 보여주었다. 이는 수술 받은 연구대상자(지정된 중재에 관계없이)의 기본 특성이 그렇지 않은 사람들의 특성과 실질적으로 다르기 때문에 중재에 따른 영향(중재준수 효과)에 비뚫림이 있었다.

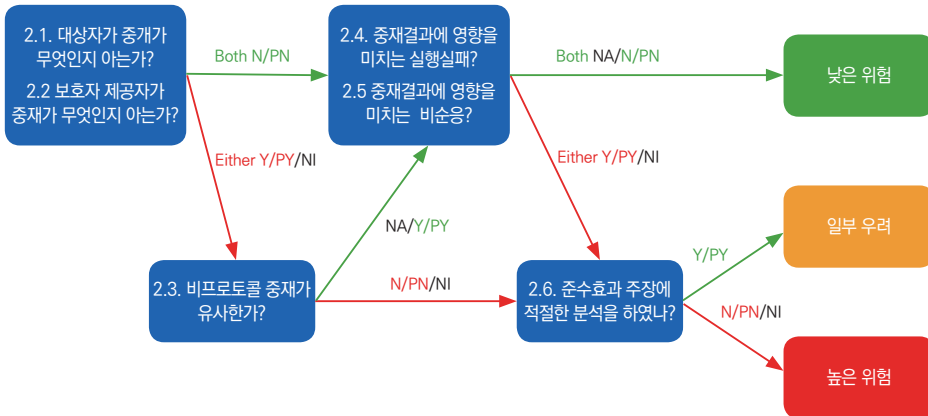
2.4. [해당되는 경우] 중재결과에 영향을 미칠 수 있는 중재 실행의 실패가 있는가?	<p>이 질문은 사전 고려사항에서 평가과정에 '중재 결과에 영향을 미칠 수 있는 중재의 실행 실패'를 다룰 것이라고 표시한 경우에만 하도록 한다.</p> <ul style="list-style-type: none"> · Y/PY: 치료를 제공하는 보건의료인에 의해 의도한 대로 중재가 실행되지 않은 경우 · N/PN: 모든 또는 대부분의 연구대상자가 중재 실행에 성공하였음 	<p style="text-align: center;">NA/Y/PY/ PN/N/NI</p>
--	--	---

예시 : 중재가 성공적으로 실행되지 않은 예

급성 충수염 환자를 표준 복강경 충수절제술 또는 단공 절개 복강경 충수절제술을 받도록 무작위 배정하였다.¹³⁾ 표준 복강경 수술은 3개의 구멍(2 개는 복강경기구 1개는 복강경 카메라)이 있는 단공 절제술은 카메라와 1개의 기구를 배치하는 단 하나의 구멍만 있다. 그러나 단공 절개술로 무작위 배정된 환자의 50% 미만이 의도한 대로 시술을 받았다. 추가기구와 구멍이 각각 51%와 10%에서 추가로 필요했다. ITT 분석에 따르면 상처 감염 위험(1차 결과), 정상 복귀 시간에서 중재군 간 차이가 없었다. 이런 분석은 중재 배정효과를 추정하는 데는 적합하지만, 중재준수에 대한 효과는 무효점에 가깝다.

신호질문	세부 설명	응답 옵션
<p>2.5. [해당되는 경우] 연구대상자의 중재 결과에 영향을 미칠만한 중재에 대한 비준수가 있는가?</p>	<p>이 질문은 사전 고려사항에서 평가 과정에 '연구대상자의 결과에 영향을 미칠 수 있는 중재에 대한 비준수'를 다룰 것이라고 표시한 경우에만 하도록 한다. 비준수에는 지속적인 중재에 대한 불완전한 이행, 중재 중단, 대조군 중재로 교차 이동, 다른 활성 중재로 전환 등이 포함된다.</p> <ul style="list-style-type: none"> · P/PY: 추적관찰 기간 동안 배정된 중재를 계속 받고 있는 연구대상자 비율에 대해 이용할 수 있는 정보를 고려하고, 비준수 비율이 문제를 제기할 정도로 높음 · N: 불완전한 준수가 불가능한 1회 투여 중재 연구나 모든 혹은 대부분의 연구대상자가 배정된 중재를 받은 경우 	<p>NA/Y/PY/ PN/N/NI</p>
<p>2.6. 2.3에 대해 N/PN/NI 혹은 2.4나 2.5에 대해 Y/PY/NI인 경우 : 중재 준수 효과 추정에 적절한 분석법을 사용하였나?</p>	<ul style="list-style-type: none"> · N/PN: (i) '배정된 대로 (ITT) 분석', (ii) '계획서 순응군 분석(PP)', (iii) '치료한 대로 분석' (iv) '받은 치료에 의한 분석('analysis by treatment received')' 이 수행된 경우, 중요한 비프로토콜 중재가 한 중재군의 모든 연구대상자에게 시행된 경우 · Y/PY: (1) 기저 상태에서만 단일 중재로 시행되고 준수가 전무 혹은 모두인 경우 표준 치료와 비교되는 임상 시험에서 중재의 영향을 추정하기 위한 도구변수분석(instrumental variable analyses), (2) 지속적인 치료 전략에 대한 임상시험에서 중도 절단된 연구대상자에 대해 보정해 주는 역확률 가중치(inverse probability weighting)를 이용하였고, 의도한 중재부터 이탈된 사람에 대한 정보가 있으며, 분석방법 자체가 적절하고 합당함 	<p>NA/Y/PY/ PN/N/NI</p>
비뚤림위험 평가	알고리즘 참조	낮음/높음/일부 우려
<p>선택사항: 의도한 중재에서 이탈로 인한 비뚤림의 예측 방향은 어떠한가?</p>	<p>비뚤림의 방향을 예측할 수 있으면, 이를 기술하는 것이 도움이 된다. 방향은 무효점(혹은 무효점의 반대쪽)을 향한, 혹은 임상시험 중재선호, 비교중재 선호로 기술한다.</p>	<p>적용불가 / 임상시험 중재선호 / 비교중재선호/무효점을 향한/ 무효점에서 멀어짐/예측불가</p>

(6) 의도한 중재에서 이탈로 인한 비뚤림위험 판단 알고리즘 (중재 준수 효과)



차) 영역 3. 중재결과 자료의 결측으로 인한 비뚤림

(1) 일러두기

- 이 영역은 중재결과 자료의 결측 또는 이를 해결하기 위한 결측값 대처법 적용 절차에 의해 발생한 비뚤림의 위험을 다룬다. 중재결과 자료가 결측가능한 이유는 다음과 같다.
 - 연구대상자가 연구 참여를 철회하거나 추적 불가능함 또는 탈락
 - 연구대상자가 결과를 측정해야 하는 연구 방문에 오지 않음 (추적관찰 실패, 탈락)
 - 연구대상자가 연구 방문했지만 관련 자료는 제공하지 않음
 - 자료가 소실되었거나 다른 이유로 사용할 수 없는 경우
 - 연구대상자에게 더 이상 중재결과가 발생할 수 없음 (예: 사망)
- 일부 연구대상자는 중재결과 자료 결측 이외의 이유로 분석에서 제외될 수 있다. 특히 '계획서 순응군 (per protocol, PP)' 분석은 의도한 중재를 받은 연구대상자로 제한된다. 그러한 분석에 의해 발생하는 비뚤림위험이나 중재결과 자료를 이용할 수 있는 적격 연구대상자가 다른 이유로 배제되어 발생하는 잠재적 비뚤림은 '의도한 중재에서 이탈로 인해 발생하는 비뚤림' 영역에서 다루어지며, 최종 신호 질문은 분석 접근법이 적절한지 여부를 검토한다.
- 중재 배정 효과를 추정할 때 중재결과 자료의 결측: 중재 배정 효과는 '배정된 대로 (intended to treat, ITT)' 분석의 기본 원칙은 다음과 같다.
 - (1) 실제 받은 중재에 관계없이 무작위 배정된 중재 그룹의 모든 연구대상자를 분석한다.
 - (2) 분석에는 모든 무작위 연구대상자를 포함하며, 이를 위해서는 모든 연구대상자의 중재 결과를 측정해야 한다.

- 중재 준수 효과 추정에 중재결과 자료의 결측: 연구대상자가 중재결과 자료 결측으로 인해 배제되는 경우, 중재에 따른 효과를 추정하기 위한 적절한 방법(예를 들어, 도구를 이용해 중재의 효과 변수를 추정하는 분석)조차도 비뚤림이 있는 추정치를 산출할 수 있다.
- 모든 또는 거의 모든 무작위 연구대상자의 중재결과 자료가 제공된다면 비뚤림의 위험은 낮을 것이다. 맥락상 '거의 모두'의 의미란 중재결과 자료가 결측된 연구대상자의 수가 너무 적기 때문에 결과가 어쨌든 간에 중재의 추정 효과에 중요한 차이가 없다는 것이다. 만약 그렇다면, 더 이상의 신호 질문에 답할 필요가 없다. 중재결과 자료가 결측된 정도에 대한 정보(예를 들어, 시험 보고서에 CONSORT 흐름도가 제공되지 않은 경우)가 없으면 이 영역에 대한 비뚤림위험이 높다고 판단하게 된다.
- 임상시험 연구자 또는 문헌고찰 저자에 의해 수행된 민감도 분석에 의해 결과가 중재결과 자료 결측치에 대한 개연성있는 가정들에서, 연구결과가 거의 변하지 않는다는 것을 확인한다면 비뚤림의 위험은 낮을 것이다. 이러한 분석은 결측된 자료의 양이 중재의 추정 효과에 대한 잠재적 영향이 상당할 정도로 충분히 큰 경우에 특히 유용할 수 있다. 민감도 분석에 의해 결과가 견고하다는 것이 확인되면, 그 결과는 비뚤림의 위험이 낮은 것으로 간주될 수 있으며 더 이상의 신호 질문에 답할 필요가 없다.
- 중재 결과의 결측 여부가 중재결과와 참값과 무관한 경우, 각 중재군 내에서 중재결과 자료가 결측되면 비뚤림이 생기지 않는다. 그러나 (1) 중재결과와 결측 여부가 중재군과 중재결과와 참값 모두에 따라 다를 경우, (2) 중재결과와 결측 여부는 참값과 관련이 있고 임상시험군 중재의 효과가 대조군 중재의 결과와 다를 경우 중재결과 자료가 결측되면 비뚤림이 발생한다.
- 중재결과 자료 결측의 비율의 차이: 실험 중재군 및 대조 중재군에서 중재결과 자료 결측의 비율이 다르면 잠재적 비뚤림의 근거를 제공한다. 그러나, 중재결과 자료 결측의 비율이 중재군 간에 달라도 중재결과 자료가 결측되는 것이 비뚤림을 초래하지 않을 수 있는데, 이는 결과가 결측될 가능성이 참값과 관련이 없는 경우에만 해당한다.
- 중재결과 자료가 결측된 이유 검토: 비뚤림으로 이어질 가능성이 가장 높은 상황은 중재결과 자료가 결측된 이유가 중재군 간에 다를 때, 예를 들어 심각한 상태가 된 연구대상자들이 대조군에서 탈락하는 동안 중재군에서 호전되어 탈락하는 경우이다.
- 사건발생 시간 자료의 경우, 문헌고찰 저자는 중도절단 비율의 차이와 중도절단 이유의 차이가 비교군간에 다른지 여부를 고려해야 한다.
- 중재결과 자료 결측을 처리하는 통계적 방법: 단일 결측값 대체방법 ('최종 관찰 이월'이나 '기저상태 관찰 이월')은 결과의 결측이 결과의 참값과 관련 있다면, 비뚤림을 제거하지 못할 가능성이 높다. 다중 결측값 대체법은 중재결과와 결측이 참값에 영향을 미치지 않는다는 가정에 대한 합당한 이유를 제시했을 경우에만 비뚤림위험이 낮을 수 있다.

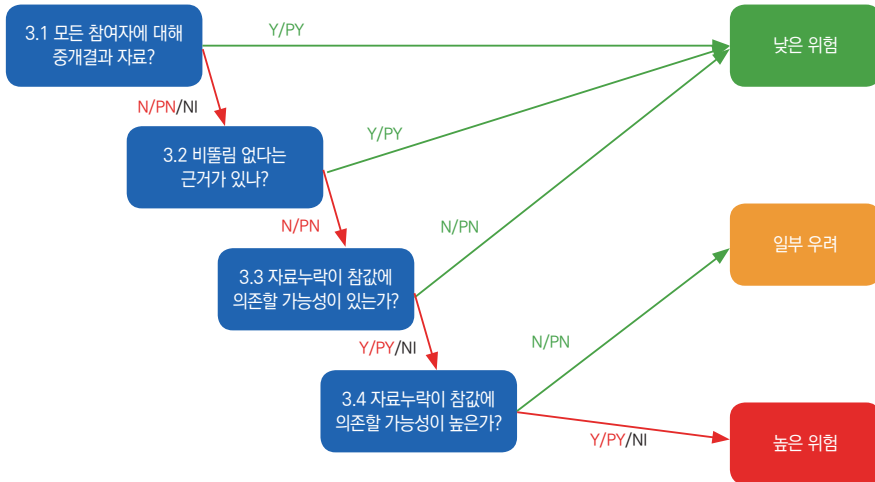
(2) 결측지로 인한 비뚫림위험 평가 가이드

신호질문	세부 설명	응답 옵션
<p>3.1 현재 중재결과에 대해 모든 혹은 거의 모든 연구대상자의 자료가 이용가능한가?</p>	<p>중재 배정 효과 분석(ITT effect analysis)을 위해서는 무작위 임상시험 연구대상자 전부가 필요하다.</p> <ul style="list-style-type: none"> • PY: 결측치가 있는 연구대상자의 수가 충분히 적기 때문에 결과가 어땠든 간에 중재의 추정 효과에 중요한 차이가 없는 것으로 해석해야 한다. 연속변수 중재결과의 경우, 연구대상자의 95%에서 자료를 이용할 수 있으면 충분하다고 할 수 있을 것이다. 이분변수 중재 결과의 경우, 요구되는 비율은 사건의 위험과 직접적으로 연관된다. 관찰된 사건 수가 결측치가 있는 연구대상자 수보다 훨씬 많은 경우, 비뚫림은 매우 적을 것이다. • NI: 임상시험 보고서에 결측치의 범위에 대한 정보를 제공하지 않는 경우에만 '정보 없음'으로 응답한다. 이러한 상황에서는 대개 '결측치로 인한 비뚫림위험'을 높다고 판단하게 될 것이다. • 대처된 결측자료는 결측치로 간주되어야 하며, 이 질문의 맥락에서는 '중재결과 자료'로 간주되지 않아야 한다. 	<p>Y/PY/ PN/N/NI</p>
<p>3.2 3.1에 대해 N/PN/NI인 경우: 결측치로 인해 연구결과의 비뚫림이 없다는 근거가 있는가?</p>	<p>(1) 결측치로 인해 결과가 편향되지 않았다는 증거는 (1) 비뚫림을 교정하는 분석 방법, 또는 (2) 결측 여부와 그 참값 사이의 관계에 대한 그럴듯한 가정 하에 연구결과가 거의 변하지 않는다는 것을 보여주는 민감도 분석에서 비롯될 수 있다.</p> <p>(2) 그러나 '최종관찰 이월'과 같은 방법 또는 중재군에만 복수의 결측치 대처를 결측치로 인한 비뚫림위험을 교정하는 것으로 가정해서는 안 된다.</p>	<p>NA/Y/PY/ PN/N</p>
<p>3.3 3.2에 대해 N/PN인 경우: 중재 결과의 결측여부가 참값에 의존할 가능성이 있는가?</p>	<p>(1) 추적관찰 실패 또는 연구 참여 철회는 연구대상자의 건강 상태와 관련될 수 있다면, 중재 결과의 결측여부는 참값에 의해 영향을 받을 수 있다. 그러나 모든 결측치가 중재 결과와 무관하다는 것이 문서화 되어 있으면, 결측치로 인한 비뚫림위험은 낮을 것이다.</p> <p>(2) 사건발생시간 분석에서, 예를 들어 연구 참여 철회로 인해 추적관찰이 중단된 경우는 일부 추적 관찰 자료가 분석에 포함되더라도 결측치로 간주해야 한다. 이러한 연구대상자는 CONSORT 흐름도에는 분석에 포함된 것으로 표시된다.</p>	<p>NA/Y/PY/ PN/N/NI</p>
<p>예시 : <u>비뚫림위험 낮음</u> 측정 장치의 고장 또는 일상적인 데이터 수집 중단</p>		
<p>3.4 3.3에 대해 Y/PY/NI: 중재 결과의 결측여부가 참값에 의존할 가능성이 높은가?</p>	<p>이 질문은 (i) 결측여부가 참값에 좌우될 가능성이 있는 상황('일부 우려'로 평가됨)과 (ii) 자료 누락이 참값에 좌우될 가능성이 높은 상황('비뚫림위험'이 높음으로 평가됨)을 구별한다. '예'라고 대답하는 다섯 가지 이유는 다음과 같다.</p>	<p>NA/Y/PY/ PN/N/NI</p>

신호질문	세부 설명	응답 옵션
	<ul style="list-style-type: none"> • Y/PY: <ol style="list-style-type: none"> 1) 중재군간 결측치 비율의 차이. 실험 중재와 대조 중재가 중재결과에 미치는 영향이 차이가 있고, 결측여부가 참값에 영향을 받으면, 결측치의 비율이 중재군간에 다를 수 있다. 이러한 차이는 결측치로 인한 비뚤림위험을 시사한다. 왜냐하면 시험 결과는 참값에 따른 중재결과의 누락에 민감할 것이기 때문이다. 사건 발생시간 분석의 경우, 중도절단 비율(추적관찰 실패)이 중재군마다 다르다는 것이 유사한 점이다. 2) 결측치 이유로 보고된 내용에서 결측여부가 참값에 의존한다는 근거를 제공한다. 3) 결측치 이유로 보고된 내용이 군간에 다르다. 4) 임상시험의 정황으로 보아 결측여부는 참값에 의존할 가능성이 높다. 예를 들어, 정신분열증을 치료하기 위한 중재 임상시험에서 증상이 지속되는 경우 탈락이 더 많다는 사실은 널리 알려져 있다. 5) 사건발생시간 분석에서, 연구대상자가 약물 독성 때문에 또는 암 임상 시험에서 2차 화학 요법으로 전환할 때, 배정된 중재를 중지하거나 변경하는 경우 중단이 발생하는 경우 • N/PN: <p>결측여부와 참값 사이의 관계를 설명하는 것이 연구대상자 특성인 경우</p> 	

비뚤림위험 평가	알고리즘 참조	낮음/높음/일부 우려
<p>선택사항: 중재결과 자료의 결측으로 인한 비뚤림의 예측 방향은 어떠한가?</p>	<p>비뚤림의 방향을 예측할 수 있으면, 이를 기술하는 것이 도움이 된다. 방향은 무효점(혹은 무효점의 반대쪽)을 향한, 혹은 실험 중재선호, 비교중재 선호로 기술한다.</p>	<p>적용불가 / 실험 중재선호 / 비교중재선호 / 무효점을 향한 / 무효점에서 멀어짐 / 예측불가</p>

(3) 중재결과 자료의 결측으로 인한 비뚫림위험 판단 알고리즘



(4) 중재결과 자료의 결측으로 인한 비뚫림위험 판단의 예시

신호질문	서술	응답
3.1 현재 중재결과에 대해 모든 혹은 거의 모든 연구대상자의 자료가 이용가능한가?	Six patients who dropped out before treatment and two patients with no posttreatment efficacy data were excluded from all analyses. One patient who was randomized to the 600 mg peramivir group mistakenly received 300 mg peramivir and was thus included in the 300 mg peramivir group. Therefore, 1,091 patients (364 receiving 300 mg peramivir, 362 receiving 600 mg peramivir, and 365 receiving oseltamivir) were included in the ITTI population, the primary efficacy analysis population.	Y
3.2 3.1에 대해 N/PN/NI인 경우: 결측치로 인해 연구결과의 비뚫림이 없다는 근거가 있는가?		NA
3.3 3.2에 대해 N/PN인 경우: 중재 결과의 결측여부가 참값에 의존할 가능성이 있는가?		NA
3.4 3.3에 대해 Y/PY/NI: 중재 결과의 결측여부가 참값에 의존할 가능성이 높은가?		NA
비뚫림위험 평가		낮은위험

카) 영역 4. 중재결과 측정의 비뚤림

(1) 일러두기

- 연구대상자의 중재결과 변수 측정 오류는 측정값이 참값 또는 기초값(underlying value)과 같지 않을 때 발생한다. 이 같은 오류는 무작위 임상시험의 중재 효과 추정에 비뚤림을 유발할 수 있다. 이러한 오류는 연속형 결과의 경우 측정 오류, 이분형 또는 범주형 결과의 경우 오분류, 사건인 경우 과소확인/과다확인(under/over-ascertainment)으로 표현된다.
- 차등적 측정 오류는 중재 배정과 관련이 있다. 그러한 오류는 실험군과 대조군 간의 체계적 차이로 나타나며, 결과 평가자가 중재 배정에 눈가림되어 있다면 발생 가능성이 낮다. 이 영역은 차등적 오류와 관련이 있다.
- 비차등적 측정 오류는 중재 배정과 관련이 없는 오류이다. 중재 배정과 무관한 중재결과의 측정 또는 분류 오류를 '비차등적' 오류라고 한다. 그것은 일반적으로 중재군간 연속 변수의 평균 차이를 보는 중재 효과 추정치에 비뚤림을 일으키지는 않지만, 중재결과가 이분 변수이거나 범주 변수일 때 오즈비, 위험비, 상대위험비 등과 같은 중재 효과 추정치를 감소시킬 가능성(즉 무효점을 향하게 하는 비뚤림)이 높다. 비차등적 측정 오류가 효과 추정치를 무효점에서 벗어나게 하는 방향일 상황이지만, 이러한 상황은 보통 무작위 시험에서는 발생할 가능성이 낮은 것으로 간주된다.
- **연구대상자 보고 중재결과:** 연구대상자 보고 중재결과는 임상 의사 또는 연구자가 연구대상자의 반응을 해석하지 않고 건강 상태 및 특정 치료법과 관련하여 연구대상자가 어떻게 효과를 보았고 느끼는지에 대해 연구대상자가 직접 중재결과를 보고하는 것이다. 예로는 고통, 메스꺼움, 그리고 건강 관련 삶의 질이 있다.
- **판단과 무관한 관찰자 보고 중재결과:** 이러한 중재결과는 관찰자의 판단을 수반하지 않는 외부 관찰자(예: 영상의학과 전문의 같이 연구대상자에게 제공되는 치료에 관여하지 않는 독립적 연구자 또는 의사)가 보고한 결과물이다. 모든 원인 사망률 또는 자동화된 시험의 중재결과를 예로 들 수 있다. 여기서의 중재결과 평가자는 관찰자이다. 판단을 수반하지 않는 관찰자 보고 중재결과의 경우, 중재결과 평가는 일반적으로 수행된 중재 지식에 의해 영향을 받지 않을 가능성이 높다.
- **일부 판단이 수반되는 관찰자 보고 중재결과:** 이것들은 임상 검사에 관여하는 것과 같이 일부 판단을 수반하는 외부 관찰자(예: 중재 제공자)에 의해 보고된 중재결과들이다. 예를 들면, 방사선 촬영, 임상적 관찰, 사망을 제외한 임상적 사건(예: 심근경색) 등 의무기록에 기초한 판단을 포함하는 임상 시험이 있다. 여기서의 중재결과 평가자는 관찰자이다. 관찰자가 수행된 중재를 인지하는 경우, 중재결과 평가는 잠재적으로 이 지식에 의해 영향을 받을 수 있어 최소한 '일부 우려'로 판단하게 된다. 문헌고찰 저자는 중재결과 평가가 수행된 중재의 지식에 의해 영향을 받았을 가능성이 있는지 판단해야 하며, 이 경우 비뚤림의 위험이 높다고 간주된다.
- **중재 제공자가 내린 결정을 반영하는 중재결과:** 이러한 중재결과는 중재 제공자의 결정을 반영한다. 이 결정의 기록에는 어떠한 판단도 수반되지 않는다. 그러나 그 결정 자체는 받은 중재에 대한 지식에 의해 영향을 받을 수 있다. 중재 제공자 의사결정 중재결과와 일반적인 예로는 입원, 치료 중단, 타 병동 이동, 제왕절개 수행, 기계환기 중단, 퇴원 등이 있다. 여기서의 중재결과 평가자는 의사 결정권을 가진 의료 제공자다. 의료 제공자가 이러한 사실을 알고 있다면, 중재결과 평가는 일반적으로 수행된 중재에 대한 지식에 의해 영향을 받을 가능성이 높다. 이는 특히 임상시험 중재의 효과에 관한 선호도, 기대치 또는 예감이 강할 때 중요하다.

• **복합 중재결과:** 복합 중재결과는 여러 개의 중재결과를 하나의 중재결과로 결합한다. 일반적으로 특정 중재결과의 집합 중 하나를 경험한 연구대상자는 복합 중재결과를 경험한 것으로 간주한다. 대표적인 예로는 주요 심장 및 뇌혈관 유해사건(major adverse cardiac and cerebrovascular events, MACCE)이 있다. 복합 중재결과는 연속형 변수 측정으로도 구성할 수 있다. 복합 중재결과에 대한 비틀림위험 평가는 복합 중재결과의 각 구성요소의 빈도나 기여도를 고려해야 하며 가장 영향력 있는 구성요소로 인한 비틀림위험을 고려해야 한다.

(2) 결과자료 측정에서 발생하는 비틀림위험 평가 가이드

신호질문	세부 설명	응답 옵션
4.1 중재결과 측정 방법이 부적절한가?	<p>이 질문은 평가하고자 하는 중재 결과에 적합하지 않은 결과 측정 방법(데이터 수집)을 확인하는 것이 목적이다. 이 질문은 평가되는 중재 결과의 선택(예: 관심 주요 결과에 대한 대리결과 또는 중간결과)이 합리적이었는지를 평가하기 위한 것이 아니다.</p> <ul style="list-style-type: none"> • N/PN: 대부분의 경우 사전 지정된 중재 결과 측정방법이 있음 • Y/PY: <ul style="list-style-type: none"> (1) 중재효과를 확인하기 어려운 측정방법 사용(예: 중재 결과 값의 중요한 범위가 측정 방법을 사용하여 확인 가능한 수준을 벗어나는 경우) (2) 측정기기의 타당도가 불량하다는 것이 입증된 경우 	<p>Y/PY/ PN/N/NI</p>

예시 : Y/ PY

(1) 그럴듯한 중재 효과에 민감할 것 같지 않음(예: 중재 결과 값의 중요한 범위가 측정 방법을 사용하여 감지 가능한 수준을 벗어나는 경우)

(2) 측정기기의 타당도가 불량하다는 것이 입증된 경우

4.2 중재군간 중재결과 측정 혹은 확인 방법이 다른가?	<p>비교 가능한 중재결과 측정 방법(자료 수집)이란 유사한 시점에 동일한 측정 방법과 임계치로 측정하는 것이다. 중재군 간의 차이가 발생하는 경우는 중재결과 자료의 수동적 수집이라는 맥락에서 '진단 결과 확인 비틀림' 때문이거나, 중재에 의료 제공자를 추가로 방문하는 것이 포함되어서 중재결과 사건을 확인할 수 있는 기회가 추가되는 경우이다.</p>	<p>Y/PY/ PN/N/NI</p>
4.3 4.1과 4.2에 대해 N/PN/NI 인 경우: 중재 결과 평가자는 연구대상자가 받은 중재를 알고 있는가?	<p>중재 결과 평가자가 중재 상태에 눈가림되어 있으면 'N'로 응답한다. 연구대상자 보고 결과의 경우, 중재 결과 평가자는 연구대상자이다.</p>	<p>NA/Y/PY/ PN/N/NI</p>

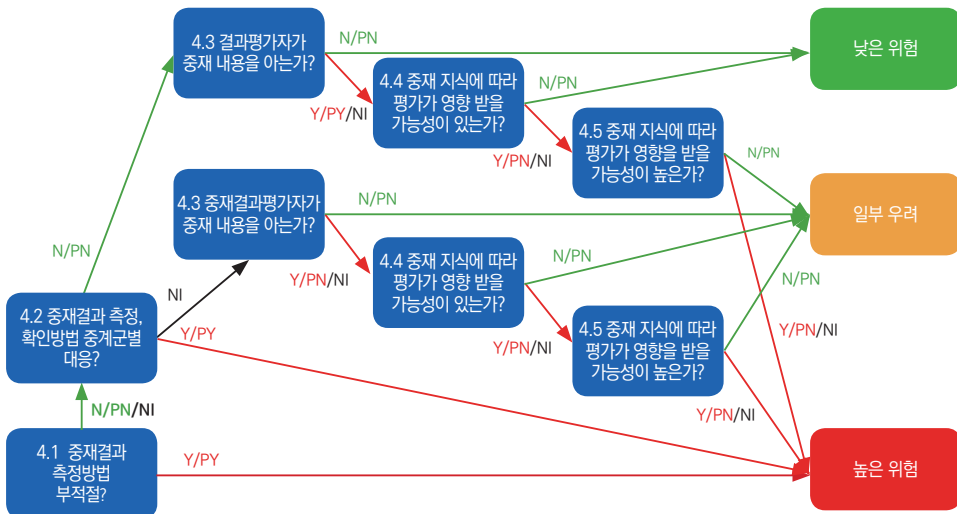
신호질문	세부 설명	응답 옵션
4.4 4.3에 대해 Y/PY/NI 인 경우: 중재결과에 대한 평가가 중재에 대한 지식에 의해 영향을 받을 가능성이 있는가?	배정된 중재에 대한 지식은 연구대상자 보고 결과(통증 수준 등), 일부 판단이 포함된 관찰자 보고 결과 및 중재 제공자의 의사결정 결과에 영향을 미칠 수 있다. 판단을 수반하지 않는 관찰자가 보고한 중재 결과(예를 들어, 모든 원인 사망률과 같은)에 영향을 미칠 가능성은 낮다.	NA/Y/PY/ <u>PN/N/NI</u>
4.5 4.4에 대해 Y/PY/NI 인 경우: 중재결과에 대한 평가가 중재에 대한 지식에 영향을 받았을 가능성이 높은가?	이 질문은 (i) 중재 상태에 대한 지식이 결과 평가에 영향을 미칠 수 있었지만 그렇게 했다고 믿을 이유는 없는 상황('일부 우려'로 평가)과 (ii) 중재 상태에 대한 지식이 결과 평가에 영향을 미칠 가능성이 있는 상황('높음'으로 평가됨)을 구별한다. 중재가 유익할거라는 혹은 유해할 거라는 믿음이 강할 때, 중재 결과가 받은 중재에 대한 지식에 영향을 받았을 가능성이 더 높다.	NA/Y/PY/ <u>PN/N/NI</u>

예시

Y/ PY: 동종요법 임상시험에서 환자 보고증상이나, 중재를 제공한 물리치료사가 기능회복에 대한 평가를 시행하는 경우 등이 예가 될 수 있다.

비뿔림위험 평가	알고리즘 참조	낮음/높음/일부 우려
선택사항: 중재결과 측정의 비뿔림의 예측 방향은 어떠한가?	비뿔림의 방향을 예측할 수 있으면, 이를 기술하는 것이 도움이 된다. 방향은 무효점(혹은 무효점의 반대쪽)을 향한, 혹은 실험 중재선호, 비교중재 선호로 기술한다.	적용불가 / 실험 중재선호 /비교 중재선호/무효점을 향한/ 무효 점에서 멀어짐/예측불가

(3) 결과자료 측정의 비뿔림위험 판단 알고리즘



(4) 결과자료 측정의 비뚫림위험 판단의 예시

신호질문	서술	응답
4.1 중재결과 측정 방법이 부적절한가?	Patients assessed their own influenza symptoms and daily living activities using an Influenza Symptom Severity scale (ISS) and a visual analogue scale (Influenza Impact Well-Being Score) ranging from 0 (unable to perform one's usual activities at all) to 10 (able to perform all usual activities fully).	<u>N</u>
4.2 중재군간 중재결과 측정 혹은 확인 방법이 다른가?	The same measurement method between groups were used.	<u>N</u>
4.3 4.1과 4.2에 대해 N/PN/NI 인 경우: 중재결과 평가자는 연구대상자가 받은 중재를 알고 있는가?	Participant-reported outcomes: Patients assessed their own influenza symptoms and daily living activities using an Influenza Symptom Severity scale. however, participants and trial staff was double-blinded, and the blindness might be maintained.	<u>PN</u>
4.4 4.3에 대해 Y/PY/NI 인 경우: 중재결과에 대한 평가가 중재에 대한 지식에 의해 영향을 받을 가능성이 있는가?		NA
4.5 4.4에 대해 Y/PY/NI 인 경우: 중재결과에 대한 평가가 중재에 대한 지식에 영향을 받았을 가능성이 높은가?		NA
비뚫림위험 평가		낮음

타) 영역5. 보고된 연구결과 선택의 비뚤림

(1) 일러두기

- 이 영역은 임상시험 연구자에 의해 계산된 여러 증재 효과 추정치 중에서 선택된(방향, 크기 또는 통계적 유의성에 기초하여) 연구결과 보고 때문에 발생하는 비뚤림을 다룬다. 우리는 이를 '보고된 연구결과 선택의 비뚤림'이라 부른다. 비뚤림위험의 고려사항으로 다음과 같은 구분이 필요하다.
 - **증재결과 영역.** 이는 관심 상태 또는 종결점이다. 어떻게 측정되는지(예 : 우울증의 중증도)와는 관계없다
 - **증재결과 측정.** 이는 결과 영역을 측정하는 구체적인 방법이다(예: 증재 6 주 후 Hamilton 등급 척도를 사용한 우울증 측정)
 - **증재결과 분석.** 이는 하나 이상의 증재결과 측정을 분석하여 얻어진 특정 연구결과이다(예: 시험군과 비교군 간의 기저 치에서 6 주후의 Hamilton 등급 척도 점수의 평균 변화량의 차이).
- 증재결과 영역, 측정, 분석의 차이에 대해, 보고된 결과를 선택할 때 비뚤림위험이 결과 누락으로 인한 비뚤림위험과 어떻게 다른지 보여준다. 이 가상의 예에서, 우울증과 불안의 두 가지 결과 영역이 관심 대상이다. 우울증은 HAM-D (Hamilton Depression Rating Scale)와 수정된 버전(mHAM-D)의 두 가지 척도를 사용하여 측정된다. 이러한 결과는 최종 값과 변경 값(최종값 - 기저치)을 모두 사용하여 분석된다. 우울증에 대해 보고된 효과 추정치(임상시험군과 비교군 간의 평균 mHAM-D 변화 값의 차이)는 임상시험적 증재에 가장 유리했기 때문에 임상시험 연구자에 의해 선택되었다. 따라서, 우울증에 대한 임상시험으로부터 이용 가능한 결과는 이 RoB 2 영역에서 보고된 연구결과 선택의 비뚤림위험이 있다. 대조적으로, 임상시험연구자들에 의해 불안이 측정되고 분석되었지만, 효과 추정이 임상시험적 증재에 유리하지 않기 때문에 자료가 보고되지 않았다. 이러한 비보고는 불안에 대한 연구결과와 메타 분석에서 비뚤림을 유발할 것이지만, 그 비보고 자체만으로는 보고된 우울증 효과 추정치나 동일한 임상시험의 다른 증재결과를 비뚤림위험에 처하게 하지는 않는다. '보고된 연구결과 선택'과 '선택적 비보고 영역'에 대한 분리는, Cochrane RoB 1 버전에서 특징적으로 달라진 것이다.
- 이 영역은 (i) 증재결과 영역 내에서 평가된 다수의 측정치로부터 특정 증재결과 측정의 선택적 보고 및 (ii) 특정 증재결과 측정의 다중 분석으로부터 특정 분석의 선택적 보고를 둘 다 고려한다. 선택이 효과 추정치의 방향, 크기 또는 통계적 유의성에 기초한 경우, 선택보고 유형 중 하나가 비뚤림으로 이어질 수 있다. 특정 증재결과 측정에 대한 선택적 보고는 여러 결과에 대한 효과 추정치 중에서 문헌고찰 저자가 관심을 가질 수 있는 증재결과 영역 내의 보고된 효과 추정치를(연구결과에 기초하여) 선택할 때 발생한다. 예를 들면 다음과 같다.
 - 증재결과가 측정된 시점 중 하나 또는 일부만을 보고하는 것(예를 들어, 6 주 및 8 주에 증재결과를 측정 했음에도 불구하고 3주 효과만 보고)
 - 측정 도구 중 하나 또는 일부만 보고(예 : 통증 척도)
 - 증재결과 평가자 중 하나 또는 일부만을 보고하는 것(예 : 환자 평가, 임상가 평가 척도는 제외)
 - 측정 도구 중 하나 또는 하위척도 일부만 보고
- 문헌고찰 저자는 사전 지칭된 분석 의도를 검색해 보는 것을 강력하게 추천한다. 그렇게 하면 임상시험 보고서에서 생략되거나 추가된 증재결과를 식별할 수 있다. 임상시험 레지스트리 항목(예: ClinicalTrials.gov 기록), 임상임상시험 프로토콜 또는 설계 문서(저널에 게시되거나 임상시험 편

드 웹 사이트를 통해 제공될 수 있음) 등을 포함하는 다양한 출처로부터 분석 의도가 확인 될 수 있다. 통계 분석 계획(SAP)은 종종 눈가림 되지 않은 결과 자료가 이용 가능하기 전에 사전 지정된 계획이 완료되었는지 여부에 대한 가장 자세한 정보를 제공하지만, 공개되지 않을 수 있다. 연구자의 사전 지정된 의도가 충분히 상세하게 이용가능하게 되면 계획된 중재결과 측정 및 분석을 발표된 보고서에 제시된 결과와 비교할 수 있다.

- 일부 임상시험의 경우 분석 의도를 쉽게 이용가능하지 않다. 그럼에도 불구하고 보고된 연구 선택의 비몰림위험을 평가하는 것은 가능하다. 예를 들어, 논문의 방법 섹션에 나열된 중재결과 측정 및 분석을 보고된 것과 비교할 수 있다. 또한, 중재결과 측정 및 분석을 임상시험 및 그 연구결과를 설명하는 여러 출처에서 비교할 수 있다. 또한 다음과 같은 질문을 하면 문헌고찰 저자가 선택적 보고를 추천하는 데 도움이 될 수 있다.
 - (1) 하위척도(subscales)가 통상적이지 않은 방식으로 종합되었나?
 - (2) 일차 및 이차 결과 또는 특정 결과의 지정에 있어 보고서 간에 불일치가 있나?
 - (3) 다수의 보정된 분석이 수행되었지만 단 하나(또는 하위세트만) 보고되었나? 하나 이상의 보정된 분석이 수행되었으나 보고된 바가 없었나?
 - (4) 연구자들은 연속형 결과 측정을 통상적이지 않은 방식으로 범주화했나? 같은 연구와 관련하여 여러 출판물에서 범주를 만들기 위한 다양한 절단점이 보고되었나?
 - (5) 분석된 연구대상자 표본에서 다른 보고서 간에 불일치가 있었나?
 - (6) 통상적이지 않은 복합 중재결과가 보고되었나? 예를 들어 “주요 부작용 사례” 또는 “경미한 부작용 사례” 범주에서 예상치 못한 부작용 사례의 다른 조합을 그룹화하는 것처럼 복합 중재결과에 대한 서로 다른 정의를 고려할 수 있었나?

(2) 보고된 연구결과 선택의 비몰림위험 평가 가이드

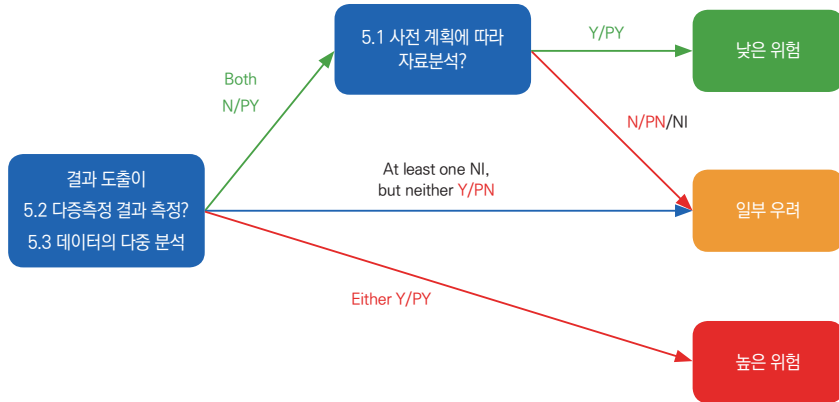
신호질문	세부설명	응답 옵션
5.1 이 연구결과를 도출한 자료는 분석할 수 있도록 미리 지정된 분석 계획에 따라 중재 결과 자료 눈가림이 해제되기 전에 분석되었는가?	(1) 사전에 지정된 연구자의 의도를 충분히 정도로 상세히 이용할 수 있는 경우, 계획된 중재 결과 측정과 보고서에 제시된 것을 비교할 수 있다. 보고된 결과의 선택 가능성을 피하기 위해, 분석 의도가 최종화되는 시점이 자료의 눈가림이 해제되는 시점보다 선행되어야 한다. (2) 자료의 눈가림이 풀리기 전에 분석계획이 변화한 경우 혹은 그러한 변화가 결과와 무관한 경우 (예를 들어, 고장난 기계로 인해 자료수집이 불가능함) 보고 결과의 선택에 대한 비몰림 우려가 제기되지 않는다.	Y/PY/ PN/N/NI
평가되고 있는 수치형 연구결과는 다음과 같은 연구결과에 기초하여 선택되었을 가능성이 있는가?		

신호질문	세부설명	응답 옵션
<p>5.2 중재 결과 영역 내에서 여러 개의 적합한 중재 연구결과 측정(예: 척도, 정의, 시점)?</p>	<p>(1) 특정 중재 결과 영역(즉, 실제 상태 또는 관심 종결점)은 여러 가지 방법으로 측정될 수 있다. 예를 들어, 통증 영역은 여러 척도(예: 시각적 아날로그 척도 및 맥길 통증 설문도구)를 사용하여 각각 여러 시점(예: 치료 후 3, 6, 12주)에서 측정할 수 있다. 복수의 측정이 이루어졌지만 결과(예: 통계적 유의성)에 근거하여 1개 또는 하나의 하위세트만 보고되는 경우, 비뿔림위험은 높을 것이다.</p> <p>(2) RoB 2 도구 사용자가 고려하기에 적격인 중재결과 만 주의를 기울이면 된다. 예를 들어 특정 측정 척도를 사용한 결과만 메타 분석(예: 해밀턴 우울증 평가 척도)에 포함할 수 있고, 임상시험에 이 중재결과가 보고된 경우, 다른 측정 척도에서 나온 결과(예: Beck depression Inventory)보다 우선적으로 이 결과가 보고되었다 하더라도(통계적 의미성 등에 의하여 했다고 하여도) 선택 문제는 없을 것이다.</p> <ul style="list-style-type: none"> • Y/PY: 한 영역을 복수의 적격 방법으로 측정했다는 명확한 근거(보통 임상시험 프로토콜이나 통계 분석 계획서의 검토를 통해)가 있지만, 여러 측정 결과 중 (사유 없이) 하나 또는 일부 subset만 완전히 보고되고, 최종 보고 결과는 그 결과(통계적 유의성 등)에 근거하여 선택되었을 가능성이 높다. 결과에 기초하여 선택하는 것은 연구결과가 뉴스 가치가 있고, 출판할 가치가 있거나, 사전 가설을 확인하기에 충분히 주목할 만한 것이어야 한다는 열망에서 비롯될 수 있다. 예를 들어, 임상시험 중재가 유익하다는 선입견을 가지고 있거나 이해상충이 있는 임상시험 연구자들은 임상시험 중재에 유리한 결과 측정을 선택적으로 보고하는 경향이 있을 수 있다. • N/PN: (일반적으로 임상시험 프로토콜 또는 통계 분석 계획의 조사를 통해) 결과 영역에 대해 모든 적격 보고 결과가 의도된 모든 중재 결과 측정과 일치한다는 명확한 근거가 있다. <p>혹은</p> <p>중재 결과 영역을 측정할 수 있는 방법은 단 한 가지뿐이다 (여러번 측정해서 하나를 선택할 기회가 없는 경우).</p> <p>혹은</p> <p>동일한 임상시험에 대한 서로 다른 보고서 간에 중재 결과 측정들이 일관되지 않지만, 임상시험 연구자들이 불일치의 이유를 제공했고, 연구결과와의 본질과 관련이 없다.</p> <ul style="list-style-type: none"> • NI: 분석 의도를 이용할 수 없거나, 분석 의도를 충분히 상세히 보고하지 않아 평가를 할 수 없고, 중재 결과 영역을 측정할 수 있었던 방법이 두 개 이상 있다. 	<p>Y/PY/ PN/N/NI</p>
<p>5.3 데이터의 다중 적격 분석?</p>	<p>(1) 특정 중재 결과는 여러 가지 방법으로 분석할 수 있다. 예를 들어, 비보정 vs 보정 모형; 최종값 vs 기저 상태로부터 변화량 vs 공분산 분석; 변수 변환; 복합 결과에 대한 다른 정의(예: '주요 유해사건'); 다양한 역할을 가진 연속형 변수를 범주형 변수로 변환; 다양한 공변량 집합에 의한 보정; 결측치를 처리하기 위한 다른 전략 등이 그것이다.</p>	<p>Y/PY/ PN/N/NI</p>

신호질문	세부설명	응답 옵션
	<p>복수의 방법을 적용하면 특정 결과 측정에 대한 복수의 효과 추정치가 생성된다. 복수의 추정치가 생성 되지만 결과(예: 통계적 유의성)에 근거하여 1개 결과 또는 하나의 하위 세트만 보고되는 경우, 최종 보고 결과는 비독립 위험이 높다.</p> <p>(2) RoB 2 도구 사용자가 적격으로 고려할 수 있는 분석으로 제한하도록 주의를 기울여야 한다. 예를 들어, 중재 시행 후 측정치만 메타 분석(예: 무작위배정 후 12주)에 포함하는 것이 적격이고, 이 결과가 보고된 경우, 기저 치료부터 변화량에 우선하여 이 결과 기초해서 보고하였다고 하여도 선택 문제는 없을 것이다.</p> <ul style="list-style-type: none"> • Y/PY:(보통 임상시험 프로토콜이나 통계 분석 계획의 조사를 통해) 측정이 복수의 적격 방법으로 분석되었다는 명확한 증거가 있지만, 분석의 한 부분 또는 일부에 대한 데이터는 충분히 보고되고, 최종 보고 결과는 하나만 혹은 하나의 하부세트만 보고한 경우(이에 대한 정당화 없이), 최종 보고 결과는 결과에 기초해서 선택되었을 가능성이 높은 경우이다. 결과에 기초하여 선택하는 것은 연구결과가 뉴스 가치가 있고, 출판할 가치가 있거나, 사전 가설을 증명하기에 충분히 주목할 만한 것이어야 한다는 열망에서 비롯될 수 있다. 예를 들어, 임상시험 중재가 유익하다는 선입견을 가지고 있거나 이해 상충이 있는 임상시험 연구자들은 임상시험 중재에 유리한 결과 측정을 선택적으로 보고하는 경향이 있을 수 있다. • N/PN: 결과 측정에 대한 모든 적격 보고 결과가 의도된 분석과 동일하다는 명확한 증거가 있다(일반적으로 임상 임상시험 프로토콜 또는 통계 분석 계획 조사를 통해). 아니면, 결과 측정을 분석할 수 있는 방법은 단 한 가지뿐이다(다중 분석에서 선택할 기회가 없는 경우). 아니면, 동일한 임상시험에 대한 서로 다른 보고서 간에 중재 결과 측정들이 일관되지 않지만, 임상시험 시행자들이 불일치의 이유를 제공했고, 결과의 성격과 관련이 없다. • NI: 다음과 같은 경우 '정보 없음'으로 응답한다 분석 의도를 이용할 수 없거나, 분석 의도를 충분히 상세히 보고하지 않아 평가를 할 수 없고 중재 결과 영역을 측정할 수 있었던 방법이 두 개 이상 있다. 	

비독립위험 평가	알고리즘 참조	낮음/높음/일부 우려
<p>선택사항: 보고된 연구결과 선택의 비독립의 예측 방향은 어떠한가?</p>	<p>비독립의 방향을 예측할 수 있으면, 이를 기술하는 것이 도움이 된다. 방향은 무효점(혹은 무효점의 반대쪽)을 향한, 혹은 실험 중재번호, 비교중재 번호로 기술한다.</p>	<p>적용불가/실험 중재번호/비교중재번호/무효점을 향한/무효점에서 멀어짐/예측불가</p>

(3) 보고된 연구결과 선택의 비뚤림위험 판단 알고리즘



(4) 보고된 연구결과 선택의 비뚤림위험 판단의 예시

신호질문	서술	응답 옵션
5.1 이 연구결과를 도출한 자료는 분석할 수 있도록 미리 지정된 분석 계획에 따라 중재 결과 자료 눈가림이 해제되기 전에 분석되었는가?	No information	NI
평가되고 있는 수치형 연구결과는 다음과 같은 연구결과에 기초하여 선택되었을 가능성이 있는가?		
5.2 중재 결과 영역 내에서 여러 개의 적합한 중재 연구결과 측정(예: 척도, 정의, 시점)?	When comparing outcome measures and analyses listed in the methods section of an article, all eligible reporting results might be consistent with all intended intervention outcome measures.	PN
5.3 데이터의 다중 적격 분석?		N
비뚤림위험 평가		일부우려

7) 전반적인 비뚤림위험 평가의 예

비뚤림위험 판단	일부우려	낮은 위험/ 높은 위험/ 일부우려
선택사항: 중재결과에 대한 전반적인 예측 비뚤림 방향은 어떠한가?	실험중재번호	적용불가/실험중재번호/비교중재번호/무효점을 함함/무효점에서 멀어짐/예측불가

PART 3

NRS 비돌림위험 평가 도구 : ROBINS-I

4장. ROBINS-I 도구

가. 개요	134
나. 국문 ROBINS-I 평가서식지	135
다. 국문 ROBINS-I	142
라. 국문 ROBINS-I 사용자 매뉴얼	159



4장. ROBINS-I 도구

가. 개요

중재 효과 체계적 문헌고찰의 목표는 하나 이상의 중재결과에 대한 효과의 인과성을 결정하는 것이다. 무작위 임상시험은 중재군과 비교군이 방법론적으로 견고하여 인과성을 추론하게 된다. 그러나 무작위 임상시험만으로 관심 질문에 대한 근거가 충분하지 않을 때, 비무작위 임상시험(NRSIs)을 포함하게 된다.

ROBINS-I (Risk Of Bias In Non-randomized Studies of Interventions) 도구는 비무작위 임상시험의 연구결과에 대한 비뚤림위험을 평가하는 도구이다. 비무작위 임상시험 유형에는 코호트 연구(cohort study), 환자-대조군 연구(case-control study), 통제 전후 연구(controlled before-and-after study), 단속 시계열 연구(interrupted-time-series), 그리고 완전한 무작위 배정에 미치지 못하는 방법을 사용하여 배정되는 통제된 임상시험('준무작위(quasi-randomized)' 연구라고 함) 같은 관찰연구가 해당된다.

이 도구는 무작위 임상시험 평가도구인 Cochrane RoB 도구(Higgins et al, 2011)¹⁾와 진단정확도 평가도구인 QUADAS 2 (Whiting et al, 2011)²⁾ 도구를 기반으로 Sterne 등³⁾에 의해 2016년에 개발되었으며, 정해진 비뚤림 영역에서 판단에 도움이 되는 신호질문에 따라 비뚤림위험을 평가하는 방법을 제공해 주어 근거를 체계적으로 정리하게 해준다.

1) Higgins JPT, Altman DG, Gøtzsche PC, Jüi P, Moher D, Oxman AD, Savovic J, Schulz KF, Weeks L, Sterne JAC, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011; 343: d5928

2) Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM: QUADAS-2 Group. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011; 155: 529-36.

3) Sterne JAC, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, Henry D, Altman DG, Ansari MT, Boutron I, Carpenter JR, Chan AW, Churchill R, Deeks JJ, Hróbjartsson A, Kirkham J, Jüni P, Loke YK, Pigott TD, Ramsay CR, Regidor D, Rothstein HR, Sandhu L, Santaguida PL, Schünemann HJ, Shea B, Shrier I, Tugwell P, Turner L, Valentine JC, Waddington H, Waters E, Wells GA, Whiting PF, Higgins JPT. ROBINS-I: a tool for assessing risk of bias in non-randomized studies of interventions. *BMJ* 2016; 355: i4919

나. 국문 ROBINS-I 평가서식지

1단계: 프토토콜 단계

문헌고찰 질문을 명확하게 기술하시오.

연구대상자	
실험 중재	
비교 중재	
중재결과	

모든 혹은 대부분의 연구에서 중요한 교란 영역의 목록을 작성하시오.

중재군에 따라 다를 수 있고 중재결과에 영향을 줄 수 있는 공동 중재의 목록을 작성하시오.

2단계: 개별 연구

연구에 특정한 목표 임상시험을 기술하시오.

연구설계	개별 무작위 / 집단 무작위 / 매칭(예: 교차)
연구대상자	
실험 중재	
비교 중재	

이 연구에 대한 리뷰 팀의 목적은?

- 중재에 대한 배정의 효과를 평가함 중재의 시작 및 준수의 효과를 평가함

중재결과를 기술하시오.

비독립위험평가를 시행하는 중재결과를 구체적으로 기술하시오(일반적으로 결과 요약표에 표시되어 있는 것 중). 이것이 중재의 이득인지 위해인지 명시할 것

평가할 수치형 연구결과를 구체적으로 기술하시오.

분석이 여러 번 시행된 경우 어떤 결과인지(예: RR=1.52 (95% CI 0.83-2.77)), 어디에서 자료를 얻었는지 명시할 것(예: 표, 그림, 단락)

교란요인에 대한 사전 고려사항

중요한 교란 영역에 대한 행을 작성하시오. i) 프로토콜에 작성된 것, ii) 이 연구의 세팅과 관련이 있거나 저자가 잠재적으로 중요하다고 간주하는 것

“중요한” 교란 영역은 본 연구의 맥락에서 보정이 중재의 추정 효과에 임상적으로 중요한 변화를 초래할 것으로 예상되는 영역이다. “타당도”는 교란 변수 또는 변수가 영역을 완전히 측정하는지 여부를 나타내며, “신뢰도”는 측정의 정밀도를 나타낸다(측정 오류가 클수록 신뢰도가 떨어짐).

(i) 문헌고찰 프로토콜에 작성된 교란 영역				
교란 영역	측정된 교란 변수	이 변수에 대한 통계가 불필요하다는 근거가 있는가?*	이 변수는 교란 영역을 타당하고 신뢰성 있게 측정하였는가?	선택 사항: 이 변수에 대한 보정 실패는 실험 중재나 비교 중재 중 어느 쪽에 유리할 것으로 예상되는가?
			그렇다 / 아니다 / 정보 없음	실험 중재 선호 / 비교 중재 선호 / 정보 없음

(ii) 이 연구의 세팅과 관련되거나 저자가 중요하다고 판단한 추가 교란 영역

교란 영역	측정된 교란 변수	이 변수에 대한 통제가 불필요하다는 근거가 있는가?*	이 변수는 교란 영역을 타당하고 신뢰성 있게 측정하였는가?	선택 사항: 이 변수에 대한 보정 실패는 실험 중재나 비교 중재 중 어느 쪽에 유리할 것으로 예상되는가?
			그렇다 / 아니다 / 정보 없음	실험 중재 선호 / 비교 중재 선호 / 정보 없음

* 특정 연구의 맥락에서, 일부 변수는 교란 요인이 아니어서 분석에 포함되지 않을 수 있다. (a) 중재 결과를 예측하지 않는 경우, (b) 중재를 예측하지 않는 경우, (c) 보정이 1차 매개 변수(파라미터)의 추정된 효과와 거의 또는 전혀 차이가 없는 경우. “통계적으로 유의미한 연관성이 없음”은 “예측하지 않음”과는 같지 않음에 주의하십시오.

공동 중재에 대한 사전 고려사항

중요한 공동 중재에 대한 행을 작성하십시오. (i) 문헌고찰 프로토콜에 작성된 것, (ii) 이 연구의 세팅과 관련이 있거나 저자가 중요하다고 간주하는 것

“중요한” 공동 중재는 본 연구의 맥락에서 보정이 중재의 추정 효과에 임상적으로 중요한 변화를 초래할 것으로 예상되는 것이다.

(i) 문헌고찰 프로토콜에 작성된 공동 중재

공동 중재	이 중재에 대한 통제가 불필요하다는 근거가 있는가?(예: 제공되지 않았기 때문에)	이 공동 중재는 실험 중재나 비교 중재 중 어느 쪽에 유리할 것으로 예상되는가?
		실험 중재 선호 / 비교 중재 선호 / 정보 없음
		실험 중재 선호 / 비교 중재 선호 / 정보 없음
		실험 중재 선호 / 비교 중재 선호 / 정보 없음
		실험 중재 선호 / 비교 중재 선호 / 정보 없음

(ii) 이 연구의 세팅과 관련되거나 저자가 중요하다고 판단한 추가 공동 중재

공동 중재	이 중재에 대한 통제가 불필요하다는 근거가 있는가?(예: 제공되지 않았기 때문에)	이 공동 중재는 실험 중재나 비교 중재 중 어느 쪽에 유리할 것으로 예상되는가?
		실험 중재 선호 / 비교 중재 선호 / 정보 없음
		실험 중재 선호 / 비교 중재 선호 / 정보 없음
		실험 중재 선호 / 비교 중재 선호 / 정보 없음
		실험 중재 선호 / 비교 중재 선호 / 정보 없음

비돌림위험 평가(코호트 타입 연구)

- **그렇다(Y) / 아마도 그렇다(PY) / 아마도 아닐 것이다(PN) / 아니다(N)** / 적용할 수 없음(NA) / 정보 없음(NI)
- **녹색으로 밑줄 친 응답**은 낮은 비돌림위험, **적색의 응답**은 비돌림위험을 나타냄

1. 교란으로 인한 비돌림		
신호질문	판단근거	응답선택
1.1 이 연구에서 중재 효과의 교란이 일어날 가능성이 있는가? 1.10i N/PN 인 경우: 연구는 교란에 의한 비돌림위험이 낮은 것으로 간주될 수 있으며, 더 이상의 신호 질문을 고려할 필요가 없다.		Y/ PY / <u>PN/N</u>
1.10i Y/PY 인 경우: 시간 변동 교란요인을 평가할 필요가 있는지 결정하십시오		
1.2 제공받은 중재에 따라, 연구대상자의 추적관찰기간을 구분하여 분석하였는가? N/PN 인 경우, 기저상태의 교란과 관련지어 대답하십시오.(1.4 ~ 1.6) Y/PY 인 경우 1.3으로 가시오.		NA / Y / PY / PN / N / NI
1.3 중재의 중단 또는 교환이 결과에 대한 예후 인자와 관련이 있을 가능성이 있는가? N/PN 인 경우, 기저선의 교란과 관련지어 대답하라.(1.4에서 1.6) Y/PY 인 경우, 기저선과 시간 변동의 교란과 관련지어 대답하라.(1.7에서 1.8)		NA / Y / PY / PN / N / NI
기저 상태 교란에만 관련이 있는 질문		
1.4 저자는 모든 중요한 교란 영역을 통제하는 적절한 분석 방법을 사용하였는가?		NA / <u>Y / PY</u> / PN / N / NI
1.5 1.40i Y/PY 인 경우: 이 연구에서 사용 가능한 변수에 의해 측정된 교란 영역이 타당하고 신뢰성 있게 통제되었는가?		NA / <u>Y / PY</u> / PN / N / NI
1.6 저자는 중재의 영향을 받을 수 있는 중재 후 변수를 통제하였는가?		NA / Y / PY / <u>PN / N</u> / NI
기저 상태와 시간 변동 교란에 대한 질문		
1.7 저자는 모든 중요한 교란 영역과 시간-변동 교란을 보정하는 적절한 분석방법을 사용하였는가?		NA / <u>Y / PY</u> / PN / N / NI
1.8 1.70i Y/PY 인 경우: 이 연구에서 이용할 수 있는 변수에 의해 측정된 교란 영역이 타당하고 신뢰성 있게 보정되었는가?		NA / <u>Y / PY</u> / PN / N / NI
비돌림위험 판단		낮음 / 중등도 / 높음 / 매우 높음 / 정보 없음
선택 사항: 교란으로 인해 예상되는 비돌림의 방향은 무엇인가?		실험 중재 선호 / 비교 중재 선호 / 예측할 수 없음

2. 연구대상자 선택의 비돌림		
신호질문	판단근거	응답선택
2.1 중재 시작 후 관찰된 연구대상자의 특성에 기초하여 연구대상자를(또는 분석 대상) 선택하였는가? 2.1이 N/PN인 경우: 2.4로 가시오.		Y / PY / <u>PN</u> / N / NI
2.2 2.1이 Y/PY인 경우: 선택에 영향을 준 중재 후 변수가 중재와 관련이 있을 가능성이 있었는가?		NA / Y / PY / <u>PN</u> / N / NI
2.3 2.2이 Y/PY인 경우: 선택에 영향을 준 중재 후 변수가 중재 결과 또는 중재 결과의 원인에 영향을 받을 가능성이 있는가?		NA / Y / PY / <u>PN</u> / N / NI
2.4 대부분의 연구대상자에서 추적 관찰의 시작 시점과 중재의 시작 시점이 일치하는가?		<u>Y</u> / PY / PN / N / NI
2.5 2.2와 2.3이 Y/PY이거나, 2.4이 N/PN인 경우: 선택 비돌림을 교정할 수 있는 보정 방법을 사용하였는가?		NA / <u>Y</u> / PY / PN / N / NI
비돌림위험 판단		낮음 / 중등도 / 높음 / 매우 높음 / 정보 없음
선택 사항: 연구대상자 선택에 의해 예상되는 비돌림의 방향은 무엇인가?		실험 중재 선호 / 비교 중재 선호 / 무효점으로 향함 / 무효점에서 멀어짐 / 예측할 수 없음
3. 중재 분류의 비돌림		
신호질문	판단근거	응답선택
3.1 중재 군이 명확하게 정의되어 있는가?		<u>Y</u> / PY / PN / N / NI
3.2 중재 시작 시 중재군들을 정의하는데 사용된 정보가 있는가?		<u>Y</u> / PY / PN / N / NI
3.3 중재 상태의 분류가 중재 결과의 지식 또는 중재 결과의 위험에 의해 영향을 받을 수 있는가?		Y / PY / <u>PN</u> / N / NI
비돌림위험 판단		낮음 / 중등도 / 높음 / 매우 높음 / 정보 없음
선택 사항: 중재 결과의 측정이나 중재에 의해 예상되는 비돌림의 방향은 무엇인가?		실험 중재 선호 / 비교 중재 선호 / 무효점으로 향함 / 무효점에서 멀어짐 / 예측할 수 없음

4. 의도한 중재에서 이탈되어 나타나는 비돌림		
신호질문	판단근거	응답선택
4.1 통상진료에서 예상되는 것 이상으로 의도한 중재로부터 이탈되는 상황이 있었는가?		Y / PY / <u>PN</u> / N / NI
4.2 4.10 Y/PY 인 경우: 의도한 중재로부터 이탈된 상황이 중재 기간에 불균형을 초래하고, 중재 결과에 영향을 미쳤을 가능성이 있는가?		NA / Y / PY / <u>PN</u> / N / NI
본 연구의 목적이 중재에 대한 시작 및 준수의 효과를 평가하는 것이라면 4.3~4.6에 응답하십시오.		
4.3 중요한 공동 중재는 중재기간 균형을 이루는가?		<u>Y</u> / PY / PN / N / NI
4.4 대부분의 연구대상자에게 중재가 성공적으로 시행되었는가?		<u>Y</u> / PY / PN / N / NI
4.5 연구대상자는 배정된 중재법을 준수하였는가?		<u>Y</u> / PY / PN / N / NI
4.6 4.3, 4.4 또는 4.50이 N/PN 인 경우: 중재의 시작 및 준수의 효과를 추정하기 위해 적절한 분석방법을 사용하였는가?		NA / <u>Y</u> / PY / PN / N / NI
비돌림위험 판단		낮음 / 중등도 / 높음 / 매우 높음 / 정보 없음
선택 사항: 의도한 중재로부터 이탈로 인해 예상되는 비돌림의 방향은 무엇인가?		실험 중재 선호 / 비교 중재 선호 / 무효점으로 향함 / 무효점에서 멀어짐 / 예측할 수 없음
5. 결측치로 인한 비돌림		
신호질문	판단근거	응답선택
5.1 모든 혹은 거의 모든 연구대상자의 중재 결과 자료를 사용할 수 있었는가?		<u>Y</u> / PY / PN / N / NI
5.2 중재 상태에 대한 결측 자료로 인해 연구대상자가 제외되었는가?		Y / PY / <u>PN</u> / N / NI
5.3 분석에 필요한 다른 변수에 대한 결측자료로 인해 연구대상자가 제외되었는가?		Y / PY / <u>PN</u> / N / NI
5.4 5.10이 PN/N 혹은 5.2나 5.30이 Y/PY인 경우: 결측된 연구대상자의 비율과 원인이 중재 기간 유사한가?		NA / <u>Y</u> / PY / PN / N / NI
5.5 5.10이 PN/N 혹은 5.2나 5.30이 Y/PY인 경우: 결측치에 관계없이 연구결과가 견고(robust)하다는 근거가 있는가?		NA / <u>Y</u> / PY / PN / N / NI
비돌림위험 판단		낮음 / 중등도 / 높음 / 매우 높음 / 정보 없음
선택 사항: 결측치에 의해 예상되는 비돌림의 방향은 무엇인가?		실험 중재 선호 / 비교 중재 선호 / 무효점으로 향함 / 무효점에서 멀어짐 / 예측할 수 없음

6. 중재결과 측정 비틀림		
신호질문	판단근거	응답선택
6.1 중재결과 측정이 제공받은 중재의 지식에 의해 영향을 받았을 수 있는가?		Y / PY / <u>PN</u> / N / NI
6.2 중재결과 평가자는 연구대상자들이 받은 중재를 알고 있는가?		Y / PY / <u>PN</u> / N / NI
6.3 중재결과 평가 방법은 중재 군 간에 유사했는가?		<u>Y</u> / PY / PN / N / NI
6.4 제공받은 중재와 관련된 중재결과 측정 시 체계적 오류가 있었는가?		Y / PY / <u>PN</u> / N / NI
비틀림위험 판단		낮음 / 중등도 / 높음 / 매우 높음 / 정보 없음
선택 사항: 중재결과 측정에 의해 예상되는 비틀림의 방향은 무엇인가?		실험 중재 선호 / 비교 중재 선호 / 무효점으로 향함 / 무효점에서 멀어짐 / 예측할 수 없음
7. 보고된 연구결과 선택의 비틀림		
신호질문	판단근거	응답선택
보고된 효과 추정치는 다음과 같은 연구결과에 기초하여 선택되었을 가능성이 있는가?		
7.1 중재 결과 영역 내에서 여러 개의 중재 결과 측정		Y / PY / <u>PN</u> / N / NI
7.2 중재-결과 관계에서의 다중 분석		Y / PY / <u>PN</u> / N / NI
7.3 다른 하위군		Y / PY / <u>PN</u> / N / NI
비틀림위험 판단		낮음 / 중등도 / 높음 / 매우 높음 / 정보 없음
선택 사항: 보고된 연구결과의 선택에 의해 예상되는 비틀림의 방향은 무엇인가?		실험 중재 선호 / 비교 중재 선호 / 무효점으로 향함 / 무효점에서 멀어짐 / 예측할 수 없음
전체 비틀림위험 판단		낮음 / 중등도 / 높음 / 매우 높음 / 정보 없음
선택 사항: 중재 결과에 대해 전반적으로 예상되는 비틀림의 방향은 무엇인가?		실험 중재 선호 / 비교 중재 선호 / 무효점으로 향함 / 무효점에서 멀어짐 / 예측할 수 없음

다. 국문 ROBINS-I

1단계: 프로토콜 단계

문헌고찰 질문을 명확하게 기술하시오.

연구대상자	
실험 중재	
비교 중재	
중재결과	

모든 혹은 대부분의 연구에서 중요한 교란 영역의 목록을 작성하시오.

중재군에 따라 다를 수 있고 중재결과에 영향을 줄 수 있는 공동 중재의 목록을 작성하시오.

2단계: 개별 연구

연구에 특정한 목표 임상시험을 기술하시오.

연구설계	개별 무작위 / 집단 무작위 / 매칭(예: 교차)
연구대상자	
실험 중재	
비교 중재	

이 연구에 대한 리뷰 팀의 목적은?

- 중재에 대한 배정의 효과를 평가함
- 중재의 시작 및 준수의 효과를 평가함

중재결과를 기술하시오.

비뚫림위험평가를 시행하는 중재결과를 구체적으로 기술하시오(일반적으로 결과 요약표에 표시되어 있는 것 중에서 구체화 함). 이것이 중재의 이득인지 위해인지 명시할 것

평가할 수치형 연구결과를 구체적으로 기술하시오.

분석이 여러 번 시행된 경우 어떤 결과인지(예: RR=1.52 (95% CI 0.83-2.77)), 어디에서 자료를 얻었는지 명시할 것(예: 표, 그림, 단락)

교란요인에 대한 사전 고려사항

중요한 교란 영역에 대한 행을 작성하시오. i) 프로토콜에 작성된 것, ii) 이 연구의 세팅과 관련이 있거나 저자가 잠재적으로 중요하다고 간주하는 것

“중요한” 교란 영역은 본 연구의 맥락에서 보정이 중재의 추정 효과에 임상적으로 중요한 변화를 초래할 것으로 예상되는 영역이다. “타당도”는 교란 변수 또는 변수가 영역을 완전히 측정하는지 여부를 나타내며, “신뢰도”는 측정의 정밀도를 나타낸다(측정 오류가 클수록 신뢰도가 떨어짐).

(i) 문헌고찰 프로토콜에 작성된 교란 영역

교란 영역	측정된 교란 변수	이 변수에 대한 통제가 불필요하다는 근거가 있는가?*	이 변수는 교란 영역을 타당하고 신뢰성 있게 측정하였는가?	선택 사항: 이 변수에 대한 보정 실패는 실험 중재나 비교 중재 중 어느 쪽에 유리할 것으로 예상되는가?
			그렇다 / 아니다 / 정보 없음	실험 중재 선호 / 비교 중재 선호 / 정보 없음

(ii) 이 연구의 세팅과 관련되거나 저자가 중요하다고 판단한 추가 교란 영역

교란 영역	측정된 교란 변수	이 변수에 대한 통제가 불필요하다는 근거가 있는가?*	이 변수는 교란 영역을 타당하고 신뢰성 있게 측정하였는가?	선택 사항: 이 변수에 대한 보정 실패는 실험 중재나 비교 중재 중 어느 쪽에 유리할 것으로 예상되는가?
			그렇다 / 아니다 / 정보 없음	실험 중재 선호 / 비교 중재 선호 / 정보 없음

* 특정 연구의 맥락에서, 일부 변수는 교란요인이 아니어서 분석에 포함되지 않을 수 있다. (a) 중재 결과를 예측하지 않는 경우, (b) 중재를 예측하지 않는 경우, (c) 보정이 1차 매개 변수(파라미터)의 추정된 효과와 거의 또는 전혀 차이가 없는 경우. "통계적으로 유의미한 연관성이 없음"은 "예측하지 않음"과는 같지 않음에 주의하십시오.

공동 중재에 대한 사전 고려사항

중요한 공동 중재에 대한 행을 작성하십시오. (i) 문헌고찰 프로토콜에 작성된 것, (ii) 이 연구의 세팅과 관련이 있거나 저자가 중요하다고 간주하는 것

"중요한" 공동 중재는 본 연구의 맥락에서 보정이 중재의 추정 효과에 임상적으로 중요한 변화를 초래할 것으로 예상되는 것이다.

(i) 문헌고찰 프로토콜에 작성된 공동 중재

공동 중재	이 중재에 대한 통제가 불필요하다는 근거가 있는가?(예: 제공되지 않았기 때문에)	이 공동 중재는 실험 중재나 비교 중재 중 어느 쪽에 유리할 것으로 예상되는가?
		실험 중재 선호 / 비교 중재 선호 / 정보 없음
		실험 중재 선호 / 비교 중재 선호 / 정보 없음
		실험 중재 선호 / 비교 중재 선호 / 정보 없음
		실험 중재 선호 / 비교 중재 선호 / 정보 없음

(ii) 이 연구의 세팅과 관련되거나 저자가 중요하다고 판단한 추가 공동 중재

공동 중재	이 중재에 대한 통제가 불필요하다는 근거가 있는가?(예: 제공되지 않았기 때문에)	이 공동 중재는 실험 중재나 비교 중재 중 어느 쪽에 유리할 것으로 예상되는가?
		실험 중재 선호 / 비교 중재 선호 / 정보 없음
		실험 중재 선호 / 비교 중재 선호 / 정보 없음
		실험 중재 선호 / 비교 중재 선호 / 정보 없음
		실험 중재 선호 / 비교 중재 선호 / 정보 없음

비뚫림위험 평가(코호트 타입 연구)

녹색으로 밑줄 친 응답은 낮은 비뚫림위험에 대한 잠재적 표시이고, 적색의 응답은 높은 비뚫림위험을 나타내는 잠재적 표시이다.

① 교란으로 인한 비뚫림

신호 질문	세부 설명(elaboration)	응답 선택
<p>1.1 이 연구에서 중재 효과의 교란이 일어날 가능성이 있는가?</p> <p>1.10이 N/PN 인 경우: 연구는 교란에 의한 비뚫림위험이 낮은 것으로 간주될 수 있으며, 더 이상의 신호 질문을 고려할 필요가 없다.</p>	<p>드물게 치료 결정에 영향을 미치는 요인과 관련이 거의 없는 경우 교란이 예상되지 않으나, 이 연구는 교란으로 인한 비뚫림위험이 낮은 것으로 간주될 수 있다. 이 신호 질문에 대한 응답 옵션에 NI(정보 없음)는 없다.</p>	<p>그렇다(Y) / 아마도 그렇다(PY) / 아마도 아닐 것이다(PN) / 아니다(N)</p>
<p>1.10이 Y/PY 인 경우: 시간 변동 교란요인을 평가할 필요가 있는지 결정하십시오.</p>		
<p>1.2 제공받은 중재에 따라, 연구대상자의 추적관찰기간을 구분하여 분석하였는가?</p> <p>N/PN 인 경우, 기저 상태의 교란과 관련지어 대답하십시오.(1.4 ~ 1.6)</p> <p>Y/PY 인 경우 1.3으로 가시오.</p>	<p>연구대상자가 중재 군간 전환이 가능한 경우 시간 변동 교란에 의해 중재와 중재 결과간의 연관성이 편향될 수 있다. 이는 예후 요인에 의해 의도한 중재군간 전환이 영향을 받을 경우 발생한다.</p>	<p>적용할 수 없음(NA) / Y / PY / PN / N / 정보 없음(NI)</p>
<p>1.3 중재의 중단 또는 교환이 결과에 대한 예후 인자와 관련이 있을 가능성이 있는가?</p> <p>N/PN 인 경우, 기저선의 교란과 관련지어 대답하라.(1.4에서 1.6)</p> <p>Y/PY 인 경우, 기저선과 시간 변동의 교란과 관련지어 대답하라.(1.7에서 1.8)</p>	<p>중재 전환이 중재 결과와 관련이 없는 경우(예: 중재 결과가 예상치 못한 위해인 경우) 시간-변동 교란은 존재하지 않으며 기저 상태 교란에 대한 통제만 필요하다.</p>	<p>NA / Y / PY / PN / N / NI</p>
<p>기저 상태 교란에만 관련이 있는 질문</p>		
<p>1.4 저자는 모든 중요한 교란 영역을 통제하는 적절한 분석 방법을 사용하였는가?</p>	<p>측정된 교란요인을 통제하는 적절한 방법으로는 층화, 회귀분석, 매칭, 표준화 및 역확률 가중(inverse probability weighting) 같은 방법들이 있다. 개별 변수 또는 추정된 성향점수로 통제할 수 있다. 역확률 가중치는 성향점수 함수를 기반으로 한다. 각각의 방법은 측정되지 않은 교란이나 잔류 교란이 없다는 가정에 기반한다.</p>	<p>NA / Y / PY / PN / N / NI</p>

<p>1.5 1.4I Y/PY 인 경우: 이 연구에서 사용 가능한 변수에 의해 측정된 교란 영역이 타당하고 신뢰성 있게 통제되었는가?</p>	<p>교란을 적절히 통제하려면 보정된 변수가 교란 영역을 타당하고 신뢰성 있게 측정해야한다. 일부 주제에서는 교란 영역에 대한 타당하고 신뢰할 수 있는 측정도구의 목록이 프로토콜에서 정해지지만, 어떤 경우는 해당 목록을 사용하지 못할 수 있다. 저자는 특정 측정 방법의 사용을 정당화하기 위해 참고문헌을 인용할 수 있다. 저자가 타당도 혹은 신뢰도 제시 없이 교란 변수를 통제하는 경우, 측정이 주관적이 될 수 있음에 주의를 기울여야 한다. 주관적 측정(예: 자가보고 기반)은 실험실 연구와 같은 객관적인 측정보다 낮은 타당도와 신뢰도를 가질 수 있다.</p>	<p>NA / <u>Y / PY</u> / PN / N / NI</p>
<p>1.6 저자는 중재의 영향을 받을 수 있는 중재 후 변수를 통제하였는가?</p>	<p>중재에 의해 영향을 받게 되는 중재 후 변수를 통제하는 것은 적절하지 않다. 매개 변수에 대한 통제는 중재의 직접적인 효과를 추정하고 비돌림을 유발할 수 있다. 중재와 결과의 일반적인 효과에 대해 통제하면 비돌림이 생기게 된다.</p>	<p>NA / <u>Y / PY</u> / <u>PN / N / NI</u></p>

기저 상태와 시간 변동 교란에 대한 질문

<p>1.7 저자는 모든 중요한 교란 영역과 시간-변동 교란을 보정하는 적절한 분석방법을 사용하였는가?</p>	<p>RCT와 NRSI에서 중재의 시작 및 준수 효과를 추정하기 위해서는 시간-변동 교란에 대한 보정이 필요하다. 적절한 방법으로는 역확률 가중치를 기반으로 하는 방법이 있다. 시간에 따라 변하는 교란이 있다면 시간-업데이트된 교란요인을 포함하는 표준 회귀 모델이 문제가 될 수 있다.</p>	<p>NA / <u>Y / PY</u> / PN / N / NI</p>
<p>1.8 1.7I Y/PY 인 경우: 이 연구에서 이용할 수 있는 변수에 의해 측정된 교란 영역이 타당하고 신뢰성 있게 보정되었는가?</p>	<p>1.5를 참고하십시오.</p>	<p>NA / <u>Y / PY</u> / PN / N / NI</p>
<p>비돌림위험 판단</p>	<p>표 1을 보시오.</p>	<p>낮음 / 중등도 / 높음 / 매우 높음 / 정보 없음</p>
<p>선택 사항: 교란으로 인해 예상되는 비돌림의 방향은 무엇인가?</p>	<p>하나 이상의 중요한 교란 영역이 통제되지 않았기 때문에 실제 효과 추정값이 연구에서 추정된 효과보다 크거나 작을 것으로 예상할 수 있는가? 이 질문에 답하려면 다른 연구에서의 전문가적 지식과 결과에 기반해야하기 때문에 근거가 되는 모든 연구를 검토한 후에야 완성될 수 있다. 측정되지 않은 각 영역의 잠재적 영향과 분석에서 통제되지 않은 모든 중요한 교란 영역이 동일한 방향으로 추정치를 변경할 가능성이 있는지 또는 분석에서 통제되지 않은 하나의 중요한 교란 영역이 지배적인 영향을 미칠 가능성이 있는지 고려해야 한다.</p>	<p>실험 중재 선호 / 비교 중재 선호 / 예측할 수 없음</p>

② 연구대상자 선택의 비돌림

신호 질문	세부 설명(elaboration)	응답 선택
<p>2.1 중재 시작 후 관찰된 연구대상자의 특성에 기초하여 연구대상자를 (또는 분석 대상) 선택하였는가?</p> <p>2.1이 N/PN 인 경우: 2.4로 가시오.</p>	<p>이 영역은 중재 시작 후 관찰된 연구대상자의 특성에 기초한 선택에만 관심이 있다. 중재의 시작전에 관찰된 특성에 따라 선택하였다면, 실험 중재와 비교 중재 사이의 불균형을 중재 결과에 대한 예측적인 기저 특성에 따라 통제함으로써 해결할 수 있다 (기저 상태 교란).</p>	<p>Y / PY / PN / N / NI</p>
<p>2.2 2.1이 Y/PY 인 경우: 선택에 영향을 준 중재 후 변수가 중재와 관련이 있을 가능성이 있었는가?</p>	<p>선택 비돌림은 선택이 중재의 영향 또는 중재의 원인과 중재 결과 또는 중재 결과의 원인과 관련될 때 발생한다. 따라서 연구대상자의 선택이 중재와 중재 결과 모두와 관련이 있다면 연구결과는 선택 비돌림이 있을 것이다.</p>	<p>NA / Y / PY / PN / N / NI</p>
<p>2.3 2.2이 Y/PY 인 경우: 선택에 영향을 준 중재 후 변수가 중재 결과 또는 중재 결과의 원인에 영향을 받을 가능성이 있는가?</p>		<p>NA / Y / PY / PN / N / NI</p>
<p>2.4 대부분의 연구대상자에서 추적 관찰의 시작 시점과 중재의 시작 시점이 일치하는가?</p>	<p>연구대상자가 중재의 시작부터 추적 되지 않는다면, 추적관찰 기간에서 제외될 것이며, 중재 직후 시행된 중재 결과는 결측될 것이다. 이 문제는 중재의 신규 사용자보다 기존 사용자가 분석에 포함될 때 발생할 것이다.</p>	<p>Y / PY / PN / N / NI</p>
<p>2.5 2.2와 2.3이 Y/PY 이거나, 2.4 이 N/PN인 경우: 선택 비돌림을 교정할 수 있는 보정 방법을 사용하였는가?</p>	<p>예를 들어, 역확률 가중치를 사용하여 선택 비돌림이 제거된 유사 모집단을 생성하거나 혹은 결측된 연구대상자의 분포를 모형화하거나 추적관찰기간과 결과 발생을 모형화하여 결측자료 분석 방법론을 사용하여 선택 비돌림을 교정하는 것이 원칙적으로 가능하다. 그러나, 이러한 방법은 거의 사용되지 않으며 이 질문에 대한 대답은 일반적으로 “아니다”가 될 것이다.</p>	<p>NA / Y / PY / PN / N / NI</p>
<p>비돌림위험 판단</p>	<p>표 1을 보시오.</p>	<p>낮음 / 중등도 / 높음 / 매우 높음 / 정보 없음</p>
<p>선택 사항: 연구대상자 선택에 의해 예상되는 비돌림의 방향은 무엇인가?</p>	<p>비돌림의 방향을 예측하는 것이 가능하다면, 이를 기술하는 것이 도움이 될 것이다. 그 방향이 무효점 (null)으로 향하던지(혹은 멀어지는 쪽으로 향하던지) 혹은 중재 어느 한쪽으로 유리하게 작용.</p>	<p>실험 중재 선호 / 비교 중재 선호 / 무효점으로 향함 / 무효점에서 멀어짐 / 예측할 수 없음</p>

③ 중재 분류의 비돌림

신호 질문	세부 설명(elaboration)	응답 선택
3.1 중재 군이 명확하게 정의되어 있는가?	중재의 적절한 비교를 위한 전제 조건은 중재가 잘 정의되어 있어야 한다는 것이다. 모호한 정의는 연구대상자를 분류하는데 비돌림을 초래할 수 있다. 개인 수준의 중재의 경우, 각 중재를 받는 개인을 고려하는 기준, 예를 들어 유형, 세팅, 용량, 빈도, 강도, 중재 시기 등이 명확하고 명시적이어야 한다. 인구집단 수준의 중재(예: 대기오염 방지 대책)의 경우, 그 질문은 인구집단을 명확하게 정의하고 있다면, 대답은 '그렇다'일 가능성이 있다.	Y / PY / PN / N / NI
3.2 중재 시작 시 중재군들을 정의하는데 사용된 정보가 있는가?	일반적으로 시행된 중재에 관한 정보가 후속 결과에 의하여 영향을 받을 수 없는 자료로부터 이용가능한 것이라면, 중재 상태의 차등적 오분류는 거의 일어나지 않을 것이다. 중재 당시 정보를 수집하면 이러한 오분류를 쉽게 피할 수 있다. 인구집단 수준의 중재(예: 대기오염 방지 대책)의 경우 이 질문에 대한 대답은 '그렇다'일 가능성이 높다.	Y / PY / PN / N / NI
3.3 중재 상태의 분류가 중재 결과의 지식 또는 중재 결과의 위험에 의해 영향을 받을 수 있는가?	중재 시점에서 정보 수집은 비돌림을 피하기 충분하지 않을 수 있다. NRSI의 목적을 위해 자료를 수집하는 방법도 오분류를 피해야 한다.	Y / PY / PN / N / NI
비돌림위험 판단	표 1을 보시오.	낮음 / 중등도 / 높음 / 매우 높음 / 정보 없음
선택 사항: 중재 결과의 측정이나 중재에 의해 예상되는 비돌림의 방향은 무엇인가??	비돌림의 방향을 예측하는 것이 가능하다면, 이를 기술하는 것이 도움이 될 것이다. 그 방향이 무효점으로 향하던지(혹은 멀어지는 쪽으로 향하던지) 혹은 중재 중 한쪽으로 유리할 수 있다.	실험 중재 선호 / 비교 중재 선호 / 무효점으로 향함 / 무효점에서 멀어짐 / 예측할 수 없음

④ 의도한 중재에서 이탈되어 나타나는 비돌림

본 연구의 목적이 중재에 대한 부정 효과를 평가하는 것이라면 4.1과 4.2에 응답하십시오		
신호 질문	세부 설명(elaboration)	응답 선택
4.1 통상진료에서 예상되는 것 이상으로 의도한 중재로부터 이탈되는 상황이 있었는가?	중재 후 통상진료에서 발생한 이탈(예: 급성 독성으로 인한 약물 투여의 중단)은 의도한 중재의 일부이므로 중재 부정 효과에 비돌림을 유발하지 않는다. 이탈은 중재군과 비교 중재 간의 차이에 대한 기대로 인해 발생할 수 있다(예를 들어, 연구대상자들이 비교 중재에 배정된 것을 불운하다고 느끼고, 따라서, 적극적 중재 혹은 그러한 요소 혹은 다른 중재를 찾기 때문임). 이러한 이탈은 통상진료의 일부가 아니므로 편향된 효과 추정치를 초래할 수 있다. 그러나, 통상진료에서 개인에 대한 관찰연구에서는 이러한 사항이 예상되지 않는다.	Y / PY / PN / N / NI
4.2 4.1이 Y/PY 인 경우: 의도한 중재로부터 이탈된 상황이 중재 기간에 불균형을 초래하고, 중재 결과에 영향을 미쳤을 가능성이 있는가?	통상진료를 반영하지 않은 의도한 중재로부터 나타나는 이탈은 중재 결과에 영향을 미친다면 중요할 것이나 그렇지 않다면 중요하지 않을 것이다. 또한 두 군에 걸친 이탈의 불균형이 있을 경우에만 비돌림이 발생할 것이다.	NA / Y / PY / PN / N / NI

본 연구의 목적이 중재에 대한 시작 및 준수의 효과를 평가하는 것이라면 4.3~4.6에 응답하십시오.

4.3 중요한 공동 중재는 중재기간 균형을 이루는가?	계획되지 않은 공동 중재가 중재의 추정 효과를 편향시키는 방식으로 시행될 경우 비돌림의 위험이 더 높아질 것이다. 공동 중재가 중재 결과에 영향을 미칠 때는 중요하나, 그렇지 않은 경우에는 중요하지 않다. 비돌림은 중재군 사이에 공동 중재의 불균형이 있을 경우에만 발생할 것이다. 중재 결과에 영향을 미칠 가능성이 있고, 이 연구에서 투여되었을 가능성이 있는 사전-명시된 공동 중재를 포함한 공동 중재를 고려해야 한다. 이러한 공동 중재는 중재 군 간에 균형이 맞는지 고려해야 한다.	Y / PY / PN / N / NI
4.4 대부분의 연구대상자에게 중재가 성공적으로 시행되었는가?	예를 들어, 연구 기간동안 보건의료전문가가 의도한대로 중재가 시행되지 않는다면 비돌림위험이 높을 것이다. 대부분의 연구대상자에게 중재가 성공적으로 수행됐는지 고려해야 한다.	Y / PY / PN / N / NI

<p>4.5 연구대상자는 배정된 중재법을 준수하였는가?</p>	<p>연구대상자들이 의도한대로 중재를 준수하지 않는다면 비둘림위험이 높을 것이다. 미준수에는 불완전한 준수, 중재의 중단, 중재군과 비교 중재의 교차, 다른 적극적인 중재로의 전환 등이 있다. 추적 관찰 기간 동안 배정된 중재를 계속해서 준수한 연구대상자의 비율에 대한 이용가능한 정보를 고려해야 하며, 만약 이 비율이 우려를 일으킬 만큼 충분히 높다면 '아니다', 또는 '아마도 아니다'라고 응답한다. 불완전한 준수가 불가능하도록 단 한번 시행되는 중재 연구에 대해서는 '그렇다'라고 응답한다.</p> <p>중재 전환된 경우, 중재 전환(중재의 중단 포함) 후 추적 관찰기간 동안을 1)새로운 중재에 배당하는 분석 혹은 2) 최초 중재로 배당하는 분석으로 구분한다. 새로운 중재로 배당하는 분석은 시간-변동 교란효과로 처리하는 것으로 여기서는 더 이상 고려하지 않아야 한다.</p>	<p>Y / PY / PN / N / NI</p>
<p>4.6 4.3, 4.4 또는 4.5이 N/PN 인 경우: 중재의 시작 및 준수의 효과를 추정하기 위해 적절한 분석방법을 사용하였는가?</p>	<p>의도한 중재로부터 몇가지 유형의 이탈을 수정하는 분석을 수행할 수 있다. 적절한 분석 전략의 예로 역확률 가중치나 도구변수 추정이 있다. 논문에서 의도한 중재로부터의 이탈에 대한 정보를 보고하지 않고 그러한 분석을 보고하는 것이 가능할 수는 있지만, 정보가 없는 경우 그러한 분석이 적절하다고 판단하기는 어렵다. 이러한 접근법을 사용한 연구를 평가하려면 전문가의 조언이 필요할 수 있다.</p> <p>만약 한 그룹의 모든 사람이 공동 중재를 받았다면 이를 극복하기 위해 보정을 할 수는 없다.</p>	<p>NA / Y / PY / PN / N / NI</p>
<p>비둘림위험 판단</p>	<p>표 2를 보시오.</p>	<p>낮음 / 중등도 / 높음 / 매우 높음 / 정보 없음</p>
<p>선택 사항: 의도한 중재로부터 이탈로 인해 예상되는 비둘림의 방향은 무엇인가?</p>	<p>비둘림의 방향을 예측하는 것이 가능하다면, 이를 기술하는 것이 도움이 될 것이다. 그 방향이 무효점으로 향하던지(혹은 멀어지는 쪽으로 향하던지) 어느 한쪽으로 유리하게 작용.</p>	<p>실험 중재 선호 / 비교 중재 선호 / 무효점으로향함/ 무효점에서멀어짐/ 예측할 수 없음</p>

⑤ 결측치로 인한 비독립

신호 질문	세부 설명(elaboration)	응답 선택
5.1 모든 혹은 거의 모든 연구대상자의 중재 결과 자료를 사용할 수 있었는가?	“거의 모든”은 “연구결과를 신뢰할 만큼 충분하다”로 해석되어야 하며, 적절한 비율은 맥락에 따라 달라질 수 있다. 어떤 상황에서는 두 중재 군 간 상당히 공통적으로 관심 사건이 발생한다면, 연구대상자의 95%(혹은 가능한 90%) 자료를 이용할 수 있는 것으로 충분할 수 있다. 이것의 한가지 측면은 문헌고찰 저자가 해당 연구의 정확한 분석 계획을 찾으려고 이상적으로 시도한다는 것이다.	Y / PY / PN / N / NI
5.2 중재 상태에 대한 결측 자료로 인해 연구대상자가 제외되었는가?	중재상태가 누락되면 문제가 될 수 있다. 이를 위해서 의도한 연구 표본이 명확해야 하나, 실제로는 그렇지 않을 수 있다.	Y / PY / PN / N / NI
5.3 분석에 필요한 다른 변수에 대한 결측자료로 인해 연구대상자가 제외되었는가?	이 질문은 특히 분석에서 통제된 교란요인에 대한 정보가 누락되었기 때문에 분석으로부터 제외된 연구대상자와 관련이 있다.	Y / PY / PN / N / NI
5.4 5.1이 PN/N 혹은 5.2나 5.3이 Y/PY인 경우: 결측된 연구대상자의 비율과 원인이 중재군간 유사한가?	이 질문은 (i)결측된 관찰값의 비율의 차이 혹은 (ii) 결측된 관찰값에 대한 원인의 차이가 다루고 있는 질문에 대한 우리의 답변 능력에 실질적으로 영향을 줄 수 있는 여부를 도출하는 것을 목적으로 한다. “비슷하다”는 우연에 의해 기대되는 것만큼 중재 군 간 약간의 불일치한 것을 의미한다.	NA / Y / PY / PN / N / NI
5.5 5.1이 PN/N 혹은 5.2나 5.3이 Y/PY인 경우: 결측치에 관계없이 연구결과가 견고(robust)하다는 근거가 있는가?	견고성(robustness)에 대한 근거는 결측치를 다룬 분석 방법과 조사자에 의해 수행된 민감도 분석 혹은 때때로 체계적 문헌고찰 수행자에 의해 시행된 추가 분석 여부에서 따라 나타날 수 있다. 분석에 사용된 가정이 명확하고 타당한지를 평가하는 것이 중요하다. 이를 위해서는 연구 내용에 대한 지식과 통계 전문지식이 필요하다. 예를 들어, 다중 대체법(multiple imputation)과 같은 통계 방법의 사용이 적절한 답변을 보장하는 것은 아니다. 문헌고찰 저자는 비교를 위해 나이브한 (완료된 사례) 분석을 찾아야 하며, 완료된 사례와 다중 대체법 기반의 결과가 다른 경우 분석 방법의 타당성에 대해 주의 깊은 평가가 이어져야 한다.	NA / Y / PY / PN / N / NI
비독립위험 판단	표 2를 보시오.	낮음 / 중등도 / 높음 / 매우 높음 / 정보 없음
선택 사항: 결측치에 의해 예상되는 비독립의 방향은 무엇인가?	비독립의 방향을 예측하는 것이 가능하다면, 이를 기술하는 것이 도움이 될 것이다. 그 방향이 무효점으로 향하던지(혹은 멀어지는 쪽으로 향하던지) 혹은 중재 중 어느 한쪽으로 유리하게 작용.	실험 중재 선호 / 비교 중재 선호 / 무효점으로 향함 / 무효점에서 멀어짐 / 예측할 수 없음

⑥ 중재결과 측정 비돌림

신호 질문	세부 설명(elaboration)	응답 선택
6.1 중재결과 측정이 제공받은 중재의 지식에 의해 영향을 받았을 수 있는가?	일부 중재결과 측정에는 평가자 판단을 무시할 만한 것들이 있다. 예를 들어 모든 원인 사망률 또는 반복 불가능한 자동화된 실험실 지표 같은 것들이다. 이러한 중재결과 측정으로 인한 비돌림위험은 낮을 것으로 예상된다.	Y / PY / PN / N / NI
6.2 중재결과 평가자는 연구대상자들이 받은 중재를 알고 있는가?	만약 중재결과 평가자가 중재 상태에 눈가림이 되어 있다고 한다면, 이 질문에 대한 대답은 '아니다' 일 것이다. 다른 상황에서, 임상 연구자들에 의해 적극적으로 눈가림을 시행하지 않았다고 하더라도, 중재결과 평가자들이 연구대상자들이 받은 중재를 인지하지 못할 수도 있을 것이다. 이런 경우 이 질문에 대한 답은 '아니다'가 될 것이다. 예를 들어 설문지와 같이 중재결과를 자가 보고하는 연구에서, 중재결과 평가자는 연구대상자가 된다. 관찰 연구에서 연구대상자가 직접 중재결과를 보고할 때 이 질문에 대한 답은 일반적으로 '그렇다'가 될 것이다.	Y / PY / PN / N / NI
6.3 중재결과 평가 방법은 중재 기간에 유사했는가?	유사한 평가 방법(즉, 자료 수집)에는 동일한 중재결과 추출 방법과 임계치, 동일 시점, 동일 정의 및 동일 측정치 같은 것들이 있다.	Y / PY / PN / N / NI
6.4 제공받은 중재와 관련된 중재결과 측정 시 체계적 오류가 있었는가?	이 질문은 중재결과와 차등적 오분류에 관한 것이다. 중재 결과 측정에 체계적 오류가 있는 경우, 중재 혹은 중재-중재 결과 관계의 교란 변수와 관련이 있다면 비돌림이 발생할 수 있다. 이는 일반적으로 중재 결과 평가자가 제공받은 중재를 인지하고 있거나 혹은 중재 결과 평가 방법의 비교 불가능성(non-comparability)에 기인하지만 이러한 통제에도 불구하고 발생하는 차등적 오분류의 예가 있다.	Y / PY / PN / N / NI
비돌림위험 판단	표 2를 보시오.	낮음 / 중등도 / 높음 / 매우 높음 / 정보 없음
선택 사항: 중재결과 측정에 의해 예상되는 비돌림의 방향은 무엇인가?	비돌림의 방향을 예측하는 것이 가능하다면, 이를 기술하는 것이 도움이 될 것이다. 그 방향이 무효점으로 향하던지(혹은 멀어지는 쪽으로 향하던지) 혹은 중재 중 어느 한쪽으로 유리하게 작용.	실험 중재 선호 / 비교 중재 선호 / 무효점으로 향함 / 무효점에서 멀어짐 / 예측할 수 없음

⑦ 보고된 연구결과 선택의 비틀림

신호 질문	세부 설명(elaboration)	응답 선택
보고된 효과 추정치는 다음과 같은 연구결과에 기초하여 선택되었을 가능성이 있는가?		
7.1 중재결과 영역 내에서 여러 개의 중재결과 측정	특정 중재결과 영역의 경우, 서로 다른 측정에 대해 다수의 효과 추정치를 생성할 수 있다. 다수의 측정이 이루어졌지만 하나 또는 일부만 보고된 경우, 연구결과에 따른 선택적 보고의 위험이 있다.	Y / PY / PN / N / NI
7.2 중재-중재결과 관계에서의 다중 분석	효과 분석에 대한 비무작위 연구로부터 사용된 자료의 한계 때문에(교란의 통제 필요, 상당한 결측치 등), 통계 분석자들은 이러한 제한점을 해결하기 위해 다른 분석 방법을 수행할 수 있다. 예를 들어, 비보정 vs 보정 모형; 최종값 vs 기저 상태에서부터 변화량 vs 공분산 분석; 변수 변환; 다양한 역할을 가진 연속형 변수를 범주형 변수로 변환; 다양한 공변량 집합에 의한 보정; 결측치를 처리하기 위한 다른 전략 등이 그것이다. 복수의 방법을 적용하면 특정 결과 측정에 대한 복수의 효과 추정치가 생성된다. 만약 통계분석자들이 적용할 방법을 사전에 정의하지 않고, 다수의 추정치가 생성됐음에도 불구하고 하나 혹은 하위집단 중재 결과만 보고한 경우, 결과에 기초한 선택적 보고의 위험이 있다.	Y / PY / PN / N / NI
7.3 다른 하위군	일상적인 자료원으로부터 종종 사용할 수 있는 큰 코호트의 경우 여러 하위군에 대해 다수의 효과 추정치를 생성하거나 원래 코호트의 다양한 비율을 단순히 생략할 수 있다. 다수의 추정치가 생성되지만 하나 또는 일부만 보고되는 경우 연구결과에 따른 선택적 보고의 위험이 있다.	Y / PY / PN / N / NI
비틀림위험 판단	표 2를 보시오.	낮음 / 중등도 / 높음 / 매우 높음 / 정보 없음
선택 사항: 보고된 연구결과와 선택에 의해 예상되는 비틀림의 방향은 무엇인가?	비틀림의 방향을 예측하는 것이 가능하다면, 이를 기술하는 것이 도움이 될 것이다. 그 방향이 무효점으로 향하던지(혹은 멀어지는 쪽으로 향하던지) 혹은 중재 중 어느 한쪽으로 유리하게 작용.	실험중재선호/비교중재선호/무효점으로 향함/무효점에서 멀어짐/예측할 수 없음

전체 비틀림위험 판단

비틀림위험 판단	표 3을 보시오.	낮음 / 중등도 / 높음 / 매우 높음 / 정보 없음
선택 사항: 중재 결과에 대해 전반적으로 예상되는 비틀림의 방향은 무엇인가?		실험 중재 선호 / 비교 중재 선호 / 무효점으로 향함 / 무효점에서 멀어짐 / 예측할 수 없음

표 1. ROBINS-I에서 비돌림위험 판단 : 중재 전 및 중재 중 영역

판단	교란으로 인한 비돌림	연구대상자 선택의 비돌림	중재 분류의 비돌림
낮은 비돌림위험 (잘 수행된 무작위 임상시험과 유사 함)	교란이 예상되지 않았음	(i) 목표 임상시험에 적절할 것 같은 연구 대상자가 연구에 포함되었음 그리고 (ii) 각 연구대상자에 대해 추적 조사의 시작과 중재의 시작이 일치하였음	(i) 중재 상태가 잘 정의됨 그리고 (ii) 중재의 정의는 중재 시 수집된 정 보에만 근거함
중등도 비돌림위 험 (비무작위 연구에 적합하지만 잘 수 행된 무작위 임상 시험과 유사하지 않음)	(i) 교란이 예상되 었고, 모든 알려진 중요한 교란 영역 이 적절하게 측정 되고 통제되었음 그리고 (ii) 중요한 영역에 대한 측정의 신뢰 도와 타당도는 충 분하였으며, 심각 한 잔류 교란이 예 상되지 않음	(i) 연구대상자 선택이 중재와 중재결과 에 관련이 있을 수 있음 그리고, 선택 비돌림에 대해 적절한 방 법으로 보정하였음 혹은, (ii) 추적 조사의 시작과 중재의 시작이 모든 연구대상자에서 일치하지는 않음 그리고 (a) 이에 해당하는 연구대상자의 비 율이 너무 낮아 중요한 비돌림을 유 발할 수 없음 혹은 (b) 선택 비돌림에 대해 적절한 방법 으로 보정됨 혹은 (c) 문헌고찰 저자가 중재 효과에 대 한 발생비(rate ratio)가 시간에 지남 에 따라 일정하다고 확신함	(i) 중재 상태가 잘 정의됨. 그리고 (ii) 중재 상태의 배 정 시 일부 양상이 후향적으로 결정됨
높은 비돌림위험 (몇 가지 중요한 문제가 있음)	(i) 하나 이상의 알 려진 중요한 영역 이 적절하게 측정 되지 않았거나 통 제되지 않았음 또는 (ii) 중요한 영역에 대한 측정의 신뢰 도 또는 타당도가 심각한 잔류 교란 을 예상할 수 있을 정도로 낮았음	(i) 연구대상자 선택이 중재와 중재결과에 관련이 있었지만 매우 강력하지는 않음 그리고 분석에서 보정될 수 없음 혹은 (ii) 추적 조사의 시작과 중재의 시작이 일치하지 않음 그리고 잠재적으로 중요한 추적관찰 시간이 분 석에서 제외됨 그리고 발생비(rate ratio)가 시간에 따라 일정 하지 않음	(i) 중재 상태가 잘 정의되지 않음 혹은 (ii) 중재 상태의 배 정 시 주요 측면이 중재 결과에 대한 지식에 의해 영향을 받을 수 있는 방식 으로 결정됨

판단	교란으로 인한 비돌림	연구대상자 선택의 비돌림	중재 분류의 비돌림
매우 높은 비돌림 위험 (중재 효과에 대한 근거로 사용하기에는 너무 문제가 많음)	(i) 본질적으로 통제할 수 없는 교란 또는 (ii) 음성 비교 중재의 사용이 측정되지 않은 교란을 강력하게 시사함	(i) 연구대상자 선택은 중재와 중재결과에 매우 밀접하게 관련이 있었음 그리고 분석에서 보정될 수 없음 혹은 (ii) 상당한 양의 추적 관찰 시간의 양이 분석에서 누락된 것으로 보임 그리고 발생비(rate ratio)가 시간에 따라 일정하지 않음	(일반적이지 않음) 중재 상태의 오분류의 양이 극도로 많음 (예: 비정상적으로 강한 회상 비돌림)
정보 없음 (이 영역에서 비돌림위험에 대한 판단을 하기에)	교란이 있는지에 대한 정보가 없음	연구대상자 선택이나, 추적 조사의 시작과 중재의 시작이 일치하는지 여부에 대한 정보가 보고되지 않음	중재에 대한 정의나 중재 상태의 정보원에 대한 설명이 보고되지 않음

표 2. ROBINS-I에서 비풀림위험 판단: 중재 후 영역

판단	의도한 중재에서 이탈되어 나타나는 비풀림	결측치로 인한 비풀림	중재결과 측정 비풀림	보고된 연구결과 선택의 비풀림
낮은 비풀림 위험 (잘 수행된 무작위 임상시험과 유사함)	<p>중재 배정의 효과: (i) 의도한 중재로부터의 이탈이 일상적인 임상상을 반영함 혹은 (ii) 일상적인 임상에서의 이탈이 중재결과에는 영향을 미칠 것 같지 않음</p> <p>중재 시작 및 준수의 효과: 중요한 공동 중재가 중재 군간에 균형을 이루었고, 중재 결과에 영향을 줄 수 있는 의도한 중재(적용이나 준수 면에서)로부터 이탈은 없었음</p>	<p>(i) 자료가 합리적으로 완료됨 혹은 (ii) 결측된 연구대상자의 비율이나 이유가 중재군간에 비슷하였음 혹은 (iii) 분석을 통해 결측치를 해결하였고 비풀림위험을 제거한 것으로 보임</p>	<p>(i) 중재결과 평가 방법이 군간에 유사함 그리고 (ii) 중재결과 측정이 연구대상자가 받은 중재에 대한 지식에 의해 영향을 받지 않았거나 (즉 객관적임), 중재결과 평가자가 연구대상자가 받은 중재를 알지 못했음 그리고 (iii) 중재결과 측정에서의 오차가 중재 상태와 관련이 없음</p>	<p>모든 보고된 연구결과가 모든 의도한 중재 결과, 분석 및 하위 코호트에 해당한다는 명확한 근거가 있음(보통 사전에 등록된 프로토콜 또는 통계분석 계획을 통해)</p>
중등도 비풀림 위험 (비무작위 연구에 적합하지만 잘 수행된 무작위 임상시험과는 유사하지 않음)	<p>중재 배정의 효과: 일상적인 임상에서의 이탈이 있으나, 그 중재 결과에 미치는 영향이 미미할 것으로 예상됨</p> <p>중재 시작 및 준수의 효과: (i) 의도한 중재와 이탈이 있으나, 그 중재 결과에 미치는 영향이 미미할 것으로 예상됨 혹은 (ii) 중요한 공동 중재가 중재 군간에 균형을 이루지 않았거나, 의도한 중재와의 이탈이 있어서 중재 결과에 영향을 줄 것으로 보임(적용 및 준수 면에서) 그리고 중재 결과에 대한 영향이 있을 것 같은 이탈을 감안하여, 중재 시작 및 준수의 영향을 추정하는데 분석이 적절하였음(적용, 준수, 그리고 공동 중재 면에서)</p>	<p>(i) 결측된 연구대상자의 비율과 이유가 군간에 약간 다름 그리고 (ii) 분석으로 결측치에서 나타나는 비풀림위험을 제거하였을 것 같지 않음</p>	<p>(i) 중재결과 평가 방법이 중재 군간에 유사함 그리고 (ii) 중재결과 측정이 연구대상자가 받은 중재에 대한 지식에 의해 영향을 최소로만 받았음 그리고 (iii) 중재결과 측정에서의 오차가 중재 상태와 최소한으로만 관계가 있음</p>	<p>(i) 중재결과 측정과 분석이 사전 계획과 일치함 혹은 명확하게 정의되고 내부적, 외부적으로 일관성이 있음 그리고 (ii) 여러 분석 중에서 보고된 분석을 선택한다는 표시가 없음 그리고 (iii) 연구결과에 기초한 분석 및 보고를 위한 코호트 또는 하위군을 선택한다는 표시가 없음</p>

판단	의도한 중재에서 이탈되어 나타나는 비뚫림	결측치로 인한 비뚫림	중재결과 측정 비뚫림	보고된 연구결과 선택의 비뚫림
<p>높은 비뚫림 위험 (몇 가지 중요한 문제가 있음)</p>	<p>중재 배정의 효과: 중재 군간에 불균형이 있었고 중재 결과에 영향을 미칠 가능성이 있을 만한, 일상적인 임상에서의 이탈이 있었음</p> <p>중재 시작 및 준수의 효과: (i) 중요한 공동 준재가 중재 군간에 균형을 이루지 못하거나, 중재 결과에 영향을 미칠 가능성이 있는 의도한 중재로부터의 이탈이 있음(적용 및 준수 면에서) 그리고 (ii) 중재 결과에 대한 영향이 있을 것 같은 이탈을 감안하여, 중재 시작 및 준수의 영향을 추정하는데 분석이 적절하지 않았음(적용, 준수, 그리고 공동 중재 면에서)</p>	<p>(i) 결측된 연구대상자의 비율이 중재 간에 상당히 큼 혹은 결측의 이유가 중재 간에 상당히 다름 그리고 (ii) 분석으로 결측치로 인한 비뚫림위험을 제거할 것 같지 않음 혹은 결측치가 분석에서 적절하지 않게 다루어짐 혹은 결측치가 적절한 분석을 통해 비뚫림위험을 제거할 수 없는 특성을 가지고 있음</p>	<p>(i) 중재결과 측정 방법이 중재 군간에 유사하지 않음 혹은 (ii) 중재결과 측정이 주관적임(즉, 연구대상자가 받는 중재에 대한 지식에 의한 영향에 취약함) 그리고 연구대상자가 어떤 중재를 받는지 알고 있는 평가자에 의해 중재결과가 평가됨 혹은 (iii) 중재결과 측정에서의 오차가 중재 상태와 관련이 있음</p>	<p>(i) 중재결과가 논문의 방법 부분과 결과 부분, 혹은 해당 연구의 다른 논문에서 다르게 정의되어 있음 혹은 (ii) 여러 분석 중에서 선택적으로 보고된 높은 위험이 있음 혹은 (iii) 코호트나 하위군이 분석을 위해 더 큰 연구에서 선택되었으며, 연구결과에 기초해서 보고된 것으로 보임</p>
<p>매우 높은 비뚫림위험 (중재 효과에 대한 근거로 사용하기에는 너무 문제가 많음)</p>	<p>중재 배정의 효과: 중재 군간에 불균형이 있었고 중재 결과에 영향을 미칠 것 같은 정도로, 통상진료에서 상당한 이탈이 있었음</p> <p>중재 시작 및 준수의 효과: (i) 그룹 전체에서 중요한 공동 중재에서 상당한 불균형이 있거나, 중재 결과에 영향을 미칠 가능성이 높은 의도한 중재와 상당한 이탈이 있음(적용 및/혹은 준수 측면에서)</p>	<p>(i) (일반적이지 않음) 결측치가 있는 연구대상자의 중재 사이에 치명적인 차이점이 있음 그리고 (ii) 결측치가 적절한 분석을 통해서 다루어지지 않거나 다룰 수가 없음</p>	<p>중재결과 평가하는 방법이 중재군 간에 합리적으로 비교할 수 없을 정도로 달랐음</p>	<p>(i) 연구결과를 선택적으로 보고하였다는 근거나 강한 의심이 있음 그리고 (ii) 보고된 연구결과와는 상당히 다른 보고되지 않은 연구결과가 있는 것 같음</p>

판단	의도한 중재에서 이탈되어 나타나는 비플립	결측치로 인한 비플립	중재결과 측정 비플립	보고된 연구결과 선택의 비플립
	그리고 (ii) 중재 결과에 대한 영향이 있을 것 같은 이탈을 감안하여, 중재 시작 및 준수의 영향을 추정하는데 분석이 적절하지 않았음(적용, 준수, 그리고 공동 중재 면에서)			
정보 없음 (이 영역에서 비플립위험에 대한 판단을 하기에)	의도한 중재로부터 이탈이 있었는지에 대해 보고된 정보가 없음	결측치나 결측되었을 가능성이 있는지에 대해 보고된 정보가 없음	중재결과 평가 방법에 대해 보고된 정보가 없음	판단을 하기에 정보가 너무 없음(예: 해당 연구에 대해서 초록만 이용할 수 있는 경우)

표 3. ROBINS-I에서 영역 수준과 전체적인 비플립위험 판단 해석하기

판단	각 영역 내에서	영역 간	기준
낮은 비플립 위험	이 영역과 관련해서 잘 수행된 무작위 임상시험과 유사함	잘 수행된 무작위 임상시험과 유사함	모든 영역에서 낮은 비플립위험으로 판단됨.
중등도 비플립 위험	이 영역과 관련해서 비무작위 연구로 적합하지만 잘 수행된 무작위 임상시험과는 유사하지 않음	비무작위 연구로 적합하지만 잘 수행된 무작위 임상시험과 유사하지 않음	모든 영역에서 낮거나 중간 비플립위험으로 판단됨
높은 비플립 위험	이 영역에서 몇 가지 중요한 문제가 있음	몇가지 중요한 문제를 가지고 있음	최소한 하나의 영역에서 높은 비플립위험으로 판단되나 어떤 영역에서도 매우 높은 비플립위험은 아님
매우 높은 비플립 위험	중재의 효과에 대한 일반적인 근거를 제공하기엔 이 영역에서 너무 문제가 많음	일반적인 근거를 제공하기엔 너무 문제가 많아서 어떤 합성에도 포함시켜서는 안됨	최소한 하나의 영역에서 매우 높은 비플립위험으로 판단됨
정보 없음	이 영역에서 비플립위험을 판단할 만한 기초가 되는 정보가 없음	비플립위험 평가를 판단할 만한 기초 정보가 없음	높거나 매우 높은 비플립위험이 있다는 명백한 명시가 없고, 하나 이상의 주요 비플립 영역에 정보가 부족함(이에 대한 판단 필요)

라. 국문 ROBINS-I 사용자 매뉴얼

1) 개요

ROBINS-I는 비무작위 임상시험(NRSIs)에 대한 방법론적 질을 평가하기 위한 도구이다. 이를 위해서 평가자는 방법론적 지식뿐만 아니라 내용에 대한 전문적 지식이 필요하다.

평가 단계는 첫째, NRSIs에 대한 교란, 공동 중재와 같은 잠재적 문제를 확인하고 비뚫림 위험에 처할 수 있는 모든 방법을 고려하여 예비 고려사항을 도출한다. 둘째, 동일 대상자 및 중재에 대한 가상의 무작위 임상시험(목표 무작위 임상시험)을 설정하여 관련된 비뚫림 위험을 평가한다.

(1) 1단계: 프로토콜 단계- 연구질문 구체화하기, 교란 영역 작성하기, 공동중재 작성하기

- 연구질문에서는 인구집단, 실험 중재군, 대조군 및 관심 결과를 다룬다. 비교군 중재는 무중재, 일상적인 진료, 혹은 대안적인 중재가 모두 가능하다.
- 교란 영역은 어떤 사람이 하나 또는 다른 관심 중재를 받는지를 예측하는 예후 인자이다.
- 공동-중재는 어떤 사람이 관심 중재의 시작과 함께 혹은 그 이후에 받을 수 있는 중재 혹은 노출이며, 이는 그들이 받는 중재와 관련이 있고 관심 결과에 대해 예후적이다.
- 발생할 수 있는 문제를 알아채는 것은 부분적으로 연구 주제 전문가의 문헌 지식에 기초할 것이다. 또한 연구팀은 이해 상충이 전문가의 판단에 영향을 미칠 수 있는지 여부를 다루어야 한다.
- 교란 영역과 공동-중재는 문헌고찰 그룹의 구성원인 연구 주제 전문가의 지식을 통해서, 그리고 문헌고찰 초창기 범주화를 통해서 정해질 수 있다.

(2) 2단계: 개별연구 평가 셋업하기, 개별연구 교란 영역과 공동 중재 평가, 비뚫림위험 평가

- 목표 임상시험은 비뚫림이 없는 경우 고려중인 NRSI의 연구결과와 동일한 가설적인 무작위 임상시험이다. 가설일 뿐이기 때문에 윤리적인 문제와 실제 가능한지에 대해서는 고려할 필요가 없다
- 목표 임상시험에서 어떤 특정 연구문제에 대한 관심 효과는 추적 관찰기간 동안 받는 중재의 정도에 관계없이 기저 상태에서 중재에 배정하는 효과이거나 프로토콜에서 명시한 중재의 시작 및 준수의 효과이다. 중재의 배정 효과는 연구대상자가 무작위 배정된 중재 그룹에서 분석되는, 배정된 대로(ITT:intention-to-treat) 분석을 통해 추정

될 수 있다. 프로토콜 준수효과(per protocol effect)는 중재를 시작하고 준수하는 효과이며, 중재의 유해효과(또는 의도하지 않은 효과)를 고려할 때 관심을 가질 만한 효과이기도 하다.

- 임상시험 프로토콜에 명시된 중요한 교란요인과 공동 중재가 현재 연구에서 측정 또는 관리되는지, 그리고 추가적인 교란요인과 공동중재를 확인한다.
- 7가지 비풀림 영역은 아래와 같다.

영역		관련 용어	설명
중재 전	교란으로 인한 비풀림	배정 비풀림; 증례-혼합 비풀림; Channelling 비풀림	기저 상태 교란 시간 변동 교란 요인
	연구대상자 선택의 비풀림	개시 비풀림(inception bias), 시간 지연 비풀림(lead time bias), 불멸의 시간 비풀림 (immortal time bias)	중재와 결과에 모두 연결되어 있는 일부 적격 연구 대상자 혹은 일부 참가자의 초기 추적시간 혹은 일부 결과 발생을 배제하는 경우
중재 중	중재 분류의 비풀림	오분류 비풀림; 정보 비풀림; 회상 비풀림; 측정 비풀림; 관찰자 비풀림	중재 상태의 차등적 또는 비차등적 오분류
중재 후	의도한 중재에서 이탈되어 나타나는 비풀림	실행 비풀림; 시간-변이 교란	실험 중재군과 대조군 사이에 제공된 관리가 체계적으로 다를 때
	결측치로 인한 비풀림	탈락 비풀림	포함되었던 연구대상자가 추적 관찰에서 누락되어 발생하는 비풀림 중재 상태나 교란 변수와 같은 변수의 정보가 누락되어 발생한 비풀림
	중재결과 측정 비풀림	결과 확인(detection) 비풀림; 회상 비풀림; 정보 비풀림; 오분류 비풀림; 관찰자 비풀림; 측정 비풀림	결과의 측정에서 차등적 또는 비차등적 오분류에 의해 발생하는 비풀림
	보고된 연구결과 선택의 비풀림	결과 보고 비풀림; 분석 보고 비풀림	연구결과에 따라 선택적으로 결과를 보고

- 비플립 영역 내 신호 질문의 응답은 (1) 그렇다 (2) 아마도 그렇다 (3) 아마도 아니다 (4) 아니다 (5) 정보 없음으로 **녹색으로 밑줄 친 대답**은 비플립위험이 낮을 가능성이 있는 표시이고 **빨간색**으로 표시된 것은 비플립이 발생할 가능성이 있는 표시이다. “정보 없음”의 범주는 판단을 허용하기에 충분하지 않은 자료가 보고된 경우에만 사용해야 한다.

- 각 영역의 비플립위험 판단에서 응답 옵션은 다음과 같다.

- ① 낮은 비플립위험(연구는 이 영역에 관련하여 잘 수행된 무작위 임상시험과 유사하다)
- ② 중등도 비플립위험(연구는 이 영역과 관련하여 비무작위 연구에 적합하지만 잘 수행된 무작위 임상시험과 유사하다고 간주할 수 없다)
- ③ 높은 비플립위험(연구는 이 영역에서 몇 가지 중요한 문제가 있다)
- ④ 매우 높은 비플립위험(연구는 이 영역에서 중재의 효과에 대해 유용한 근거를 제공하기에는 너무 문제가 많다)
- ⑤ 이 영역에 대한 비플립위험을 판단하기 위한 근거가 되는 정보가 없다.

- 프리 텍스트 박스: 비플립위험 판단과 함께 뒷받침 하는 이유를 기술한다. 특히 높은 비플립위험, 매우 높은 비플립위험은 필수적으로 작성한다.

- 비플립의 방향: 영역 및 전체에 대한 비플립 방향을 판단하는 선택적 구성요소가 포함되어 있다. 일부 영역의 경우, 비플립은 무효점(null)을 향하거나 무효점에서 멀어지는 것으로 매우 쉽게 생각할 수 있다. 예를 들어 통계적으로 유의하지 않은 연구결과의 선택적 비보고에 대한 의혹이 있다면 무효점에서 멀어지는 비플립을 암시함. 다른 일부 영역(특히 교란효과, 선택 비플립, 차등적 오분류와 같은 측정 비플립 형태)은 무효점과 관련하여 판단하지 말고 효과 추정치의 증가 또는 감소로 생각해야 한다. 예를 들어 교란 비플립은 실제 위험비(risk ratio)가 1보다 크면 무효점으로 향하게 하여 효과 추정치를 감소시킨다. 단, 문헌고찰 저자가 판단할 명확한 논리적 근거가 없는 경우 비플립의 방향을 추측하려고 시도해서는 안된다.

- 전체적 비돌림위험 판단

응답 옵션	기준
<p><u>낮은 비돌림위험</u>(연구는 이 영역에 관련하여 잘 수행된 무작위 임상시험과 유사하다)</p>	<p>연구는 모든 영역에서 낮은 비돌림위험에 있다고 판단된다.</p>
<p><u>중등도 비돌림위험</u>(연구는 이 영역과 관련하여 비무작위연구에 적합하지만 잘 수행된 무작위 임상시험과 유사하다고 간주할 수 없다)</p>	<p>연구는 모든 영역에서 낮거나 중등도의 비돌림위험에 있다고 판단된다.</p>
<p><u>높은 비돌림위험</u>(연구는 이 영역에서 몇 가지 중요한 문제가 있다)</p>	<p>연구는 최소한 하나의 영역에서 높은 비돌림위험에 있다고 판단되나, 어떤 영역에서도 매우 높은 비돌림위험에 있지는 않다.</p>
<p><u>매우 높은 비돌림위험</u>(연구는 이 영역에서 중재의 효과에 대해 유용한 근거를 제공하기에는 너무 문제가 많다)</p>	<p>연구는 최소한 하나의 영역에서 매우 높은 비돌림위험에 있다고 판단된다.</p>
<p>이 영역에 대한 비돌림위험을 판단하기 위한 근거가 되는 <u>정보가 없다</u>.</p>	<p>연구가 높거나 매우 높은 비돌림위험에 있다는 명확한 내용이 없고, 하나 이상의 핵심 비돌림 영역에서 정보가 부족하다.</p>

※ 각 세부 단계에 대한 자세한 설명은 부록을 참조하십시오.

2) 평가 적용의 예

가) 평가에 사용한 문헌

Tokarek T, et al. Clinical outcomes in patients after surgical and transcatheter aortic valve replacement. Pol Arch Med Wewn, 2015;125(10):755-64.

나) 1단계: 프로토콜 단계

문헌고찰 질문을 명확하게 기술하시오.

연구대상자	patients who underwent SAVR and transfemoral TAVR for Severe Aortic Stenosis Treatment
실험 중재	transfemoral TAVR(transcatheter aortic valve replacement)
비교 중재	SAVR
중재결과	<ul style="list-style-type: none"> - primary end-points : death from any cause and major adverse cardiac and cerebrovascular events(MACCE) at 1 year <li style="padding-left: 20px;">* MACCE: death from any cause, stroke, myocardial infarction, percutaneous coronary intervention and CABG - secondary endpoints : cerebrovascular accidents, acute myocardial infarction, repeat hospitalization due to cardiac reasons, and acute heart failure

모든 혹은 대부분의 연구에서 중요한 교란 영역의 목록을 작성하시오.

- 환자 기본 특성 : 성, 연령, 체질량 지수, 흡연, 과거력(당뇨병, 고혈압, 심각한 간질환/간경변, COPD, 신경학적 질환, 말초동맥질환, 심근경색, 불안정형 협심증, 폐동맥 고혈압, 급성 신생물 질환, 만성신장질환), 과거 수술력(과거 심장수술, PCI, 풍선 대동맥판 성형술 기왕력)
- 시술전 환자 상태 : NYHA 기능분류, STS score, LogEuroSCORE I / II, 활동성 심내막염, 대동맥판 형태, 승모판 역류증, 대동맥판 역류증, 좌심실구혈률, 노쇠 점수

중재군에 따라 다를 수 있고 중재결과에 영향을 줄 수 있는 공동 중재의 목록을 작성하시오.

- 항고혈압약제, 당뇨병 치료제(인슐린 포함), 항응고제, 심장질환 치료제 등
- 행동변화중재, 생활습관 변화 중재(예: 운동, 금연중재, 식이요법 등)

다) 2단계: 개별 연구

연구에 특정한 목표 임상시험을 기술하십시오.

연구설계	개별 무작위
연구대상자	Severe Aortic Stenosis
실험 중재	transfemoral TAVR
비교 중재	SAVR

이 연구에 대한 리뷰 팀의 목적은?

- 중재에 대한 배정의 효과를 평가함 중재의 시작 및 준수의 효과를 평가함

중재결과를 기술하십시오.

비돌림위험평가를 시행하는 중재결과를 구체적으로 기술하십시오(일반적으로 결과 요약표에 표시되어 있는 것 중). 이것이 중재의 이득인지 위해인지 명시할 것

death from any cause and major adverse cardiac and cerebrovascular events(MACCE)

평가할 수치형 연구결과를 구체적으로 기술하십시오.

분석이 여러 번 시행된 경우 어떤 결과인지(예: RR=1.52 (95% CI 0.83-2.77)), 어디에서 자료를 얻었는지 명시할 것(예: 표, 그림, 단락)

- Death from any cause HR 0.99, 95% CI 0.72-1.35, p=0.936
- MACCE HR 1.03, 95% CI 0.78-1.36, p=0.831
- 1년 시점에서의 값도 제시됨(표4 참조)

교란요인에 대한 사전 고려사항

중요한 교란 영역에 대한 행을 작성하십시오. i) 프로토콜에 작성된 것, ii) 이 연구의 세팅과 관련이 있거나 저자가 잠재적으로 중요하다고 간주하는 것

“중요한” 교란 영역은 본 연구의 맥락에서 보정이 중재의 추정 효과에 임상적으로 중요한 변화를 초래할 것으로 예상되는 영역이다. “타당도”는 교란 변수 또는 변수가 영역을 완전히 측정하는지 여부를 나타내며, “신뢰도”는 측정의 정밀도를 나타낸다(측정 오류가 클수록 신뢰도가 떨어짐).

(i) 문헌고찰 프로토콜에 작성된 교란 영역

교란 영역	측정된 교란 변수	이 변수에 대한 통제가 불필요하다는 근거가 있는가?*	이 변수는 교란 영역을 타당하고 신뢰성 있게 측정하였는가?	선택 사항: 이 변수에 대한 보정 실패는 실험 중재나 비교 중재 중 어느 쪽에 유리할 것으로 예상되는가?
기본 특성	성	아니다	그렇다	정보 없음
	연령	아니다		실험 중재 선호
	체질량 지수	아니다		정보 없음
	흡연	아니다	아니다	정보 없음
환자상태 (수술 위험 등)	NYHA 기능분류, STS score, LogEuroSCORE I / II	아니다	그렇다	실험 중재 선호
	활동성 심내막염, 대동맥판 형태, 승모판 역류증, 대동맥판 역류증, 좌심실 구혈률, 노쇠 점수	아니다	그렇다	실험 중재 선호
과거력	심장수술/PCI/풍선 대동맥판 성형술 기왕력	아니다	그렇다	실험 중재 선호
동반질환	당뇨병, 고혈압, 심각한 간 질환/간경변, COPD, 신경학적 질환, 말초동맥질환, 심근경색, 불안정형 협심증, 폐동맥 고혈압, 급성 신생물 질환, 만성신장질환	아니다	그렇다	실험 중재 선호

(ii) 이 연구의 세팅과 관련되거나 저자가 중요하다고 판단한 추가 교란 영역

교란 영역	측정된 교란 변수	이 변수에 대한 통제가 불필요하다는 근거가 있는가?*	이 변수는 교란 영역을 타당하고 신뢰성 있게 측정하였는가?	선택 사항: 이 변수에 대한 보정 실패는 실험 중재나 비교 중재 중 어느 쪽에 유리할 것으로 예상되는가?
환자상태	Creatinine	아니다	그렇다	실험 중재 선호
	투석치료	아니다		실험 중재 선호
	알부민	아니다		정보 없음
	헤모글로빈	아니다		실험 중재 선호

환자상태	산소투여 여부	아니다	그렇다	실험 중재 선호
	Critical preoperative state	아니다	정보 없음	정보 없음

* 특정 연구의 맥락에서, 일부 변수는 교란 요인이 아니어서 분석에 포함되지 않을 수 있다. (a) 중재 결과를 예측하지 않는 경우, (b) 중재를 예측하지 않는 경우, (c) 보정이 1차 매개 변수(파라미터)의 추정된 효과와 거의 또는 전혀 차이가 없는 경우. “통계적으로 유의미한 연관성이 없음”은 “예측하지 않음”과는 같지 않음에 주의하십시오.

공동 중재에 대한 사전 고려사항

중요한 공동 중재에 대한 행을 작성하십시오. (i) 문헌고찰 프로토콜에 작성된 것, (ii) 이 연구의 세팅과 관련이 있거나 저자가 중요하다고 간주하는 것

“중요한” 공동 중재는 본 연구의 맥락에서 보정이 중재의 추정 효과에 **임상적으로 중요한 변화를 초래할 것으로 예상되는 것이다.**

(i) 문헌고찰 프로토콜에 작성된 공동 중재

공동 중재	이 중재에 대한 통제가 불필요하다는 근거가 있는가?(예: 제공되지 않았기 때문에)	이 공동 중재는 실험 중재나 비교 중재 중 어느 쪽에 유리할 것으로 예상되는가?
항고혈압약제	아니다	정보 없음
당뇨병 치료제(인슐린 포함)	아니다	정보 없음
항응고제	아니다	정보 없음
심장질환 치료제	아니다	정보 없음
만성질환 치료제	아니다	정보 없음
생활습관 및 행동변화중재	아니다	정보 없음

(ii) 이 연구의 세팅과 관련되거나 저자가 중요하다고 판단한 추가 공동 중재

공동 중재	이 중재에 대한 통제가 불필요하다는 근거가 있는가?(예: 제공되지 않았기 때문에)	이 공동 중재는 실험 중재나 비교 중재 중 어느 쪽에 유리할 것으로 예상되는가?
없음	적용할 수 없음	적용할 수 없음

비돌림위험 평가(코호트 타입 연구)

- **그렇다(Y) / 아마도 그렇다(PY) / 아마도 아닐 것이다(PN) / 아니다(N) / 적용할 수 없음(NA) / 정보 없음(NI)**
- **녹색으로 밑줄 친 응답**은 낮은 비돌림위험, **적색의 응답**은 비돌림위험을 나타냄

1. 교란으로 인한 비돌림		
신호질문	판단근거	응답선택
1.1 이 연구에서 중재 효과의 교란이 일어날 가능성이 있는가? 1.1이 N/PN 인 경우: 연구는 교란에 의한 비돌림위험이 낮은 것으로 간주될 수 있으며, 더 이상의 신호 질문을 고려할 필요가 없다.	중요 교란변수들이 있음(목록 참고), 교란으로 인한 비돌림위험이 예상됨	PY
1.1이 Y/PY 인 경우: 시간 변동 교란요인을 평가할 필요가 있는지 결정하십시오		
1.2 제공받은 중재에 따라, 연구대상자의 추적관찰기간을 구분하여 분석하였는가? N/PN 인 경우, 기저 상태의 교란과 관련지어 대답하십시오.(1.4 ~ 1.6) Y/PY 인 경우 1.3으로 가시오.	추적기간을 구분하여 분석하지 않음	N
1.3 중재의 중단 또는 교환이 결과에 대한 예후 인자와 관련이 있을 가능성이 있는가? N/PN 인 경우, 기저선의 교란과 관련지어 대답하라.(1.4에서 1.6) Y/PY 인 경우, 기저선과 시간 변동의 교란과 관련지어 대답하라.(1.7에서 1.8)	중재 중단 또는 교환이 중재 결과에 대한 예후 인자와 관련되어 일어나기 어려운 것으로 고려됨	PN
기저 상태 교란에만 관련이 있는 질문		
1.4 저자는 모든 중요한 교란 영역을 통제하는 적절한 분석 방법을 사용하였는가?	측정된 교란요인을 통제하는 방법으로 회귀분석, 성향점수 매칭과 같은 방법이 활용됨	<u>Y</u>
1.5 1.4이 Y/PY 인 경우: 이 연구에서 사용 가능한 변수에 의해 측정된 교란 영역이 타당하고 신뢰성 있게 통제되었는가?	The propensity score includes the following variables: age; sex; previous percutaneous coronary intervention; previous balloon aortic valvuloplasty; previous cardiac surgery; diabetes; chronic obstructive pulmonary disease; smoking; previous myocardial infarction; peripheral arteriopathy; creatinine; critical preoperative state; unstable angina; neurological dysfunction; pulmonary hypertension(systolic pulmonary arterial pressure >60 mm Hg);	<u>PY</u>

	chronic liver disease; active neoplastic disease; New York Heart Association functional class; frailty score [Geriatric Status Scale]; left ventricular ejection fraction; coronary artery disease; urgency status; and mitral regurgitation.	
1.6 저자는 중재의 영향을 받을 수 있는 중재 후 변수를 통제하였는가?	중재 후 변수를 통제하지는 않음. 중재 전 기저상태의 변수를 통제에 활용함	<u>N</u>
기저 상태와 시간 변동 교란에 대한 질문		
1.7 저자는 모든 중요한 교란 영역과 시간-변동 교란을 보장하는 적절한 분석방법을 사용하였는가?		
1.8 1.7이 <u>Y/PY</u> 인 경우: 이 연구에서 이용할 수 있는 변수에 의해 측정된 교란 영역이 타당하고 신뢰성 있게 보장되었는가?		
비돌림위험 판단	교란이 예상되었고, 모든 알려진 중요한 교란 영역이 적절하게 측정되고 통제되었음 그리고 중요한 영역에 대한 측정의 신뢰도와 타당도는 충분하였으며, 심각한 잔류 교란이 예상되지 않음	중등도
선택 사항: 교란으로 인해 예상되는 비돌림의 방향은 무엇인가?	중요 교란요인이 충분히 다루어졌다고 할 수 없으므로 실험 중재군에 유리하게 작용했을 수 있을 것임	실험 중재 선호

설명	<ul style="list-style-type: none"> - 치료 결정에 영향을 미치는 요인과 관련이 거의 없는 경우 교란이 예상되지 않아, 교란으로 인한 비돌림위험이 낮은 것으로 간주될 수 있다. - 연구대상자가 중재 군간 전환이 가능한 경우 시간 변동 교란에 의해 중재와 중재 중재결과간의 연관성이 편향될 수 있다. 중재 전환이 중재 결과와 관련이 없는 경우 시간-변동 교란은 존재하지 않으며 기저 상태 교란에 대한 통제만 필요하다. - 측정된 교란요인을 통제하는 적절한 방법으로는 층화, 회귀분석, 매칭, 표준화 및 역확률 가중 (inverse probability weighting) 같은 방법들이 있다. - 교란 영역에 대한 타당하고 신뢰할 수 있는 측정도구의 목록이 프로토콜에서 정해지지만, 어떤 경우는 해당 목록을 사용하지 못할 수 있다. 저자는 특정 측정 방법의 사용을 정당화하기 위해 참고문헌을 인용할 수 있다. 저자가 타당도 혹은 신뢰도 제시 없이 교란 변수를 통제하는 경우, 측정이 주관적이 될 수 있음에 주의를 기울여야 한다. - 중재에 의해 영향을 받게 되는 중재 후 변수를 통제하는 것은 적절하지 않다. 매개 변수에 대한 통제는 중재의 직접적인 효과를 추정하고 비돌림을 유발할 수 있다. 중재와 결과의 일반적인 효과에 대해 통제하면 비돌림이 생기게 된다. - 중재의 시작 및 준수 효과를 추정하기 위해서는 시간-변동 교란에 대한 보정이 필요하다. 적절한 통계분석 방법으로는 역 확률 가중치를 기반으로 하는 방법이 있다.
-----------	---

2. 연구대상자 선택의 비돌림		
신호질문	판단근거	응답선택
2.1 중재 시작 후 관찰된 연구대상자의 특성에 기초하여 연구대상자를(또는 분석 대상) 선택하였는가? 2.1이 N/PN인 경우: 2.4로 가시오.	연구대상자는 중재 전 관찰된 특성에 따라 선택됨. 실험 중재(TAVR)와 비교 중재(SAVR) 사이의 불균형을 중재 결과에 대한 예측적인 기저 특성에 따라 통제함으로써 해결하려고 함	PN / N
2.2 2.1이 Y/PY인 경우: 선택에 영향을 준 중재 후 변수가 중재와 관련이 있을 가능성이 있었는가?		NA
2.3 2.2이 Y/PY인 경우: 선택에 영향을 준 중재 후 변수가 중재 결과 또는 중재 결과의 원인에 영향을 받을 가능성이 있는가?		NA
2.4 대부분의 연구대상자에서 추적 관찰의 시작 시점과 중재의 시작 시점이 일치하는가?	연구대상자가 중재 시작부터 추적 관찰되어 결과 보고됨	Y
2.5 2.2와 2.3이 Y/PY이거나, 2.4이 N/PN인 경우: 선택 비돌림을 교정할 수 있는 보정 방법을 사용하였는가?		NA
비돌림위험 판단	목표 임상시험에 적절할 것 같은 연구대상자가 연구에 포함되었음 각 연구대상자에 대해 추적 조사의 시작과 중재의 시작이 일치하였음	낮음
선택 사항: 연구대상자 선택에 의해 예상되는 비돌림의 방향은 무엇인가?		예측할 수 없음

설명	<ul style="list-style-type: none"> - 이 영역은 중재의 시작전에 관찰된 특성에 따라 선택하였다면, 실험 중재와 비교 중재 사이의 불균형을 중재 결과에 대한 예측적인 기저 특성에 따라 통제함으로써 해결할 수 있다. - 연구대상자의 선택이 중재와 중재 결과 모두와 관련이 있다면 연구결과는 선택 비돌림이 있을 것이다. - 연구대상자가 중재의 시작부터 추적되지 않는다면, 중재 직후 시행된 중재 결과는 결측될 것이다. 이 문제는 중재의 신규 사용자보다 기존 사용자가 분석에 포함될 때 발생할 것이다. 결측자로 분석 방법론을 사용하여 선택 비돌림을 교정하는 것이 원칙적으로 가능하다. 그러나, 거의 사용되지 않는다.
-----------	---

3. 중재 분류의 비돌림		
신호질문	판단근거	응답선택
3.1 중재 군이 명확하게 정의되어 있는가?	Patient eligibility criteria, study design and data collection modalities have been previously described (10). Briefly, OBSERVANT was a national observational, prospective, multicenter cohort study that enrolled consecutive AS patients who underwent TAVR or SAVR at 93 Italian centers (34 hemodynamic centers and 59 cardiac surgery centers) between December 2010 and June 2012, and was run by the Italian National Health Institution in cooperation with the Italian Ministry of Health, the National Agency for Regional Health Services, Italian Regions, and Italian scientific societies and federations representing Italian professionals involved in the management of AS. Hospitals invited to participate were those where a procedural (SAVR and/or TAVR) treatment could be offered to AS patients for the complete list of executive working group members, participating centers, and investigators). For the purposes of the present analysis, patients who underwent an associated procedure or a trans-aortic and/or transapical TAVR and patients who reported having a porcelain aorta, hostile thorax, and those who underwent combined coronary artery bypass grafting (CABG) or percutaneous coronary intervention were excluded.	Y
3.2 중재 시작 시 중재군들을 정의 하는데 사용된 정보가 있는가?		Y
3.3 중재 상태의 분류가 중재 결과의 지식 또는 중재 결과의 위험에 의해 영향을 받을 수 있는가?	중재 분류의 명확성(중재법의 차이)으로 인해 중재 결과에 따라 중재 분류가 영향을 받지 않는 것으로 판단됨	N
비돌림위험 판단	중재 상태가 잘 정의됨. 중재의 정의는 중재 시 수집된 정보에만 근거함	낮음
선택 사항: 중재 결과의 측정이나 중재에 의해 예상되는 비돌림의 방향은 무엇인가?		예측할 수 없음

설명	<ul style="list-style-type: none"> - 개인 수준의 중재의 경우, 각 중재를 받는 개인을 고려하는 기준(유형, 세팅, 용량, 빈도, 강도, 중재 시기 등)이 명확하고 명시적이어야 한다. - 중재에 관한 정보가 후속 결과에 의하여 영향을 받을 수 없는 자료로부터 이용가능한 것이라면, 중재 상태의 차등적 오분류는 거의 일어나지 않을 것이다. - 중재 시점에서 정보 수집은 비돌림을 피하기 충분하지 않을 수 있다.
-----------	---

4. 의도한 중재에서 이탈되어 나타나는 비둘림		
신호질문	판단근거	응답선택
4.1 통상진료에서 예상되는 것 이상으로 의도한 중재로부터 이탈되는 상황이 있었는가?	통상진료에서 중재로부터의 이탈이 거의 발생할 수 있는 상황은 아닌 것으로 고려됨	PN
4.2 4.1이 Y/PY 인 경우: 의도한 중재로부터 이탈된 상황이 중재 기간에 불균형을 초래하고, 중재 결과에 영향을 미쳤을 가능성이 있는가?	통상진료를 반영하지 않은 의도한 중재로부터 나타나는 이탈은 중재 결과에 영향을 미친다면 중요할 것이나 그렇지 않다면 중요하지 않을 것이다. 또한 두 군에 걸친 이탈의 불균형이 있을 경우에만 비둘림이 발생할 것이다.	NA
본 연구의 목적이 중재에 대한 시작 및 준수의 효과를 평가하는 것이라면 4.3~4.6에 응답하십시오.		
4.3 중요한 공동 중재는 중재기간 균형을 이루는가?		
4.4 대부분의 연구대상자에게 중재가 성공적으로 시행되었는가?		
4.5 연구대상자는 배정된 중재법을 준수하였는가?		
4.6 4.3, 4.4 또는 4.5이 N/PN 인 경우: 중재의 시작 및 준수의 효과를 추정하기 위해 적절한 분석방법을 사용하였는가?		
비둘림위험 판단	중재로부터의 이탈이 거의 발생할 수 있는 상황은 아님	낮음
선택 사항: 의도한 중재로부터 이탈로 인해 예상되는 비둘림의 방향은 무엇인가?		예측할 수 없음

설명	<p>- 중재에 대한 배정 효과를 평가시</p> <ul style="list-style-type: none"> • 중재 후 통상진료에서 발생한 이탈은 의도한 중재의 일부이므로 중재 배정 효과에 비둘림을 유발하지 않는다. 이탈은 중재군과 비교 중재 간의 차이에 대한 기대로 인해 발생할 수 있다. 그러나, 통상진료에서 개인에 대한 관찰연구에서는 이러한 상황이 예상되지 않는다. • 통상진료를 반영하지 않은 의도한 중재로부터 나타나는 이탈은 중재 결과에 영향을 미친다면 중요할 것이나 그렇지 않다면 중요하지 않을 것이다. 또한 두 군에 걸친 이탈의 불균형이 있을 경우에만 비둘림이 발생할 것이다.
-----------	--

설명	<p>- 중재에 대한 시작 및 준수의 효과를 평가시</p> <ul style="list-style-type: none"> • 계획되지 않은 공동 중재가 중재의 추정 효과를 편향시키는 방식으로 시행될 경우 비둘림의 위험이 더 높아질 것이다. 공동 중재가 중재 결과에 영향을 미칠 때는 중요하나, 그렇지 않은 경우에는 중요하지 않다. 비둘림은 중재군 사이에 공동 중재의 불균형이 있을 경우에만 발생할 것이다. • 연구대상자들이 의도한대로 중재를 준수하지 않는다면 비둘림위험이 높을 것이다. 미준수에는 불완전한 준수, 중재의 중단, 중재군과 비교 중재의 교차, 다른 적극적인 중재로의 전환 등이 있다. • 의도한 중재로부터 몇가지 유형의 이탈을 수정하는 분석을 수행할 수 있다. 적절한 분석 전략의 예로 역확률 가중치나 도구변수 추정이 있다. 논문에서 의도한 중재로부터의 이탈에 대한 정보를 보고하지 않고 그러한 분석을 보고하는 것이 가능할 수는 있지만, 정보가 없는 경우 그러한 분석이 적절하다고 판단하기는 어렵다.
-----------	--

5. 결측치로 인한 비둘림		
신호질문	판단근거	응답선택
5.1 모든 혹은 거의 모든 연구대상자의 중재 결과 자료를 사용할 수 있었는가?	모든 혹은 거의 모든 연구대상자의 중재결과 자료를 보고함	<u>Y</u>
5.2 중재 상태에 대한 결측 자료로 인해 연구대상자가 제외되었는가?	결측 자료에 따라 대상자수 제외는 없었음	<u>N</u>
5.3 분석에 필요한 다른 변수에 대한 결측자료로 인해 연구대상자가 제외되었는가?	사전에 성향점수에 따른 매칭 population으로 모든 대상자에 대한 분석결과가 제시됨	<u>N</u>
5.4 5.1이 PN/N 혹은 5.2나 5.3이 Y/PY 인 경우: 결측된 연구대상자의 비율과 원인이 중재 군간 유사한가?		NA
5.5 5.1이 PN/N 혹은 5.2나 5.3이 Y/PY 인 경우: 결측치에 관계없이 연구결과가 견고(robust)하다는 근거가 있는가?		NA
비둘림위험 판단	자료가 합리적으로 완료됨	낮음
선택 사항: 결측치에 의해 예상되는 비둘림의 방향은 무엇인가?		예측할 수 없음

설명	<ul style="list-style-type: none"> - “거의 모든”은 “연구결과를 신뢰할 만큼 충분하다”로 해석되어야 하며, 적절한 비율은 맥락에 따라 달라질 수 있다. 어떤 상황에서는 두 중재 군 간 상당히 공통적으로 관심 사건이 발생한다면, 연구대상자의 95%(혹은 가능한 90%) 자료를 이용할 수 있는 것으로 충분할 수 있다. - 중재 상태가 누락되면 문제가 될 수 있다. 이를 위해서 의도한 연구 표본이 명확해야 한다. - 결측된 관찰값의 비율의 차이와 원인의 차이에서 “비슷하다”는 우연에 의해 기대되는 것만큼 중재군 간 약간의 불일치한 것을 의미한다. 견고성에 대한 근거는 분석에 사용된 가정이 명확하고 타당한지를 평가하는 것이 중요하다. 이를 위해서는 연구 내용에 대한 지식과 통계 전문지식이 필요하다. 예를 들어, 다중 대체법(multiple imputation)과 같은 통계 방법의 사용이 적절한 답변을 보장하는 것은 아니다.
-----------	--

6. 중재결과 측정 비틀림		
신호질문	판단근거	응답선택
6.1 중재결과 측정이 제공받은 중재 의 지식에 의해 영향을 받았을 수 있는가?	일차결과변수인 모든 원인 사망과 MACCE(사망, 뇌졸중, 심근경색, PCI & CABG)를 고려할 때 중재 지식에 영향을 거의 받지 않는다고 판단됨(의료기록에 기반)	PN
6.2 중재결과 평가자는 연구대상자들이 받은 중재를 알고 있는가?	record linkage with the National Hospital Discharged Records(HDR) database(for in-hospital events) and with the Tax Registry Information System(TRIS) DATA QUALITY ASSESSMENT. Specific quality assessment activities have been arranged to evaluate the reliability of the OBSERVANT database. In particular, independent observers, following sSpecific standard operating procedures, monitored the participating hospitals to assess the completeness of the enrolled cohort and to compare the collected data to those reported in the original clinical charts.	PN
6.3 중재결과 평가 방법은 중재 군간에 유사했는가?	대부분의 결과변수 6.2 판단근거 내용에 따라 의료기록에 기반하여 평가하였음 Echocardiographic criteria post-procedure(prosthesis performance and paravalvular regurgitation) were defined according to the Valve Academic Research Consortium definitions(14).The endpoint definitions are reported in the Online Appendix.	Y
6.4 제공받은 중재와 관련된 중재 결과 측정 시 체계적 오류가 있었는가?	평가자의 눈가림이나 중재결과 평가방법의 유사성 확보 이외에 중재 결과 측정에 따른 차등적 오분류 가능성은 낮은 것으로 고려함(의료기록에 기반한 의료결과(진단, 수술 등)로 나타나는 결과변수가 대부분)	PN
비틀림위험 판단	중재결과 평가 방법이 군간에 유사함 중재결과 측정이 연구대상자가 받은 중재에 대한 지식에 의해 영향을 받지 않았거나(즉 객관적임), 중재결과 평가자가 연구대상자가 받은 중재를 알지 못했음 중재결과 측정에서의 오차가 중재 상태와 관련이 없음	낮음
선택 사항: 중재결과 측정에 의해 예상되는 비틀림의 방향은 무엇인가?		예측할 수 없음

설명	<ul style="list-style-type: none"> - 일부 중재결과에서 측정(모든 원인 사망률 또는 반복 불가능한 자동화된 실험실 지표 등)으로 인한 비돌림위험은 낮을 것으로 예상된다. - 만약 중재결과 평가자가 중재 상태에 눈가림이 되어 있다고 한다면, 이 질문에 대한 대답은 “아니다” 일 것이다. 적극적으로 눈가림을 시행하지 않았다고 하더라도, 중재를 인지하지 못할 수도 있을 것이다. 이런 경우 이 질문에 대한 답은 “아니다”가 될 것이다. - 중재결과와 차등적 오분류에 있어서 중재결과 측정에 체계적 오류가 있는 경우, 중재 혹은 중재-중재결과 관계의 교란 변수와 관련이 있다면 비돌림이 발생할 수 있다.
-----------	---

7. 보고된 연구결과 선택의 비돌림		
신호질문	판단근거	응답선택
보고된 효과 추정치는 다음과 같은 연구결과에 기초하여 선택되었을 가능성이 있는가?		
7.1 중재 결과 영역 내에서 여러 개의 중재 결과 측정	특정 중재결과 영역에서 서로 다른 측정의 효과 추정치를 보고하지는 않음	<u>N</u>
7.2 중재-결과 관계에서의 다중 분석	비무작위연구이기는 하나 교란통제 및 결측치 등의 제한점 극복을 위해 다른 분석방법을 다양하게 사용하여 효과 추정치를 다수 생성하지는 않음	<u>N</u>
7.3 다른 하위군	여러 하위군에 대한 다수의 효과 추정치를 생성하지는 않음	<u>N</u>
비돌림위험 판단	모든 보고된 연구결과가 모든 의도한 중재 결과, 분석 및 하위 코호트에 해당한다는 명확한 근거가 있음	낮음
선택 사항: 보고된 연구결과와 선택에 의해 예상되는 비돌림의 방향은 무엇인가?		예측할 수 없음

설명	<ul style="list-style-type: none"> - 특정 중재결과 영역의 경우, 서로 다른 측정에 대해 다수의 효과 추정치를 생성할 수 있다. 다수의 측정이 이루어졌지만 하나 또는 일부만 보고된 경우, 연구결과에 따른 선택적 보고의 위험이 있다. - 비무작위 연구로부터 사용된 자료의 한계 때문에 이러한 제한점을 해결하기 위해 다른 분석 방법을 수행할 수 있다. 복수의 방법을 적용하면 특정 결과 측정에 대한 복수의 효과 추정치가 생성된다. 만약 통계분석자들이 적용할 방법을 사전에 정의하지 않고, 다수의 추정치가 생성됐음에도 불구하고 하나 혹은 하위집단 중재 결과만 보고한 경우, 결과에 기초한 선택적 보고의 위험이 있다. - 큰 코호트의 경우 여러 하위군에 대해 다수의 효과 추정치를 생성하거나 원래 코호트의 다양한 비율을 단순히 생략할 수 있다. 다수의 추정치가 생성되지만 하나 또는 일부만 보고되는 경우 연구결과에 따른 선택적 보고의 위험이 있다.
-----------	---

전체 비뚫림위험 판단	모든 영역에서 낮거나 중간 비뚫림위험으로 판단됨	중등도
선택 사항: 중재 결과에 대해 전반적으로 예상되는 비뚫림의 방향은 무엇인가?	중요 교란요인이 충분히 고려되었다고 판단하기 어려우므로 실험중재에 보다 선호될 것으로 판단됨	실험 중재 선호

