

SVEUČILIŠTE U SPLITU
MEDICINSKI FAKULTET

Ružica Bojčić

Učestalost i način korištenja alata AMSTAR 2 u procjeni sustavnih pregleda

Doktorski rad

Akadska godina: 2024./2025.

Mentorica:

prof. dr. sc. Livia Puljak, dr. med.

U Splitu, kolovoz 2025.

ZAHVALA

Zahvaljujem

SADRŽAJ

POPIS KRATICA	1
1.UVOD	2
1.1. Sustavni pregled i meta-analiza.....	2
1.2. Alat AMSTAR 2 za kritičku procjenu metodologije sustavnih pregleda	3
1.3. Značaj alata AMSTAR 2 za znanstvenu i kliničku praksu	4
1.4. Razlozi za razvoj alata AMSTAR 2	5
1.5. Sadržaj i objašnjenja domena alata AMSTAR 2.....	6
1.6. Način korištenja alata AMSTAR 2	14
1.7. Usporedba alata AMSTAR 2 i AMSTAR.....	17
1.8. Korištenje AMSTAR-a 2 za procjenu SR-ova koji ne analiziraju istraživanja intervencija	18
2. CILJEVI RADA I HIPOTEZE	19
3. JEDINICE ANALIZE I POSTUPCI.....	20
3.1. Jedinice analize	20
3.2. Etička načela	20
3.3. Ustroj istraživanja	20
3.4. Kriteriji uključenja i isključenja.....	20
3.5. Pretraživanje literature	21
3.6. Probir literature	21
3.7. Prikupljanje podataka.....	22
3.8. Anketiranje ispitanika	22
3.9. Ishodi istraživanja	23
3.9.1. Ishodi prvog istraživanja	23
3.9.2. Ishodi drugog istraživanja	23
3.9.3. Ishodi trećeg istraživanja.....	24
3.10. Statistička analiza	25

4. REZULTATI.....	26
4.1. Rezultati prvog istraživanja.....	26
4.1.1. Bibliometrijska analiza.....	26
4.1.2. Izvori informacija koji su referencirani za korištenje istraživanih alata	29
4.1.3. Rezultati ankete	29
4.2. Rezultati drugog istraživanja.....	33
4.2.1. Rezultati pretraživanja.....	33
4.2.2. Značajke uključenih sustavnih pregleda	34
4.2.3. Metodološka procjena kvalitete uključenih sustavnih pregleda.....	35
4.3. Rezultati trećeg istraživanja	38
4.3.1. Rezultati pretraživanja.....	38
4.3.2. Značajke analiziranih istraživanja	38
4.3.3. Pristup izvještavanju i način korištenja alata AMSTAR 2.....	40
5. RASPRAVA.....	44
5.1. Prvo istraživanje	44
5.2. Drugo istraživanje	47
5.3. Treće istraživanje	50
6. ZAKLJUČCI	55
6.1. Zaključak prvog istraživanja	55
6.2. Zaključak drugog istraživanja	55
6.3. Zaključak trećeg istraživanja.....	55
7. LITERATURA.....	57
8. SAŽETAK.....	66
9. SUMMARY	68
10. PRILOZI.....	70
10.1. Prilog 1. Strategija pretraživanja literature	70

10.2. Prilog 2. Programski kôd korišten za nasumični odabir istraživanja za analizu s popisa uključenih istraživanja.....	71
10.3. Prilog 3. Tekst ankete poslane e-mailom u tijeku prvog istraživanja	72
11. ŽIVOTOPIS	73

POPIS KRATICA

AMSTAR - engl. *A MeaSurement Tool to Assess Systematic Reviews*

EBM - medicina utemeljena na dokazima, engl. *Evidence-Based Medicine*

MA - meta-analiza, engl. *Meta-analysis*

MeSH - medicinski predmetni pojmovi, engl. *Medical Subject Headings*

MOOSE - engl. *Meta-analysis of Observational Studies in Epidemiology*

NRSI - nerandomizirana istraživanja intervencija, engl. *Non-Randomized Studies of Interventions*

OSF - engl. *Open Science Framework*

OSR - pregled sustavnih pregleda, engl. *Overview of Systematic Reviews*

PRISMA - engl. *Preferred Reporting Items for Systematic Reviews and Meta-Analyses*

R-AMSTAR - engl. *Revised AMSTAR*

RCT - randomizirani kontrolirani pokus, engl. *Randomized Controlled Trial*

RoB - rizik od pristranosti, engl. *Risk of Bias*

ROBINS-I - engl. *Risk of Bias In Non-randomized Studies - of Interventions*

SR - sustavni pregled, engl. *Systematic Review*

1.UVOD

1.1. Sustavni pregled i meta-analiza

Sustavni pregled (engl. *Systematic Review*, SR) označava vrstu znanstvenog istraživanja kojoj je cilj sažeti sva postojeća istraživanja koja zadovoljavaju unaprijed određene kriterije kako bi se odgovorilo na specifično istraživačko pitanje (1). SR koristi eksplicitne metode koje su odabrane na način da u najvećoj mogućoj mjeri smanje pristranost, čime se osiguravaju pouzdaniji dokazi iz kojih se mogu izvlačiti zaključci i donositi odluke (2,3). Metodologija SR-a, koju je uveo i razvio Cochrane, visoko je strukturirana, transparentna i ponovljiva (4). To uključuje: unaprijed jasno definirano istraživačko pitanje; jasnoću opsega SR-a i prihvatljivost istraživanja za uključivanje; ulaganje svih napora da se pronađu sva relevantna istraživanja i da se u obzir uzme rizik od pristranosti (engl. *Risk of Bias*, RoB) u uključenim istraživanjima; te analiziranje uključenih istraživanja na nepristran i objektivan način (1).

Važan korak u SR-u je pažljivo razmatranje prikladnosti kombiniranja numeričkih rezultata svih, ili možda samo nekih, uključenih istraživanja. U slučaju kada se odluči napraviti statističku kombinaciju rezultata iz dvaju ili više zasebnih istraživanja, to se naziva meta-analiza (engl. *Meta-analysis*, MA). Potencijalne prednosti MA uključuju poboljšanje preciznosti, mogućnost odgovaranja na pitanja koja pojedinačna istraživanja ne postavljaju te priliku za rješavanje kontroverzi koje proizlaze iz suprotstavljenih tvrdnji. Međutim, MA također ima potencijal za ozbiljno zavaravanje, osobito ako se pažljivo ne razmotre specifični ustroji istraživanja, pristranosti i varijacije među istraživanjima, te pristranosti u izvještavanju (1).

Neophodno je istaknuti da SR-ovi predstavljaju najvišu razinu dokaza u medicini (5). Hijerarhija dokaza u medicini prikazana je na Slici 1. SR-ovi također čine okosnicu medicine utemeljene na dokazima (engl. *Evidence-Based Medicine*, EBM). Pojam EBM formalno je uveden i populariziran 1990.-ih kao novi pristup medicinskoj edukaciji i kliničkoj praksi, s naglaskom na integraciju najboljih dostupnih znanstvenih dokaza s kliničkom stručnošću i preferencijama pacijenta (5,6).

S obzirom na sve brže povećanje broja objavljenih SR-ova i činjenicu da na isto istraživačko pitanje različiti SR-ovi daju različite odgovore, pojavila se potreba za istraživanjima koja objedinjuju već postojeće SR-ove. Takva vrsta istraživanja naziva se pregled sustavnih pregleda (engl. *Overview of Systematic Reviews* ili *umbrella review*, OSR) (1). U skladu s time, pojavila se još veća potreba za procjenom metodološke kvalitete SR-ova (7,8).

Slika 1. Hijerarhija dokaza u medicini



1.2. Alat AMSTAR 2 za kritičku procjenu metodologije sustavnih pregleda

AMSTAR (engl. *A Measurement Tool to Assess Systematic Reviews*) je alat za kritičku procjenu metodologije SR-ova. Prva inačica alata objavljena je 2007. i sadrži 11 domena procjene metodološke kvalitete SR-ova randomiziranih kontroliranih pokusa (engl. *Randomized Controlled Trial*, RCT) (7). Nova inačica alata nazvana je AMSTAR 2 i objavljena je u rujnu 2017. (8).

Alat AMSTAR 2 omogućuje detaljniju procjenu SR-ova u odnosu na AMSTAR. Također omogućuje procjenu SR-ova koji uključuju randomizirana ili nerandomizirana istraživanja intervencija u području biomedicine i zdravstva, ili oboje. AMSTAR 2 sadrži 16 domena procjene. Za svaku domenu, alat postavlja pitanje na koje odgovor u jedanaest domena može biti „da“ ili „ne“, a u pet domena postoji i dodatni odgovor „djelomično da“. Prema autorima alata, prilikom procjene SR-ova nije predviđena dodjela bodova za pojedine domene AMSTAR 2 alata kao ni određivanje ukupne brojčane ocjene. Umjesto toga, autori su definirali sedam ključnih domena na temelju čega se SR-ovi mogu svrstati u 4 kategorije – SR visoke, umjerene, niske ili kritično niske pouzdanosti (engl. *high, moderate, low or critically low confidence*) (8).

Obje inačice alata objavili su Shea i sur. (7,8). U međuvremenu, preciznije 2010., Kung i sur. objavili su R-AMSTAR (engl. *Revised AMSTAR*) kojem je cilj bio „kvantificirati kvalitetu“ SR-ova. Oni su, sadržajno se držeći pitanja iz originalnog alata, osmislili način dodjeljivanja bodova za svaku domenu i generiranja ukupnog brojčanog rezultata (9). Autori istraživanja koje je na istom uzorku SR-ova za procjenu kvalitete koristilo AMSTAR i R-AMSTAR, zaključili su da su ocjene kvalitete dobivene tim alatima usporedive (10). Međutim, istraživanje Piepera i sur. zaključilo je kako R-AMSTAR, za razliku od AMSTAR-a, nije dobar mjerni instrument (11).

1.3. Značaj alata AMSTAR 2 za znanstvenu i kliničku praksu

Jedan je članak iz 2010. pokazao da se još u to vrijeme dnevno objavljivalo 75 RCT članaka i 11 SR-ova koji su uključili RCT-ove (12). Drugo istraživanje, objavljeno 2020., pokazalo je eksponencijalno povećanje broja članaka koji opisuju RCT-ove objavljenih u časopisima s čimbenikom utjecaja većim od deset u razdoblju od 1965. do 2017. god.. Pri tome je, od ukupnog broja članaka objavljenih u navedenom razdoblju, čak 60 % bilo objavljeno nakon 2000. (13).

Obzirom na vrlo brzo povećanje broja objavljenih istraživanja iz područja biomedicine i zdravstva, može se reći da je praćenje rezultata primarnih istraživanja postalo nemoguća misija za zdravstvene djelatnike (7,8). Stoga bi bilo pragmatično oslanjati se na SR-ove prilikom donošenja odluka o zdravstvenoj skrbi (14). Upravo SR-ovi pružaju mogućnost da se donošenje odluka temelji na točnim, sažetim i vjerodostojnim sažecima najboljih dostupnih dokaza o nekoj temi (14).

SR-ovi, kao najviša razina dokaza u medicini, čine temelj pisanja smjernica koje se onda koriste za donošenje važnih kliničkih odluka u svakodnevnoj praksi (8). Međutim, više je istraživanja pokazalo da SR-ovi nisu nužno provedeni u skladu s metodološkim očekivanjima postavljenima u AMSTAR-u ili AMSTAR-u 2. Na primjer, autori istraživanja koje je pomoću AMSTAR-a procijenilo kvalitetu 97 SR-ova intervencijskih RCT-ova na temu neuropatske boli izvijestili su da njihova metodološka kvaliteta nije bila optimalna (10). Novije istraživanje koje je procjenjivalo metodološku kvalitetu SR-ova o stomatološkim *in vitro* istraživanjima, a koje je uključilo 185 SR-ova, pokazalo je da je 68 % bilo kritično niske pouzdanosti prema AMSTAR 2 alatu (15). Nadalje, istraživanje o kvaliteti SR-ova o vremenu uvođenja dohrane za prevenciju alergija u ranom djetinjstvu, koje je uključilo 12 SR-ova s RCT-ovima, otkrilo je da je među njima čak devet bilo kritično niske do niske pouzdanosti prema AMSTAR-u 2 (16). Istraživanje o metodološkoj kvaliteti SR-ova o liječenju depresije kod odraslih zaključilo je kako je nužno poboljšanje njihove

metodologije prema AMSTAR-u 2, budući da su od 60 uključenih SR-ova čak 53 bila kritično niske pouzdanosti (17).

Iz svega navedenog, jasno je da nekritično prihvatanje rezultata samo jednog SR-a može biti rizično i da je važno razvijati svjesnost da nisu svi SR-ovi jednako pouzdani. Korištenje loših i pristranih metoda može dovesti do nepouzdanih SR-ova. Stoga je važno da postoji alat kao što je AMSTAR 2, kako bi korisnici SR-ova mogli razlikovati one visoke pouzdanosti od onih s neprikladnom metodologijom (8).

Još jedna potencijalno važna primjena alata AMSTAR 2 je mogućnost njegova korištenja za pomoć tijekom učenja o SR-ovima. Drugim riječima, AMSTAR 2 može poslužiti kao sažete smjernice za one koji provode SR-ove da to naprave bolje (8). Ipak, obzirom da mu navedeno nije primarna svrha, tj. AMSTAR 2 ne objašnjava u detalje metodologiju provođenja SR-ova, njegovi autori one koji traže sveobuhvatne materijale za učenje upućuju na Cochraneov priručnik o izradi sustavnih pregleda (engl. *Cochrane Handbook*) (1,8).

1.4. Razlozi za razvoj alata AMSTAR 2

U znanosti je izuzetno važno standardizirati izvještavanje o pojedinim vrstama istraživanja kako bi ti izvještaji bili usporedivi, transparentni i dovoljno detaljni. Propisno pisanje izvještaja o istraživanju omogućuje prikladnu procjenu i ponavljanje istraživanja. Za standardizaciju izvještavanja o SR-ovima koriste se smjernice PRISMA (engl. *Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) i MOOSE (engl. *Meta-analysis of Observational Studies in Epidemiology*) (18,19). Kvaliteta izvještavanja SR-a može, međutim, točnije odražavati sposobnost autora da pišu na razumljiv i detaljan način, a ne sam način provođenja sustavnog pregleda. Zbog toga se ukazala potreba za stvaranjem alata AMSTAR i AMSTAR 2, odnosno alata koji ocjenjuje način na koji se SR-ovi planiraju i provode (1,8,20).

Cochraneov priručnik o izradi SR-ova pruža sveobuhvatan vodič za njihove autore, ali ne sadrži sažet i praktičan instrument kritičke procjene već objavljenih SR-ova (1). Premda postoje različiti instrumenti za procjenu pojedinačnih istraživanja uključenih u SR, kao i instrumenti za procjenu nekih koraka u izradi SR-a (npr. meta-analiza, testiranje pristranosti objavljivanja), zapravo je relativno malo instrumenata koji obuhvaćaju procjenu svih važnih koraka u tijeku planiranja i provođenja sustavnog pregleda (21–27). Upravo takav instrument bio je originalni AMSTAR iz 2007. koji je postao jedan od najčešće korištenih alata za procjenu metodologije SR-ova (7,28,29).

AMSTAR je praktičan alat koji je omogućio zdravstvenim djelatnicima, istraživačima, autorima smjernica u zdravstvu i donositeljima zakona iz područja zdravstvene skrbi, koji nužno ne moraju imati naprednu razinu poznavanja epidemiologije i metodologije, relativno brzu, lako ponovljivu i kritičnu procjenu SR-ova intervencijskih RCT-ova (7). Od objave AMSTAR-a objavljeno je više konstruktivnih kritika tog alata. Također su autori AMSTAR-a dobili bitne povratne informacije od korisnika alata na različitim radionicama. Nadalje, uzimajući u obzir nova dostignuća u znanosti o SR-ovima, pojavila se potreba za revidiranjem originalnog alata, odnosno razvojem AMSTAR-a 2 (8,9,30–32).

1.5. Sadržaj i objašnjenja domena alata AMSTAR 2

Mnoge od ukupno 16 domena u AMSTAR-u 2 napisane su na način da su razumljive same po sebi. No, budući da alat pokriva procjene u širokom spektru intervencija, neke su domene podložne različitim tumačenjima pa ih je nužno detaljnije pojasniti. Detaljna objašnjenja svih domena objavljena su u uputama za korisnike kao dodatak izvornom članku koji opisuje AMSTAR 2 (8).

Domena 1.: Jesu li istraživačka pitanja i kriteriji uključenja za SR sadržavali PICO elemente?

Kratica PICO dolazi od prvih slova engleskih riječi: *Population* (populacija), *Intervention* (intervencija), *Comparator/Control group* (komparator/kontrolna skupina) i *Outcome* (ishod). Uobičajena je praksa koristiti PICO opis kao organizacijski okvir za istraživačko pitanje. Ponekad je potrebno dodati i dodatnu domenu za vremenski okvir (*engl. Time*) ako je to kritično u određivanju vjerojatnosti da istraživanje obuhvati relevantne kliničke ishode (npr. kada se učinak intervencije očekuje tek nakon nekoliko godina) pa se u tom slučaju koristi kratica PICOT. Okvir PICO definira elemente koje treba detaljno opisati u SR-u kako bi se jasno postavilo istraživačko pitanje. AMSTAR 2 ne zahtijeva da elementi PICO-a budu eksplicitno napisani, ali traži da se mogu jasno prepoznati pažljivim čitanjem sažetka, uvoda i metoda. Da bi ocijenili s „da“ ovu domenu, procjenitelji bi trebali biti sigurni da su sva četiri elementa PICO-a opisana (8).

Domena 2.: Je li izvještaj o SR-u sadržavao izričitu izjavu da su metode bile definirane prije provedbe SR-a i, ukoliko je bilo značajnih odstupanja od protokola, jesu li za njih navedena opravdanja?

Izrada dobro definiranog protokola prije početka istraživanja te pridržavanje istog smanjuju RoB u SR-u. Autori trebaju u izvještaju napisati da su SR proveli prema unaprijed definiranom protokolu

koji bi morao biti javno dostupan. Autori javnu provjeru protokola mogu omogućiti njegovom registracijom na internetu – npr. u registru kao što je PROSPERO (33) ili na mrežnoj stranici OSF-a (engl. *Open Science Framework*) (34), kao i objavom u časopisu.

Dakle, prije početka provedbe SR-a, istraživačka pitanja, strategija pretraživanja literature, kriteriji uključenja i isključenja, te planirani način procjene RoB-a trebaju biti detaljno opisani. Za ocjenu „djelomično da“ u izvještaju mora barem pisati da je potonje učinjeno. Za ocjenu „da“, uključujući sve prethodno navedeno, autori moraju imati registriran protokol i uz to trebaju zadovoljiti sljedeće: plan za sintezu dokaza, odnosno MA ukoliko istu bude prikladno napraviti, plan za proučavanje uzroka heterogenosti, te objašnjenje razloga bilo kakvog odstupanja od protokola. Procjenitelji trebaju usporediti objavljeni izvještaj s registriranim protokolom. Ako utvrde odstupanja od protokola koja nisu objašnjena, uputno je smanjiti ocjenu za ovu domenu (8).

Domena 3: Jesu li autori objasnili odabir vrsta istraživanja uključenih u SR?

Postoje SR-ovi koji uključuju samo RCT-ove, ili samo nerandomizirana istraživanja intervencija (engl. *Non-Randomized Studies of Interventions*, NRSI), ili oboje (1). Odabir vrsta istraživanja koja će se uključiti u SR ne smije biti proizvoljan, nego ga autori trebaju obrazložiti. Autori bi se najprije trebali pitati bi li SR koji uključuje samo RCT-ove dao nepotpun odgovor na njihovo istraživačko pitanje. To je slučaj kada nema relevantnih RCT-ova, kada su ishodi u dostupnim RCT-ovima nedostadni, neprikladne statističke snage ili ograničene populacije ispitanika. U takvim je slučajevima opravdano uz RCT-ove uključiti i NRSI. Drugim riječima, uključivanje i RCT-ova i NRSI može biti opravdano kako bi se dobila potpunija slika učinkovitosti i štetnosti povezanih s ispitivanom intervencijom, pri čemu autori AMSTAR-a 2 preporučuju da se te dvije vrste istraživanja procjenjuju zasebno.

Naprotiv, da bi se opravdalo ograničavanje SR-a samo na RCT-ove, autori bi trebali navesti da je za njihov SR postojao dostatan broj uključivih RCT-ova za koje smatraju da mogu dati cjelovit odgovor na istraživačko pitanje i bez da uključe NRSI. Obrazloženje može biti i nastojanje autora da u što većoj mjeri smanje RoB.

Ograničavanje SR-a samo na NRSI opravdano je u slučajevima kada RCT-ovi ne mogu pružiti potrebne podatke o ishodu ili kada već postoje dostupni SR-ovi RCT-ova, a SR koji uključuje samo NRSI će nadopuniti ono što je već poznato.

Važno je napomenuti da je ovo donekle zanemareno područje, stoga ovu domenu može biti teško procijeniti. Često je nužno pažljivo pročitati cijeli SR prije donošenja procjene. Za ocjenu „da“

autori bilo gdje u izvještaju o SR-u moraju navesti obrazloženje odabira vrste uključenih istraživanja, bez obzira na to koju su od gore navedenih mogućnosti izabrali (8).

Domena 4.: Jesu li autori SR-a koristili sveobuhvatnu strategiju pretraživanja literature?

Dobro je utvrđeno koliko je značenje prikladnog pretraživanja literature u tijeku izrade SR-ova (1). Potrebno je pretražiti najmanje dvije bibliografske baze podataka (npr. CENTRAL, EMBASE i MEDLINE) te navesti datum pretraživanja. Obavezno je napisati ključne riječi i/ili medicinske predmetne pojmove (engl. *Medical Subject Headings*, MeSH). Ukoliko cjelovita strategija pretraživanja nije navedena u izvještaju o SR-u (u glavnom dijelu ili u prilogima), treba biti navedeno da je dostupna na zahtjev. Pretraživanje treba biti nadopunjeno provjerom objavljenih SR-ova, specijaliziranih registara, kontaktiranjem stručnjaka za pojedina područja istraživanja, te uvidom u popis referenci iz uključenih istraživanja. Ponekad je potrebno kontaktirati autore izvornih istraživanja kako bi se razjasnili rezultati ili dobili obnovljeni podaci. Treba tražiti publikacije na svim relevantnim jezicima i dati obrazloženje kada postoje jezična ograničenja. U nekim slučajevima potrebno je pretražiti i takozvanu sivu literaturu (engl. *grey literature*), odnosno literaturu koja nije objavljena u znanstvenim časopisima. Da bi 4. domenu ocijenili s „da“, procjenitelji bi se trebali uvjeriti da su svi upravo opisani relevantni aspekti pretraživanja učinjeni. Za ocjenu „djelomično da“ dovoljno je da su pretražene barem dvije baze podataka, da su navedene ključne riječi ili strategija pretraživanja, te da su obrazložena jezična ili druga ograničenja u pretraživanju (ako ih je bilo) (8).

Domena 5.: Jesu li autori SR-a radili dvostruki neovisni probir pronađenih zapisa istraživanja?

Najbolja praksa zahtijeva da tijekom faze probira (engl. *screening*) dva autora neovisno procijene prihvatljivost svih zapisa pronađenih pretraživanjem literature (1). Poželjno je opisati koliko su se dva ocjenjivača slagala oko uključivanja pojedinih istraživanja te da, ukoliko je bilo neslaganja, opišu način njihovog razrješavanja (npr. rasprava i konzultacija s trećim autorom). Druga je mogućnost da na manjem uzorku dva autora izvrše neovisno probir i da se navede rezultat slaganja među procjeniteljima (engl. *inter-rater agreement*), primjerice u obliku kappa statistike, gdje bi rezultat trebao biti 0,8 ili veći, što upućuje na veliko slaganje među ocjenjivačima. Nakon toga se provedba SR-a nastavlja tako da samo jedan autor procjenjuje je li pojedino istraživanje zadovoljava kriterije uključenja. Da bi 5. domenu ocijenili s „da“ potrebno je da probir bude učinjen na jedan od dvaju opisanih načina (8,35).

Domena 6.: Jesu li autori SR-a izvršili dvostruku neovisnu ekstrakciju podataka?

Sukladno prethodnoj domeni koja se odnosi na probir literature, optimalno je da najmanje dva autora neovisno rade ekstrakciju podataka. U slučaju kada ekstrakciju podataka radi samo jedan autor, potrebno je da istu prethodno na manjem uzorku istraživanja naprave neovisno dva autora i da se navede kappa rezultat usporedbe slaganja među procjeniteljima, koji bi trebao biti 0,8 ili veći. Za ocjenu „da“ potrebno je da ekstrakcija podataka bude učinjena na jedan od ova dva načina. (1,8,35). Treba istaknuti da je pitanje neovisne dvostruke ekstrakcije podataka od velike važnosti, a osobito u slučajevima kada SR uključuje NRSI. Stručna skupina koja je sudjelovala u izradi AMSTAR-a 2 upozorila je da ekstrakcija podataka iz NRSI može biti iznimno složena jer u pravilu uključuje ekstrakciju mjera ishoda koje su već prilagođene s obzirom na potencijalne zbunjujuće čimbenike (engl. *confounding factors*). Za razliku od toga, iz RCT-ova se vade neobrađeni podaci o ishodima (8).

Domena 7.: Jesu li autori SR-a priložili popis isključenih istraživanja i naveli razloge isključenja?

Ova domena zahtijeva od autora objavu cjelovitog popisa istraživanja koja su isključena nakon razmatranja cjelovitog teksta uz navođenje razloga za isključenje svakog pojedinog istraživanja. Postoji niz razloga za isključenje, na primjer zbog neprikladne ili nerelevantne populacije, intervencije ili kontrole. Neopravdano isključenje može utjecati na rezultate SR-a. Stoga je bitno da autori transparentno navedu koja su istraživanja isključili u fazi probira cjelovitih tekstova i zašto. Da bi se 7. domena ocijenila s „djelomično da“ potrebno je da autori objave (u glavnom dijelu teksta ili kao prilog) popis svih istraživanja isključenih u tijeku probira cjelovitih tekstova. Za ocjenu „da“ još je potrebno da uz svako istraživanje s popisa navedu razlog isključenja (8).

Domena 8.: Jesu li autori SR-a dovoljno detaljno opisali uključena istraživanja?

U SR-u autori trebaju opisati svako uključeno istraživanje u skladu s PICO strukturom kliničkog pitanja. Taj bi opis trebao omogućiti da na temelju njega čitatelji SR-a vide jesu li uključena odgovarajuća istraživanja u odnosu na postavljeno istraživačko pitanje, odnosno koliko je određeni SR relevantan za njihovu kliničku praksu. Opis uključenih istraživanja ujedno omogućuje proučavanje heterogenosti u kasnijem tijeku AMSTAR 2 procjene. Za ocjenu „djelomično da“ autori za svako uključeno istraživanje moraju opisati sve od navedenog: populaciju, intervencije, kontrole, ishode i ustroj. Uz to, za ocjenu „da“ moraju detaljno opisati populaciju, intervencije i kontrole, te opisati mjesto/okruženje provođenja istraživanja, kao i vremenski okvir praćenja (8).

Domena 9.: Jesu li autori koristili odgovarajuću metodu procjene RoB-a u pojedinim istraživanjima uključenima u SR?

Procjena RoB-a ključni je dio izrade svakog SR-a, a osobito onih koji uključuju NRSI. Naime, dokaze nije dovoljno samo sažeti iz literature, već ih treba kritički procijeniti. Deveta domena alata AMSTAR 2 propituje jesu li autori SR-a koristili odgovarajuću metodu procjene RoB-a (8). Kada je SR ograničen na RCT-ove, preporučuje se koristiti Cochraneov RoB alat (36). Kada SR uključuje NRSI, autori trebaju koristiti sustavni pristup procjene RoB-a, s dobro razvijenim instrumentom ocjenjivanja. Najnoviji i najopsežniji takav instrument je Cochraneov alat ROBINS-I (engl. *Risk of Bias In Non-randomized Studies - of Interventions*) čije korištenje preporučuju autori AMSTAR-a 2 (8,37). Potrebno je naglasiti da u ovom aspektu AMSTAR 2 nije sveobuhvatan, nego se usmjerava samo na najčešće prepoznate vrste RoB-a. Ukoliko je nužna detaljnija procjena RoB-a, preporučuje se uključivanje stručnjaka iz područja znanstvene metodologije (8).

Procjenitelji trebaju utvrditi jesu li autori SR-a za procjenu RCT-ova koristili Cochraneov RoB alat. Pritom za ocjenu „djelomično da“ trebaju pokazati da su ispitali RoB zbog izostanka prikrivanja razvrstavanja (engl. *unconcealed allocation*) i RoB zbog nedostatka zasljepljenja pacijenata i procjenitelja pri ocjeni ishoda. Uz to, za ocjenu „da“ trebaju pokazati da su ispitali RoB zbog raspodjele koja nije bila uistinu nasumična i RoB zbog odabira objavljenih rezultata između svih rezultata višestrukih mjerenja (8,38).

Ako su autori SR-a uključili NRSI, onda je potrebno procijeniti jesu li autori za procjenu RoB-a koristili ROBINS-I. Pritom za ocjenu „djelomično da“ trebaju pokazati da su ispitali RoB zbog zbunjujućih čimbenika i RoB zbog pristranosti u odabiru (engl. *selection bias*). Za ocjenu „da“ također trebaju pokazati da su ispitali RoB zbog metoda korištenih za utvrđivanje izloženosti i ishoda, te RoB zbog odabira objavljenih rezultata između svih rezultata višestrukih mjerenja (8,38).

Domena 10.: Jesu li autori naveli informaciju o izvorima financiranja istraživanja koja su uključili u SR?

Nekoliko je istraživanja pokazalo da je komercijalno sponzoriranje istraživanja povezano s većom vjerojatnošću da takva istraživanja imaju rezultate koji idu u prilog proizvodu sponzora, nego u slučaju neovisno financiranih istraživanja (39,40). Zbog toga je bitno da autori SR-ova za svako uključeno istraživanje pokušaju pronaći izvore financiranja. Pomoću tih informacija moguće je u SR-u odvojeno analizirati rezultate komercijalno financiranih i neovisno financiranih istraživanja. Za ocjenu „da“ potrebno je da su autori naveli izvore financiranja istraživanja uključenih u SR ili da su naznačili da izvore financiranja nisu uspjeli pronaći (8).

Domena 11.: Ako je meta-analiza bila opravdana, jesu li autori koristili odgovarajuće metode statističkog kombiniranja rezultata?

Autori bi trebali u protokolu SR-a jasno navesti načela na kojima temelje svoju odluku o provedbi MA na podacima iz uključenih istraživanja. MA je statistička metoda koja omogućuje da se zbirno analiziraju rezultati više istraživanja. Međutim, MA bi se trebala raditi samo kad su pojedinačna istraživanja dovoljno slična (homogena) u statističkom i kliničkom smislu istraživanja. Stoga je važno voditi računa u kolikoj se mjeri uključena istraživanja podudaraju za statističko kombiniranje, s obzirom na postavljeni PICO okvir (1,8).

Ako SR uključuje RCT-ove i NRSI, autori trebaju provesti MA odvojeno. Za NRSI je važno znati da populacije mogu uvelike varirati u veličini, od malih kohorti (od desetaka ili stotina sudionika) do istraživanja stotina tisuća pojedinaca i tisuća događaja. Kad bi se ti rezultati kombinirali s onima iz manjih RCT-ova, u objedinjenoj procjeni učinka dominirali bi podaci iz NRSI. Ako se uzme u obzir da na rezultate NRSI može utjecati niz pristranosti, to onda znači da bi ukupne procjene mogle biti precizne, ali pristrane. Drugim riječima, u slučaju NRSI, statističko kombiniranje rezultata može dati vrlo precizne i "statistički značajne", ali zapravo pristrane procjene učinka. Nadalje, interval pouzdanosti (engl. *confidence interval*) računa se na temelju pretpostavke da nema pristranosti (tj. procjene su takve kao da su dobivene iz visoko kvalitetnog RCT-a s istim brojem sudionika). Iznimno je rijetko da NRSI ima tako nizak rizik od pristranosti kao visokokvalitetni RCT istog istraživačkog pitanja. Zbog toga interval pouzdanosti za NRSI (i skupne procjene temeljene na NRSI) treba oprezno tumačiti. Dakle, statistička prilagodba učinaka intervencije za zbunjujuće čimbenike može dati procjene koje se prilično razlikuju od neprilagođene procjene izvedene iz neobrađenih podataka. Općenito, pri statističkom kombiniranju rezultata NRSI pregleda autori bi trebali koristiti prilagođene procjene učinka. Bilo kakve drugačije procjene autori bi trebali jasno opravdati (8).

Za ocjenu „da“, autori trebaju obrazložiti na temelju čega su odlučili provesti MA. Pritom, u slučaju kad MA uključuje RCT-ove, trebaju objasniti izbor korištenog modela MA (MA s fiksnim ili slučajnim učincima), te definirati metode istraživanja heterogenosti. Slično tome, u slučaju kad MA uključuje NRSI, trebaju pokazati da su koristili odgovarajuću tehniku kombiniranja rezultata, prilagođavajući se heterogenosti ako je bila prisutna; te da su statistički kombinirane procjene učinka bile prilagođene s obzirom na zbunjujuće čimbenike. Također je nužno da su autori odvojeno izvijestili sažete procjene za RCT i NRSI kada su obje vrste ustroja bile uključene u MA (8,38). Osim procjene „da“ i „ne“, za ovu domenu postoji mogućnost procjene „meta-analiza nije učinjena“ kako bi se naznačilo da 11. domena nije primjenjiva za SR-ove bez MA (8,38).

Domena 12.: Ako je učinjena meta-analiza, jesu li autori procijenili potencijalni učinak RoB-a u pojedinačnim istraživanjima na rezultate meta-analize, odnosno sinteze dokaza?

U slučajevima kada su autori SR-a odlučili uključiti samo visokokvalitetne RCT-ove, tada sva uključena istraživanja imaju nizak RoB i nema razloga za razmatranje potencijalnog utjecaja pristranosti na rezultate MA. Međutim, u slučajevima kada su autori uključili RCT-ove različitih razina RoB-a, trebali bi procijeniti učinak RoB-a regresijskom analizom ili zasebnim analiziranjem više podskupina istraživanja, ovisno o njihovom procijenjenom RoB-u. Nakon što se napravi zbirna procjena učinka svih istraživanja, u analizi podskupina pojedinačno se procjenjuje učinak pojedinih skupina istraživanja ovisno o tome jesu li imali visok ili nizak RoB pa se ti rezultati usporede s analizom svih istraživanja (8).

U 12. domeni ocjenu „da“ SR-ovi dobiju u dva slučaja. Prvi je slučaj ako su autori uključili samo RCT-ove s niskim RoB-om. Drugi je slučaj ako su autori zbirne procjene u MA temeljili na RCT-ovima i/ili NRSI koji su imali različit RoB, i autori su pritom napravili odgovarajuće analize kojima su procijenili mogući utjecaj RoB-a na zbirnu procjenu učinka intervencije. Osim procjene „da“ i „ne“, za ovu domenu postoji mogućnost procjene „meta-analiza nije učinjena“ kako bi se naznačilo da 12. domena nije primjenjiva za SA bez MA.

Domena 13: Jesu li autori uzeli u obzir RoB u pojedinačnim istraživanjima prilikom interpretacije, odnosno rasprave o rezultatima SR-a?

Čak i u slučajevima kada MA (ili druge metode sinteze podataka kao što je regresijska analiza) nisu provedene, autori bi u raspravi svejedno trebali dati neke komentare vezane za vjerojatni utjecaj RoB-a na rezultate pojedinačnih istraživanja, odnosno na tumačenje rezultata SR-a. Ako su navedeno učinili, ili se radi o SR-u koji je uključio samo RCT-ove niskog RoB-a, ova se domena ocjenjuje s „da“. (8).

Domena 14: Jesu li autori dali zadovoljavajuće objašnjenje i raspravili uzroke heterogenosti u rezultatima SR-a?

Postoje brojni uzroci heterogenosti u rezultatima, osobito kad su u SR uključene NRSI. Neki od primjera su: različit ustroj istraživanja, različite metode analize, različite populacije, različiti intenziteti intervencija, odnosno doza ako se radi o lijekovima. Ako su autori SR-a istražili potencijalne uzroke heterogenosti i raspraviti njihov utjecaj na rezultate, ili u slučaju da značajna heterogenost rezultata nije postojala, ova se domena ocjenjuje s „da“ (8).

Domena 15: Ako je učinjena kvantitativna sinteza, jesu li autori proveli odgovarajuće istraživanje pristranosti u objavljivanju (engl. publication bias) i raspravili njezin mogući utjecaj na rezultate SR-a?

Pristranost u objavljivanju je pristranost koja se javlja u znanstvenim istraživanjima kada rezultati istraživanja nisu objavljeni zbog prirode i smjera rezultata. Konkretno, istraživanja koja pokazuju statistički značajne ili pozitivne rezultate imaju veću vjerojatnost da će biti objavljena, dok istraživanja s negativnim ili statistički neznačajnim rezultatima često ostaju neobjavljena. Taj fenomen može dovesti do iskrivljenog prikaza stvarne situacije u znanstvenoj literaturi zato što pozitivni rezultati postaju prekomjerno zastupljeni. Posljedično se može stvoriti lažna slika o učinkovitosti određenih intervencija, lijekova ili teorija, jer se negativni rezultati ili rezultati koji ne podržavaju određenu hipotezu ne pojavljuju u znanstvenoj literaturi. Na primjer, ako se provede deset istraživanja o učinkovitosti određenog lijeka, a samo dva od njih pokazuju pozitivan učinak, dok ostalih osam pokazuje negativne ili nejasne rezultate, postoji veća vjerojatnost da će dva pozitivna istraživanja biti objavljena. Ako čitatelji imaju pristup samo tim objavljenim istraživanjima, jedina informacija koju imaju jest da lijek djeluje puno bolje nego što zapravo jest. Pristranost u objavljivanju predstavlja ozbiljan problem jer može dovesti do pogrešnih zaključaka u svim područjima istraživanja (41).

Procjena moguće pristranosti objavljivanja u SR-ovima bitno je pitanje, ali autorima SR-ova, kao i procjeniteljima, može biti teško procijeniti razmjere te moguće pristranosti. Obično se u tu svrhu koriste statistički testovi ili grafički prikazi, tzv. grafikon u obliku lijevka (engl. *funnel plot*). Na takvom grafikonu na osi x obično se prikazuje veličina efekta istraživanja, dok se na osi y prikazuje mjera preciznosti istraživanja, najčešće standardna pogreška ili inverzni kvadrat standardne pogreške. Ako nema pristranosti, očekuje se da će istraživanja biti raspoređena simetrično oko prosječne veličine efekta. Istraživanja s većom preciznošću (manja standardna pogreška) nalaze se bliže vrhu lijevka, dok su ona s manjom preciznošću raspoređena šire na dnu. Ako grafikon pokazuje asimetriju, to može ukazivati na pristranost u objavljivanju. Na primjer, ako nedostaje istraživanja na jednoj strani grafikona, to može značiti da istraživanja s negativnim ili statistički neznačajnim rezultatima nisu objavljena. Da bi grafikon u obliku lijevka pokazao asimetriju, odnosno pristranost, potrebno je barem deset istraživanja uključenih u pojedinu meta-analizu (1). U određenoj mjeri važnost pristranosti objavljivanja ovisi o kontekstu i okruženju. Na primjer, postoji veća vjerojatnost da će postojati pristranost objavljivanja ako je uključen niz očigledno metodološki ispravnih istraživanja koje sponzorira industrija lijekova ili medicinskih uređaja (42).

Ako su autori proveli testove za procjenu pristranosti objavljivanja, te raspravili njen mogući utjecaj na rezultate SR-a, ova se domena ocjenjuje s „da“. Osim procjene „da“ i „ne“, za ovu domenu postoji mogućnost procjene „meta-analiza nije učinjena“ kako bi se naznačilo da 15. domena nije primjenjiva za SA bez MA (8).

Domena 16: Jesu li autori naveli vlastite potencijalne izvore sukoba interesa, uključujući sva sredstva koja su primili za provođenje SR-a?

Poznato je kako primarna istraživanja koja financira industrija mogu dati rezultate koji će vjerojatnije pokazati bolje rezultate za sponzoriranu intervenciju nego neovisna istraživanja. Ista se pretpostavka odnosi na SR-ove. Ako neki autori imaju financijski sukob interesa, mogu napraviti pristrani SR koji ide u prilog sponzoru (39,40). Stoga autori trebaju u SR-u opisati svoje izravne izvore financiranja, što časopisi u pravilu i zahtijevaju. Međutim, procjena sukoba interesa tu ne prestaje. Autori bi trebali opisati i druge potencijalne sukobe interesa. Pregled se može financirati neovisno, ali autori mogu biti povezani s tvrtkama koje proizvode lijekove ili proizvode uključene u SR. Nadalje, postoje i profesionalni sukobi interesa koji mogu imati veliki utjecaj, ali ih je teže uočiti jer se rijetko prijavljuju. Primjer toga je kada istraživači dugo rade na nekom polju istraživanja, a SR je u sukobu s njihovim dotadašnjim uvjerenjima. Potencijalni sukob interesa takve vrste posebno je teško procijeniti, ali se na njega može posumnjati ukoliko su istraživači ranije opsežno objavljivali u području na koje se odnosi SR, te su njihova vlastita istraživanja uključena u dotični SR (8,43).

1.6. Način korištenja alata AMSTAR 2

AMSTAR 2 alat objavljen je s detaljnim uputama za korisnike. Na svaku od 16 domena procjenitelj treba odgovoriti s „da“ ili „ne“. U pet domena (drugo, četvrto, sedmo, osmo i deveto) ponuđena je i opcija „djelomično da“. U slučaju kada procjenitelj traženu informaciju o ispunjavanju zahtjeva iz domene ne može pronaći u objavljenom članku o SR-u, niti u dostupnim popratnim materijalima, domena se ocjenjuje s „ne“. Potrebno je napomenuti da se domene pod rednim brojevima 11, 12 i 15 ispunjavaju samo ukoliko je provedena MA (8).

Prema mišljenju autora AMSTAR-a 2, svi su koraci u planiranju i provođenju SR-ova važni. Ipak, izdvojili su sedam domena koje mogu kritično utjecati na vjerodostojnost zaključaka SR-a. Nazvali su ih „ključne domene“. Među njima se čak dvije odnose na RoB – je li prikladno procijenjen i kako utječe na rezultate SR-a. Razlog tolikog naglašavanja važnosti RoB-a leži u tome što je nova verzija alata namijenjena i procjeni SR-ova koji uključuju NRSI (8). Tablica 1. sažeto prikazuje sedam ključnih domena alata AMSTAR 2 (8).

Tablica 1. Sedam ključnih domena alata AMSTAR 2 (8)

2. domena	Registracija protokola prije početka provođenja SR-a
4. domena	Prikladno pretraživanje literature
7. domena	Objašnjenje razloga isključenja potencijalno relevantnih istraživanja
9. domena	Procjena RoB-a u istraživanjima uključenim u SR
11. domena	Prikladnost metoda MA
13. domena	Razmatranje RoB-a prilikom interpretacije rezultata SR-a
15. domena	Procjena prisutnosti i vjerojatnog utjecaja pristranosti objavljivanja

Sedam navedenih ključnih domena samo su prijedlog autora AMSTAR-a 2 koji procjeniteljima također ostavljaju mogućnost da sami definiraju koje će domene biti ključne ovisno o kontekstu istraživanja (8). Potrebno je naglasiti da autori alata AMSTAR 2 ni u kojem slučaju ne preporučuju ocjenjivanje na način da se dodjeljuju bodovi po domenama i generira brojčana ukupna ocjena za SR-ove. Razlog tomu je što ponekad visoka ocjena može prikriti velike nedostatke SR-a , tj. u takvim slučajevima zbirna ocjena ne bi dala pravu sliku o SR-u (8,44,45).

Na temelju takozvanih „slabosti“ odnosno zastupljenosti odgovora „ne“ u ključnim i ne-ključnim domenama, autori su predložili način dodjeljivanja ukupne procjene pouzdanosti SR-a. Tablica 2. prikazuje ukupnu procjenu pouzdanosti SR-a prema alatu AMSTAR 2 (8).

Tablica 2. Ukupna procjena pouzdanosti sustavnog pregleda prema alatu AMSTAR 2 (8)

Ukupna procjena	Broj „ne“ odgovora u ključnim domenama	Broj „ne“ odgovora u ne-ključnim domenama	Značenje procjene
Visoka pouzdanost	0	0 ili 1	SR daje sveobuhvatan i točan sažetak rezultata dostupnih istraživanja.
Umjerena pouzdanost	0	Više od 1*	SR bi mogao dati točan sažetak rezultata uključenih istraživanja.
Niska pouzdanost	1	0, 1 ili više od 1	SR možda neće pružiti točan i sveobuhvatan sažetak dostupnih istraživanja.
Kritično niska pouzdanost	Više od 1	0, 1 ili više od 1	Na SR se ne bi trebalo oslanjati za točan i sveobuhvatan sažetak dostupnih istraživanja.
*Višestruki odgovori „ne“ u ne-ključnim domenama mogu umanjiti pouzdanost u SR pa je možda prikladno ukupnu procjenu smanjiti s umjerene na nisku pouzdanost.			

1.7. Usporedba alata AMSTAR 2 i AMSTAR

Za razliku od izvorne inačice alata AMSTAR koji je služio za procjenu SR-ova koji uključuju samo RCT-ove, AMSTAR 2 uz to omogućuje procjenu SR-ova, bilo da uključuju RCT-ove, NRSI ili oboje (7,8). Uključivanje NRSI u procjenu bilo je neophodno budući da je već u vrijeme nastanka alata AMSTAR 2 oko polovice SR-ova uključivalo i takvu vrstu istraživanja. Budući da NRSI mogu biti različita ustroja te da su podložna velikom broju pristranosti koje u RCT-ovima nisu prisutne ili su prisutne tek u manjoj mjeri, one kao takve zahtijevaju i različit način procjene RoB-a (19,46,47).

Očigledna je razlika da AMSTAR 2 ima 16 domena, u usporedbi s 11 domena u izvornoj inačici. Pritom je deset domena zadržano iz originalnog alata, uz minimalne izmjene u načinu na koji su napisane. Dvije su domene detaljnije obrađene nego u izvornom alatu – autori su svaku od njih podijelili na dvije zasebne domene. Prvo, dvostruki neovisni odabir istraživanja (peta domena) i dvostruka neovisna ekstrakcija podataka (šesta domena) su sada zasebne domene (prije su bile objedinjene). Drugo, potencijalni učinak izvora financiranja sada se razmatra odvojeno za sva istraživanja uključena u SR (10. domena) te za sam SR (16. domena). Također su dodana detaljnija razmatranja RoB-a u vidu odvojenih pod-domena za RCT i NRSI (9. domena) (7,8,36,37). Pretraživanje tzv. sive literature više nije zasebna domena, nego je priključeno domeni pretraživanja literature (4. domena). Ukupno su dodane četiri nove domene. Od toga su dvije preuzete direktno iz ROBINS-I alata: razrada istraživačkog pitanja prema PICO okviru (1. domena) i način na koji je RoB obrađen tijekom sinteze dokaza (12. domena) (8,37). Dodane su još rasprava o mogućim uzrocima i utjecaju heterogenosti (14. domena), te objašnjenje opravdanosti odabira ustroja istraživanja za uključenje u SR (3. domena) (8).

Nadalje, AMSTAR 2 ima jednostavnije kategorije odgovora nego izvorni alat. Za svaku domenu odgovor može biti „da“ ili „ne“, pri čemu „da“ znači pozitivan rezultat. Samo u pet domena, gdje su autori smatrali da je vrijedno zabilježiti i djelomično ispunjavanje zahtjeva, postoji i opcija „djelomično da“. Uklonjene su opcije odgovora iz izvorne inačice alata „nije primjenjivo“ i „ne mogu odgovoriti“, budući da autori AMSTAR-a 2 sve domene smatraju relevantnima za suvremene SR-ove zdravstvenih intervencija (8). Osim toga, AMSTAR 2 uključuje opsežniji korisnički priručnik i upute za izradu ukupne ocjene na temelju slabosti prepoznatih u ključnim domenama (7,8).

Na temelju podataka objavljenih u literaturi, korištenje AMSTAR-a 2 za procjenu SR-ova zahtijeva više vremena nego korištenje AMSTAR-a. Prvotno su autori AMSTAR-a 2 izvijestili da je za

ocjenu jednog SR-a pomoću alata AMSTAR 2 potrebno od 15 do 32 minute, što je gotovo dvostruko više nego za popunjavanje izvornog AMSTAR-a gdje je vrijeme variralo od 10 do 15 minuta (8). Banzi i sur. objavili su da je za pet ocjenjivača različitog iskustva u ocjenjivanju za potpunu procjenu prema AMSTAR-u bilo potrebno 5,8 minuta (48), dok su Pieper i sur. izvijestili da su četiri ocjenjivača različitog iskustva imala srednje vrijeme za ispunjavanje AMSTAR-a 2 od 18 minuta (49). Zaključno, korištenje AMSTAR-a 2 oduzima mnogo više vremena u odnosu na AMSTAR, što je istraživačima bitno kada treba procijeniti veliki broj SR-ova.

1.8. Korištenje AMSTAR-a 2 za procjenu SR-ova koji ne analiziraju istraživanja intervencija

Za procjenu pouzdanosti SR-ova koji ne analiziraju istraživanja intervencija (ne-intervencijski SR-ovi) ne postoji prikladan alat (50). Stoga autori ponekad u tu svrhu koriste AMSTAR 2 (51,52), premda je on izvorno namijenjen procjeni SR-ova intervencija u zdravstvu (8). Istraživanje objavljeno 2023. ispitivalo je primjenjivost alata AMSTAR 2 na ne-intervencijske SR-ove. U tom su istraživanju autori koristili uzorke od 20 SR-ova za svaku od četiri vrste SR-ova, uključujući one o točnosti dijagnostičkih testova, etiologiji i/ili rizicima, prevalenciji i/ili incidenciji, te prognozi. Tri su autora neovisno pokušala procijeniti svaki uključeni SR prema AMSTAR-u 2. Nakon toga su određivali primjenjivost svake AMSTAR 2 domene na određenu vrstu SR-a, te na bilo koju vrstu SR-a. Rezultati su pokazali da je sedam AMSTAR 2 domena primjenjivo za sve četiri ispitivane vrste SR-ova. Preciznije, to su: domena 2 (unaprijed definiran protokol istraživanja), domena 5 (dvostruki odabir istraživanja), domena 6 (dvostruka ekstrakcija podataka), domena 7 (popis isključenih istraživanja s obrazloženjem), domena10 (izvještavanje o financiranju uključenih istraživanja), domena14 (razmatranje heterogenosti) i domena 16 (sukob interesa i financiranje autora SR-a). Najveći broj domena, njih 13, ocijenjen je primjenjivim za prognostičke SR-ove. Po deset domena bilo je primjenjivo za SR-ove koji istražuju etiologiju/rizik, te za SR-ove o točnosti dijagnostičkih testova. Samo sedam domena ocijenjeno je primjenjivim za SR-ove koji istražuju prevalenciju/incidenciju. Prilikom procjene primjenjivosti svake domene AMSTAR-a 2 na bilo koju vrstu SR-a (tj. izvan navedene četiri specifične vrste), istraživači su smatrali da je osam stavki primjenjivo. Uz sedam stavki koje su već identificirane kao primjenjive na četiri specifične vrste SR-a, domena 4 (korištenje sveobuhvatne strategije pretraživanja) također je ocijenjena kao primjenjiva za bilo koju vrstu SR-a. Autori su zaključili kako je AMSTAR 2 samo djelomično primjenjiv na ne-intervencijske SR-ove, odnosno da postoji potreba za prilagodbom ili proširenjem AMSTAR 2 za ovu vrstu istraživanja (50).

2. CILJEVI RADA I HIPOTEZE

U okviru doktorske disertacije provedena su tri istraživanja.

Cilj prvog istraživanja bio je analizirati učestalost korištenja alata AMSTAR 2 tijekom prve tri godine od njegovog objavljivanja te istražiti potencijalne prepreke njegovom prihvatanju. Hipoteza je bila da će, tijekom prve tri godine od objave alata AMSTAR 2, 50% autora i dalje koristiti AMSTAR (prvu inačicu alata) za procjenu SR-ova

Cilj drugog istraživanja bio je analizirati metodološku kvalitetu i karakteristike SR-ova u kojima je navedeno da su provedeni u skladu s alatom AMSTAR 2. Hipoteze su bile da će među SR-ovima koji su naveli da su koristili AMSTAR 2 za ustroj svojeg istraživanja više od polovice njih biti kritično niske pouzdanosti prema AMSTAR-u 2, te da će manje od 10 % njih sadržavati detaljnu AMSTAR 2 procjenu vlastitog SR-a.

Cilj trećeg istraživanja bio je analizirati pristup izvještavanju i način na koji se AMSTAR 2 koristi za metodološku procjenu SR-ova. Hipoteze su bile da će više od polovice istraživanja koristiti bodove prilikom upotrebe AMSTAR-a 2, te da će manje od polovice istraživanja objaviti svoje detaljne AMSTAR 2 procjene za svaku domenu svakog procijenjenog istraživanja.

3. JEDINICE ANALIZE I POSTUPCI

3.1. Jedinice analize

U sva tri istraživanja bili su korišteni javno dostupni podaci iz publikacija indeksiranih u bazama MEDLINE i Embase. Osim toga, prvo istraživanje sadržavalo je i podatke prikupljene anketiranjem autora istraživanja koji su u definiranom vremenskom razdoblju koristili AMSTAR umjesto AMSTAR 2 alata.

3.2. Etička načela

Za anketiranje ispitanika koje je provedeno u okviru prvog istraživanja, protokol istraživanja odobrilo je 18. ožujka 2020. Etičko povjerenstvo Hrvatskog katoličkog sveučilišta (Klasa 641-03/20-01/02; Urbroj: 498-03-02-06-02/1-20-03). Svi ispitanici dali su svoju privolu za sudjelovanje putem e-pošte. Sve su metode bile provedene u skladu s relevantnim smjernicama i propisima, uključujući Etički kodeks Hrvatskog katoličkog sveučilišta i Helsinšku deklaraciju (53). Za prvi dio prvog istraživanja, kao i drugo i treće istraživanje, nije bilo potrebno odobrenje etičkog povjerenstva s obzirom na to da su u tim istraživanjima bili analizirani isključivo javno dostupni podaci iz radova indeksiranih u bazama MEDLINE i Embase.

3.3. Ustroj istraživanja

Prvo istraživanje bilo je multimetodološko (engl. *mixed-methods study*) i sastojalo se od dva dijela. Prvi dio je bilo presječno meta-istraživanje (metodološko istraživanje objavljene literature). Drugi dio sastojao se od anketiranja autora istraživanja koji su koristili AMSTAR umjesto AMSTAR-a 2. Drugo istraživanje bilo je presječno meta-istraživanje. Treće istraživanje bilo je također presječno meta-istraživanje.

3.4. Kriteriji uključenja i isključenja

U prvo istraživanje bili su uključeni znanstveni članci koji opisuju originalna istraživanja koja su koristila AMSTAR ili AMSTAR 2 za procjenu SR-ova. Isključeni su radovi koji su samo spominjali AMSTAR ili AMSTAR 2, ali nisu ga koristili za procjenu SR-ova. Također su isključeni radovi koji su samo naveli da su pripremljeni u skladu s navedenim alatima te metodološka istraživanja čiji je cilj bio opisivanje ovih alata. Uključeni radovi bili su objavljeni u razdoblju od 1. siječnja 2018. do 31. prosinca 2020. U drugi dio istraživanja, anketu, bili su uključeni autori istraživanja koji su u analiziranom razdoblju koristili AMSTAR umjesto AMSTAR 2 alata.

U drugo istraživanje bili su uključeni cjeloviti članci objavljeni u časopisima u kojima je navedeno da su provedeni/pripremljeni/osmišljeni u skladu s alatom AMSTAR 2. Isključeni su kongresni sažeci i članci koji su bili jasno povezani sa starom verzijom AMSTAR-a, kao i članci koji su spominjali AMSTAR 2 u bilo kojem drugom kontekstu. Nisu bila postavljena nikakva ograničenja u pogledu ispitanika, intervencija (uključeni su SR-ovi svih zdravstvenih intervencija, kao i ne-intervencijski SR-ovi), kontrolne skupine i ishoda. Uključeni radovi bili su objavljeni u razdoblju od 1. siječnja 2018. do 3. svibnja 2022.

U treće istraživanje bili su uključeni znanstveni članci koji su koristili AMSTAR 2 za metodološku procjenu SR-ova ili bilo koje druge vrste sinteze dokaza. Isključeni su članci koji su spominjali ili koristili AMSTAR 2 u bilo kojem drugom kontekstu. Uključeni radovi bili su objavljeni u razdoblju od 1. siječnja 2018. do 5. studenoga 2023.

Za sva tri istraživanja: početni datum pretraživanja bio je 1. siječnja 2018. zbog toga što je AMSTAR 2 objavljen u rujnu 2017. godine.

3.5. Pretraživanje literature

Pretražene su baze MEDLINE i Embase putem OVID-a s ciljem pronalaženja publikacija koje sadrže riječ AMSTAR. Ista široka strategija pretraživanja bila je korištena u sva tri istraživanja:

((((AMSTAR) ILI (AMSTAR-2)) ILI (AMSTAR 2)) ILI (R-AMSTAR).

Pretraživanje nije bilo ograničeno ni na jedan određeni dio članka (tj. nije bilo ograničenja samo na sažetak ili metode). Također nije bilo ograničenja vrste ili jezika publikacije. Zapisi su bili izvezeni u računalni program EndNote X5 (Clarivate Analytics, London, UK) nakon čega su izbrisani duplikati. Postupak pretraživanja literature opisan je u Prilogu 1.

3.6. Probir literature

Nakon automatskog brisanja duplikata, ostatak duplikata izbrisan je manualno. Prvu fazu probira (engl. *screening*) koja se odnosila na naslove i sažetke zapisa pronađenih pretraživanjem radila su neovisno dva autora. Svaki zapis bio je ocijenjen s „da“, „ne“ ili „možda da“ s obzirom na to je li zadovoljavao kriterije uključenja na temelju informacija navedenih u naslovu i sažetku. Svi zapisi koji su označeni kao „da“ i „možda da“, uključeni su u drugu fazu probira u kojoj su pronađeni i procijenjeni cjeloviti tekstovi tih istraživanja. Neslaganja oko uključivanja pojedinih istraživanja riješena su u raspravi s trećim autorom.

3.7. Prikupljanje podataka

U tijeku prvog istraživanja, izrađena je pilot tablica s uzorkom od 20 uključenih istraživanja. Dva autora neovisno su prikupljala podatke, a nakon postizanja visoke razine slaganja među autorima, dalje je prikupljanje podataka radio jedan, a provjeravao drugi autor. Neslaganja su riješena savjetovanjem s trećim autorom.

U drugom istraživanju, prikupljanje bibliografskih podataka učinjeno je prema istim načelima kao u prvom istraživanju. Metodološku procjenu pouzdanosti uključenih SR-ova prema AMSTAR 2 alatu najprije su na uzorku od pet istraživanja napravila neovisno tri autora, a rezultati su uspoređeni i detaljno raspravljani na sastanku, tj. svaka AMSTAR 2 domena i ukupna procjena svakog od tih pet istraživanja zasebno je razmotrena. Kasnije su metodološku procjenu radila neovisno dva autora, a neslaganja su bila riješena konzultacijom s trećim autorom.

U tijeku trećeg istraživanja, izrađena je pilot tablica s uzorkom od 20 uključenih istraživanja iz kojih su podatke prikupljala neovisno po tri autora. Nakon postizanja visoke razine slaganja među autorima, dalje su podatke prikupljala neovisno dva autora, a neslaganja su riješena kroz raspravu i konzultaciju s trećim autorom. Ekstrakcija podataka provedena je u elektroničkoj tablici za nasumično odabranih 50 % članaka koji su uključeni nakon druge faze probira. Programski kôd koji je korišten za nasumični odabir dostupan je u Prilogu 2.

3.8. Anketiranje ispitanika

U svibnju 2020. (za članke objavljene 2018. - 2019.), te u siječnju i veljači 2022. (za članke objavljene 2020.) poslana su e-poruke dopisnim autorima uključenih istraživanja koja nisu koristila AMSTAR 2. Korištene su isključivo adrese e-pošte navedene u objavljenim člancima. Nije bilo pokušaja pronalaženja alternativnih adrese e-pošte ako bi se poruka vratila nedostavljena ili nije bila odgovorena. Svaki potencijalni ispitanik dobio je dva podsjetnika u razmaku od tjedan dana. Kriterij uključenja za te autore bio je da je rukopis poslan nakon rujna 2017., odnosno nakon što je AMSTAR 2 bio objavljen. Tekst ankete poslana e-poštom dostupan je u Prilogu 3.

Ispitanici su odgovarali na pitanja putem ankete na *Google Forms*-u. U pozivu su ispitanici dobili detaljne informacije o svrsi istraživanja. U anketi su bila pitanja jesu li upoznati s AMSTAR-om 2, razloge zbog kojih nisu koristili AMSTAR 2 umjesto AMSTAR-a, jesu li urednici ili recenzenti predlagali korištenje AMSTAR-a 2, te da navedu bilo kakve prepreke za usvajanje AMSTAR-a 2 s kojima su se susreli ili s kojima bi se netko drugi mogao susresti. Dodatno, upitani su o godinama iskustva u sintezi dokaza ili metodološkim istraživanjima. Nakon slanja ankete, javilo se nekoliko

autora iz Kine s primjedbom da ne mogu pristupiti anketi. Iz tog su razloga pregledane afilijacije svih autora za dopisivanje. Nakon toga su svim autorima s afilijacijama u Kini pitanja iz ankete poslana u sadržaju e-pošte (bez poveznice za *Google Forms*).

3.9. Ishodi istraživanja

3.9.1. Ishodi prvog istraživanja

Glavne mjere ishoda:

- broj (postotak) istraživanja koja su koristila AMSTAR 2,
- broj (postotak) istraživanja koja su koristila AMSTAR,
- broj (postotak) autora uključenih u anketu kojima je bilo poznato da je objavljena nova verzija alata, tj. AMSTAR 2

Sporedne mjere ishoda:

Iz podataka dostupnih u uključenim člancima analizirali smo: naziv časopisa, mjesec slanja rada u časopis, mjesec prihvatanja rukopisa, mjesec objave rada *online*, vrste istraživanja (OSR, metodološko istraživanje), te reference navedene za korištenje AMSTAR-a ili AMSTAR-a 2.

Iz podataka dobivenih anketiranjem autora analizirali smo sljedeće:

- razloge zašto su koristili AMSTAR, a ne AMSTAR 2,
- jesu li urednici ili recenzenti sugerirali da bi trebali koristiti AMSTAR 2 [da/ne],
- prepreke za korištenje AMSTAR-a 2,
- broj godina iskustva u području sinteze dokaza ili metodoloških istraživanja.

3.9.2. Ishodi drugog istraživanja

Glavna mjera ishoda: broj (postotak) uključenih SR-ova za koje smo korištenjem alata AMSTAR 2 procijenili da su visoke, umjerene, niske, odnosno kritično niske pouzdanosti.

Sporedne mjere ishoda bili su podaci o analiziranom članku: ime prvog autora, godina objave članka, naziv časopisa, vrsta istraživanja, kopija dijela teksta u kojem se navodi da je istraživanje u skladu s alatom AMSTAR 2 te je li članak uključivao dodatni dokument u kojem je detaljno navedena procjena vlastitog SR-a pomoću alata AMSTAR 2 (da/ne).

3.9.3. Ishodi trećeg istraživanja

Glavne mjere ishoda:

- Broj autora koji su proveli procjenu prema AMSTAR-u 2 [jedan, dva, više od dva]
- Je li više autora neovisno provelo AMSTAR 2 procjene? [da/ne]
- Jesu li naveli razinu suglasja među procjeniteljima (engl. *inter-rater agreement*)? [da/ne]
 - Ako da, kako? [kopirati tekst]
- Jesu li autori u procjeni koristili prosudbe (da, djelomično da, ne) koje su preporučili autori alata AMSTAR 2? [da/ne]
- Jesu li autori koristili ključne domene? [da/ne]
 - Ako da, jesu li koristili sedam ključnih domena koje su preporučili autori AMSTAR-a 2 (domene 2, 4, 7, 9, 11, 13, 15) [da/ne]?
 - Ako ne, jesu li procijenili drugačije ključne domene? [da/ne]
 - Ako da, koje su domene smatrali ključnima? [kopirati tekst]
- Jesu li autori koristili samo ključne domene (bez korištenja ostalih domena)? [da/ne]
- Jesu li autori dodijelili numeričke bodove AMSTAR 2 domenama? [da/ne]
- Jesu li naveli postotak domeni u skladu s AMSTAR-om 2 za svaki sustavni pregled? [da ne]
- Jesu li naveli postotak usklađenosti po domeni? [da/ne]
- Jesu li naveli ukupan numerički rezultat za svaki sustavni pregled? [da/ne]
- Jesu li naveli konačnu ocjenu (tj. visoku, umjerenu, nisku, vrlo nisku pouzdanost) za svaki uključeni sustavni pregled? [da/ne]
 - Ako da, jesu li izvijestili o prosudbi za svaki uključeni sustavni pregled u glavnom dijelu članka ili u dodatku? [glavni dio članka/dodatak]
 - Ako da, jesu li naveli obrazloženje za prosudbu svakog sustavnog pregleda? [da ne]
 - Ako da, jesu li naveli obrazloženje prosudbe u glavnom dijelu članka ili u dodatku [glavni dio članka/dodatak]
- Jesu li izvijestili o prosudbi za svaku domenu (tj. da, djelomično da, ne) svakog uključenog članka? [da ne]

- Ako da, jesu li to prijavili u glavnom dijelu članka ili u dodatku [glavni dio članka/dodatak]
- Ako da, jesu li prijavili obrazloženje za prosudbu za svaku domenu svakog uključenog članka? [da ne]
- Ako da, jesu li to prijavili u glavnom dijelu članka ili u dodatku [glavni dio članka/dodatak]

Sekundarne mjere ishoda bili su sljedeći podaci o analiziranom radu: ime prvog autora, godina objave, naziv časopisa te koja je vrsta ustroja istraživanja uključena u analizu [na primjer, sustavni pregled, brzi pregled, pretražni sustavni pregled (engl. *scoping review*), itd.].

3.10. Statistička analiza

U sva tri istraživanja podaci su analizirani pomoću deskriptivne statistike. Podaci su predstavljeni s učestalostima i postocima korištenjem računalnog programa MedCalc (MedCalc Software bv., Ostend, Belgija).

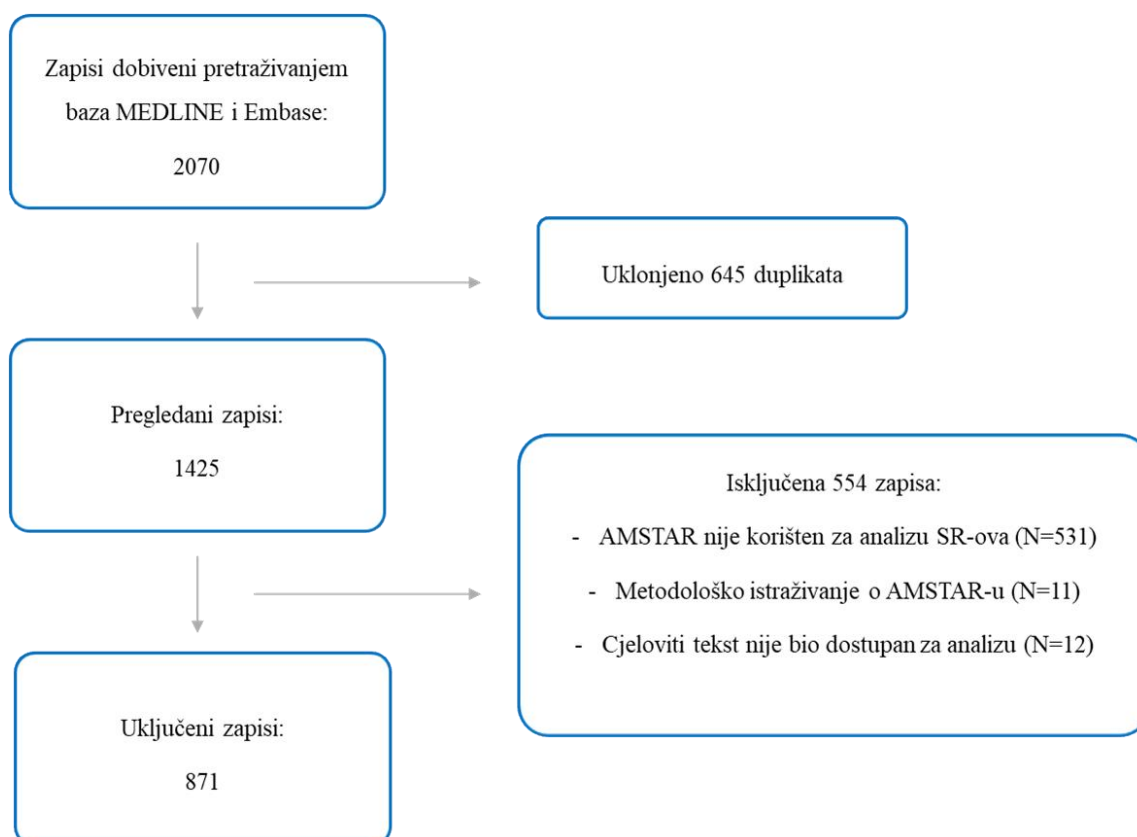
4. REZULTATI

4.1. Rezultati prvog istraživanja

4.1.1. Bibliometrijska analiza

Pretraživanjem baza podataka dobiveno je ukupno 2070 zapisa. Nakon uklanjanja duplikata, pregledano je 1425 zapisa. Isključena su 542 zapisa jer nisu odgovarala kriterijima uključenja, odnosno nisu koristila istraživani alat za procjenu SR-ova. Većina isključenih istraživanja samo je spomenula da je njihov SR bio pripremljen u skladu s AMSTAR-om. Dvanaest istraživanja nije bilo moguće analizirati zbog nedostupnosti cjelovitog teksta. Popis isključenih i nedostupnih istraživanja, s pojedinačno navedenim razlozima isključenja, kao i svi neobrađeni podaci (engl. *raw data*) prikupljeni u tijeku ovog istraživanja dostupni su u obliku Dodatne datoteke 1 i Dodatne datoteke 2, redom, koje su dostupne na poveznici <https://osf.io/gj28f/files/osfstorage>. Uključeno je 871 istraživanje. Slika 2. prikazuje dijagram tijeka prvog istraživanja.

Slika 2. Dijagram tijeka prvog istraživanja.



Većina uključenih istraživanja (70 %) bili su OSR-ovi, nakon čega su slijedila metodološka istraživanja (18 %). Najveći dio njih bili su cjeloviti tekstovi objavljeni u znanstvenim časopisima. Uključena istraživanja bila su objavljena u 513 različitih časopisa. Najviše zastupljeni časopisi bili su: *BMJ Open*, *Systematic Reviews*, *PLoS One*, *Medicine* i *Journal of Clinical Epidemiology*. AMSTAR 2, odnosno novija inačica alata, bila je korištena u 52 % istraživanja, dok je AMSTAR, odnosno starija inačica alata, bila korištena u 44 % istraživanja. Tijekom 2018., 81 % analiziranih istraživanja koristilo je AMSTAR, dok je 16 % koristilo AMSTAR 2. Tijekom 2019., 52 % je koristilo AMSTAR, dok je 44 % koristilo AMSTAR 2. Među istraživanjima koja su objavljena tijekom 2020., 28 % je koristilo AMSTAR, a 69 % AMSTAR 2. U 31 istraživanju bio je korišten R-AMSTAR. Tablica 3. detaljno prikazuje značajke uključenih istraživanja.

Tablica 3. Značajke uključenih istraživanja (N=871)

Varijabla	N (%)*
Ustroj istraživanja	
Pregled sustavnih pregleda (engl. <i>overview of systematic reviews</i> , OSR)	612 (70)
Metodološko istraživanje	160 (18)
Protokol za OSR	61 (7.0)
Smjernice	29 (3.3)
Politički sažetak sa sintezom dokaza	1 (0.01)
Procjena zdravstvenih tehnologija	1 (0.01)
Protokol za metodološko istraživanje	1 (0.01)
Metodološko istraživanje i OSR	1 (0.01)
Procjena brzog pregleda	1 (0.01)
Procjena smjernica	1 (0.01)
Protokol za smjernice	1 (0.01)
Konsenzusna izjava	1 (0.01)
Vrsta izvješća	
Cjeloviti tekst istraživanja	826 (95)
Kongresni sažetak	45 (5.2)
Najčešći časopisi u kojima su objavljena analizirana istraživanja	
BMJ Open	31 (3.6)
Systematic Reviews	23 (2.6)
PLoS One	22 (2.5)
Medicine	22 (2.5)
Journal of Clinical Epidemiology	19 (2.2)
Chinese Journal of Evidence-Based Medicine	14 (1.6)
Cochrane Database of Systematic Reviews	13 (1.5)
Evidence-based Complementary and Alternative Medicine	11 (1.3)
International Journal of Environmental Research and Public Health	11 (1.3)
Korištena inačica AMSTAR alata	
AMSTAR 2	451 (52)
AMSTAR	382 (44)
R-AMSTAR	31 (3.6)

AMSTAR plus AMSTAR 2	5 (0.006)
AMSTAR plus R-AMSTAR	1 (0.001)
Nejasno o kojoj se inačici radi	1 (0.001)
Reference koje su korištene kao potpora korištenju alata AMSTAR (N=382)	
Shea et al., 2007, BMC Medical Research Methodology [1]	216 (57)
Shea et al., 2009 [19]	94 (25)
Shea et al., 2007, PLoS One [20]	20 (5.2)
Shea et al., 2017 [2]	11 (2.9)
AMSTAR web site	6 (1.6)
Pieper et al., 2015 [13]	6 (1.6)
Sharif et al., 2013 [21]	6 (1.6)
Pollock et al., 2017 [22]	4 (1.0)
Xiong et al., 2009 [23]	3 (0.8)
Reference koje su korištene kao potpora korištenju alata AMSTAR 2 (N=451)	
Shea et al., 2017 [2]	396 (88)
Shea et al., 2007, BMC Medical Research Methodology [1]	14 (3.1)
Shea et al., 2009 [19]	9 (2.0)
Mrežna stranica alata AMSTAR	8 (1.8)
Zhang et al., 2018	3 (0.07)
Lorenz et al., 2019 [24]	2 (0.04)
Shea et al., 2007 PLoS One [20]	2 (0.04)
Ge et al., 2017 [25]	2 (0.04)
Banzi et al. 2018	1 (0.02)
Biondi-Zoccai, 2016 [26]	1 (0.02)
Brouwers et al., 2010 [12]	1 (0.02)
Ciapponi, 2017 [27]	1 (0.02)
Pieper et al., 2014 [28]	1 (0.02)
Pollock et al., 2017 [22]	1 (0.02)
Tian et al., 2017 [29]	1 (0.02)
Xiong et al., 2009 [30]	1 (0.02)
Yan et al., 2018 [31]	1 (0.02)
Reference koje su korištene kao potpora korištenju alata R-AMSTAR (N=12)	
Kung et al., 2010 [5]	23 (75)
Shea et al., 2017 [2]	3 (9.7)
Shea et al., 2007, BMC Medical Research Methodology [1]	3 (9.7)
Dosenovic et al., 2018 [17]	
Rotta et al., 2015 [32]	1
Shea et al., 2007 PLoS One [20]	

* Postoci za svaku varijablu možda neće iznositi ukupno 100 zbog zaokruživanja.

Šest istraživanja koristilo je po dva alata: pet istraživanja AMSTAR zajedno s AMSTAR-om 2, te jedno istraživanje AMSTAR zajedno s R-AMSTAR-om (Tablica 3.). U tim je istraživanjima analizirano jesu li autori naveli objašnjenje za korištenje dvaju alata. Sharma i Oremus (54) naveli su da je AMSTAR 2 objavljen neposredno prije slanja njihovog rada u časopis pa su odlučili dodati i procjene prema novijoj verziji alata. McGuire i sur. (55) nisu eksplicitno naveli zašto su koristili dva alata. Oni su u metodama opisali obilježja obaju alata, a u rezultatima su spomenuli kako

određene analize nisu mogli provesti s AMSTAR-om 2 jer [citat] *“Ovaj alat nema numerički sustav bodovanja.”* (55). Kim i sur. (56) usporedili su procjene dobivene korištenjem alata AMSTAR i AMSTAR 2 uz objašnjenje da [citat]: *“Nije jasno mogu li se slične procjene metodološke kvalitete provesti za isto istraživanje, jer AMSTAR i AMSTAR 2 imaju različit broj domena za procjenu, različit sadržaj domena, metode procjene i metode izračuna rezultata procjene.”* (56). De Santis i Kaplan (57) usporedile su rezultate procjena dobivenih AMSTAR-om i AMSTAR-om 2 objasnivši da [citat]: *“Budući da se AMSTAR i AMSTAR 2 značajno razlikuju, ostaje nejasno daju li slične ocjene kvalitete za sustavne preglede u zdravstvu.”* (57). Jeyaraman i sur. (58) koristili su i AMSTAR i AMSTAR 2 bez ikakvih komentara ili objašnjenja za korištenje oba alata. U jednom istraživanju (59) koje je koristilo i AMSTAR i R-AMSTAR bilo je navedeno [citat]: *“Također smo koristili revidiranu verziju AMSTAR-a (R-AMSTAR), koja dodjeljuje ukupnu ocjenu kvalitete sustavnom pregledu”* (59).

4.1.2. Izvori informacija koji su referencirani za korištenje istraživanih alata

U analiziranim istraživanjima autori su naveli 41 različitu referencu kako bi poduprli korištenje AMSTAR-a, R-AMSTAR-a ili AMSTAR-a 2. Medijan broja referenci korištenih za te alate bio je 1 (raspon: 0 do 3). Istraživanja koja su koristila samo izvornu inačicu AMSTAR-a (N=382) navela su 31 različitu referencu za korištenje alata. Većina (N=216; 57 %) je kao referencu navela istraživanje iz 2007. u kojemu je alat prvi put opisan, a objavili su ga Shea i sur. (7). Druge najčešće korištene reference za AMSTAR uključivale su radove Shea i sur. iz 2007. (29) i 2009. (28), koji opisuju daljnja testiranja alata, ili pak referencu na mrežnu stranicu AMSTAR-a (38). Više autora citiralo je istraživanja drugih autora koja su ispitivala AMSTAR. Jedanaest istraživanja koristilo je izvorni AMSTAR, ali su pogrešno referencirala rad Shea i sur. iz 2017., koji opisuje AMSTAR 2. Vidi Tablicu 3.

Autori istraživanja koja su koristila samo AMSTAR 2 (N=451) naveli su 17 različitih referenci za njegovo korištenje. Većina (88 %) referencirala je članak Shea i sur. iz 2017., u kojem je alat opisan. Međutim, više je autora također navodilo reference na članak koji opisuje izvornu inačicu alata ili druga istraživanja objavljena prije 2017. Jedno je istraživanje referenciralo članak o AGREE II alatu (60). Vidi Tablicu 3.

Među 31 istraživanjem koje je koristilo samo R-AMSTAR, 23 su referencirala članak Kunga i sur. iz 2010., u kojemu je R-AMSTAR opisan (9), dok su ostali navodili reference na AMSTAR ili AMSTAR 2, a neki čak i reference na druge radove koji su koristili R-AMSTAR. Vidi Tablicu 3.

4.1.3. Rezultati ankete

Kriterije za uključenje u anketiranje zadovoljila su 354 istraživanja. U 11 istraživanja nije bila navedena e-adresa autora za dopisivanje. Stoga su na sudjelovanje u anketi pozvana 343 autora. Za 19 poruka je stigla obavijest da poruka nije dostavljena. Od preostalih 324 autora, odgovorilo ih je 88 (stopa odgovora: 27 %), od čega 79 putem *Google Forms*-a, a devet putem e-pošte. Medijan godina istraživačkog iskustva autora koji su ispunili anketu bio je 13 (raspon: 1 do 25).

Među 88 ispitanika, 68 (77 %) izjavilo je da su znali da je AMSTAR 2 objavljen. Deset ispitanika (11 %) je navelo da su urednici ili recenzenti od njih zatražili korištenje AMSTAR-a 2. Među 68 ispitanika koji su znali za objavu AMSTAR-a 2, 41 (60 %) je izjavilo da su čuli za objavu AMSTAR-a 2 prije nego što su poslali svoj rad u časopis. Dvadeset i pet (38 %) od tih 68 ispitanika izjavilo je da su razmatrali korištenje AMSTAR-a 2 u svom istraživanju umjesto AMSTAR-a.

Razloge za nekorisćenje AMSTAR-a 2 navela su 44 autora. Autori su navodili da nisu koristili AMSTAR 2 jer psihometrijska svojstva tog alata nisu bila utvrđena u to vrijeme, njihov protokol je već bio zaključen, prikupljanje i/ili analiza podataka završeno je prije objave AMSTAR-a 2, nisu znali za objavu AMSTAR-a 2, nisu imali vremena za novu analizu, alat je bio dulji od izvornog AMSTAR-a, te urednici i recenzenti to nisu tražili. Tablica 4. prikazuje razloge zbog kojih autori nisu koristili AMSTAR 2.

Na pitanje postoje li prepreke za primjenu AMSTAR-a 2 koje su iskusili ili koje bi netko drugi mogao iskusiti, odgovorilo je 68 ispitanika. Većina je navela da nema prepreka. Oni koji su spomenuli prepreke naveli su sljedeće: nedostatak kvantitativnog aspekta ocjenjivanja s AMSTAR-om 2, nedostatak znanja o postojanju tog alata, duljina alata, odnosno više vremena potrebno za korištenje novog alata, nedovoljno poznavanje alata, i poteškoće s određenom domenom. Tablica 5. prikazuje prepreke koje su ispitanici naveli za korištenje AMSTAR-a 2.

Tablica 4. Razlozi zbog kojih autori nisu koristili AMSTAR 2

Razlog	N
Psihometrijska svojstva AMSTAR-a 2 u to vrijeme nisu bila utvrđena.	2
AMSTAR 2 je izašao nakon što smo završili prikupljanje podataka koristeći AMSTAR.	2
AMSTAR još nije bio objavljen u vrijeme kada smo razvili naš protokol i poslali rukopis na recenziju.	2
Već smo razvili protokol i započeli istraživanje.	1
Već smo bili završili procjenu kvalitete korištenjem AMSTAR-a	2
Istraživanje smo započeli puno prije nego što je AMSTAR 2 objavljen, a urednici ni recenzenti nisu tražili da koristimo AMSTAR 2.	1
AMSTAR su i dalje koristili drugi autori.	1
Zato što smo protokol u kojem smo se odlučili za korištenje AMSTAR-a objavili prije objave AMSTAR-a 2. Štoviše, imali smo razloga vjerovati da se rangiranje uključenih istraživanja neće značajno promijeniti upotrebom AMSTAR-a 2.	1
Budući da je pretraga provedena do ožujka 2017. godine, a kritička procjena nakon toga, to je odgovaralo završetku konačnog rukopisa prije objave AMSTAR 2.	1
Nisam za to čuo na vrijeme.	1
Čuo sam za AMSTAR 2 nakon što je rukopis bio poslan.	1
Objavljen je upravo kad sam slao članak pa nisam donio odluku o promjeni.	1
Jer je duži od AMSTAR-a	1
Članak koji sam objavio bio je dio moje doktorske disertacije, koja je imala zadani rok za prijavu obrane, što mi je onemogućilo da napravim promjene i koristim AMSTAR 2. Zaista, nedostatak vremena me spriječio da koristim taj instrument.	1
Preblizu slanja rukopisa u časopis (nije bilo vremena za ponavljanje analize).	1
Nisam bio svjestan da je objavljena nova inačica alata.	1
Dok smo provodili istraživanje, nismo bili svjesni AMSTAR-a 2. Postali smo svjesni kasnije.	1
Naš pregled bio je završen kada smo saznali za AMSTAR 2. To je bio pregled sustavnih pregleda opservacijskih istraživanja (ne RCT-ova). Nismo smatrali da bi AMSTAR 2 pridonio našim procjenama.	1
Istraživanje smo započeli 2016. godine. Do vremena kada je AMSTAR 2 objavljen 2017., završili smo prikupljanje podataka i analizu te odlučili nastaviti dalje.	1
Ne u objavljenom, ali smo koristili AMSTAR 2 u narednom istraživanju.	1
Koristio sam R-AMSTAR u svom istraživanju.	1

Tablica 5. Prepreke koje su ispitanici naveli za korištenje AMSTAR-a 2

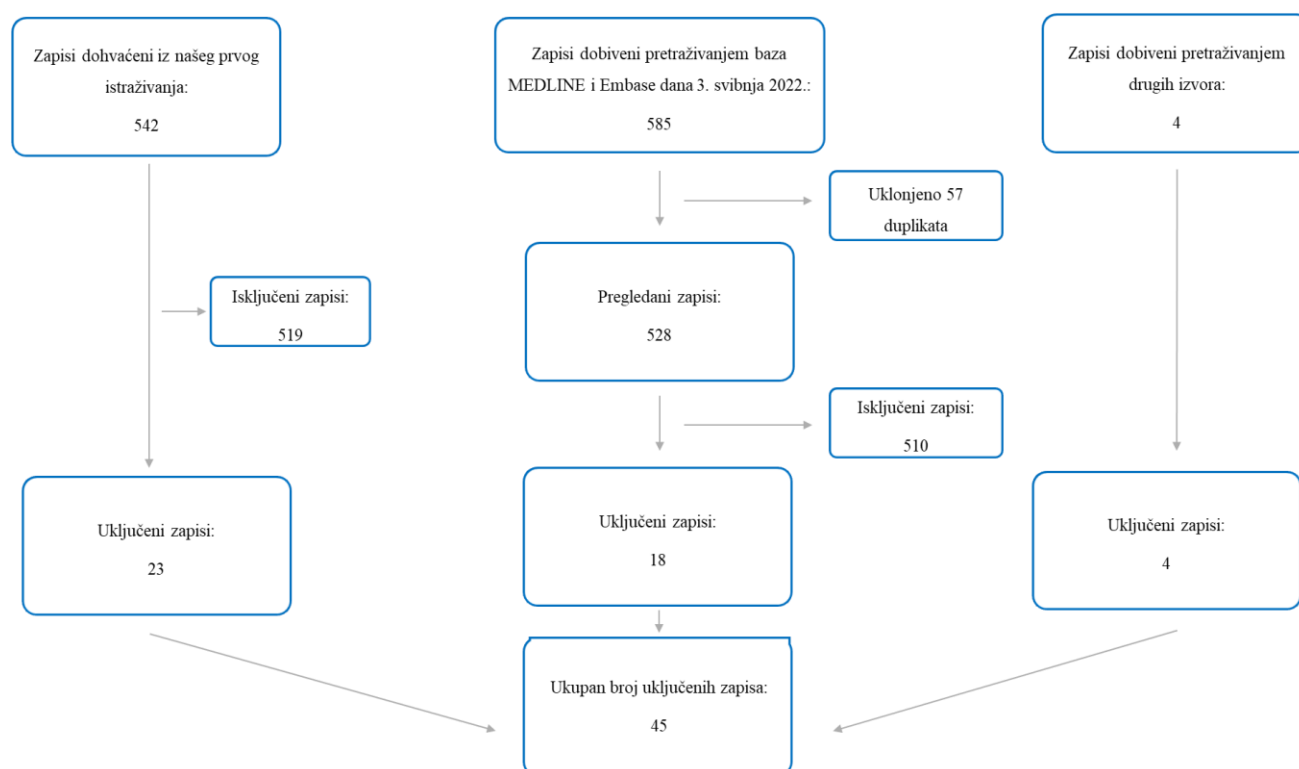
Prepreka	N
Nedovoljna informiranost.	3
Duži je, pa je moguće da zbog toga zahtijeva više vremena za korištenje u usporedbi s AMSTAR-om.	3
Nedostatno poznavanje alata i nesigurnost oko toga kako se razlikuje ili je bolji od AMSTAR-a.	1
Novi je alat, pa ljudi moraju uložiti vrijeme da nauče kako ga koristiti, dok su vjerojatno već upoznati sa starom verzijom AMSTAR-a.	1
Imam poteškoća s pitanjem o pristranosti u objavljivanju, kada nije bilo moguće provesti ovu analizu zbog malog broja studija. Nema odgovarajuće opcije.	1
AMSTAR 2 još nije u širokoj upotrebi.	1

4.2. Rezultati drugog istraživanja

4.2.1. Rezultati pretraživanja

Za istraživanja objavljena prije 2021. pregledana su 542 zapisa s popisa isključenih istraživanja iz prvog istraživanja o AMSTAR-u 2 (61) koji je javno dostupan kao Dodatna datoteka 1 na poveznici <https://osf.io/gj28f/files/osfstorage>. Od toga je isključeno 519 i uključeno 23 zapisa. Pretraživanje istraživanja objavljenih od 2021. nadalje dalo je 585 zapisa iz baza podataka. Nakon uklanjanja zapisa koji su se ponavljali, pregledano je 528 zapisa. Isključeno je 510 zapisa koji nisu ispunjavali kriterije uključenja. Uključeno 18 zapisa. Popis isključenih i nedostupnih istraživanja, s razlozima, javno je dostupan kao Dodatna datoteka 3 na poveznici <https://osf.io/gj28f/files/osfstorage>. Kad se pribroje još četiri uključena zapisa dobivena pretraživanjem drugih izvora, ukupno je uključeno 45 zapisa. Slika 3. prikazuje dijagram tijeka drugog istraživanja. Svi neobrađeni podaci prikupljeni za vrijeme istraživanja dostupni su kao Dodatna datoteka 4 na poveznici <https://osf.io/gj28f/files/osfstorage>.

Slika 3. Dijagram tijeka drugog istraživanja



4.2.2. Značajke uključenih sustavnih pregleda

Od 45 uključenih SR-ova, većina su bili SR-ovi intervencija koji su uključivali primarna istraživanja na ljudima. Više od polovice bili su SR-ovi s MA. Dva članka opisivala su protokole SR-a. Dvadeset sedam članaka eksplicitno je navelo da su napisani u skladu s AMSTAR-om 2, a za preostalih 18 članaka nije bilo jasno iz cjelovitog teksta niti iz referenci jesu li koristili AMSTAR 2 ili AMSTAR. Međutim, s obzirom na datum objavljivanja, pretpostavili smo da su koristili AMSTAR 2. Najveći broj uključenih članaka objavljen je u *International Journal of Surgery* (13 članaka) i u *Annals of Medicine and Surgery* (pet članaka). Većina analiziranih članaka objavljeni su 2020. i 2021. Tablica 6. prikazuje značajke uključenih sustavnih pregleda.

Tablica 6. Značajke uključenih sustavnih pregleda (N=45)

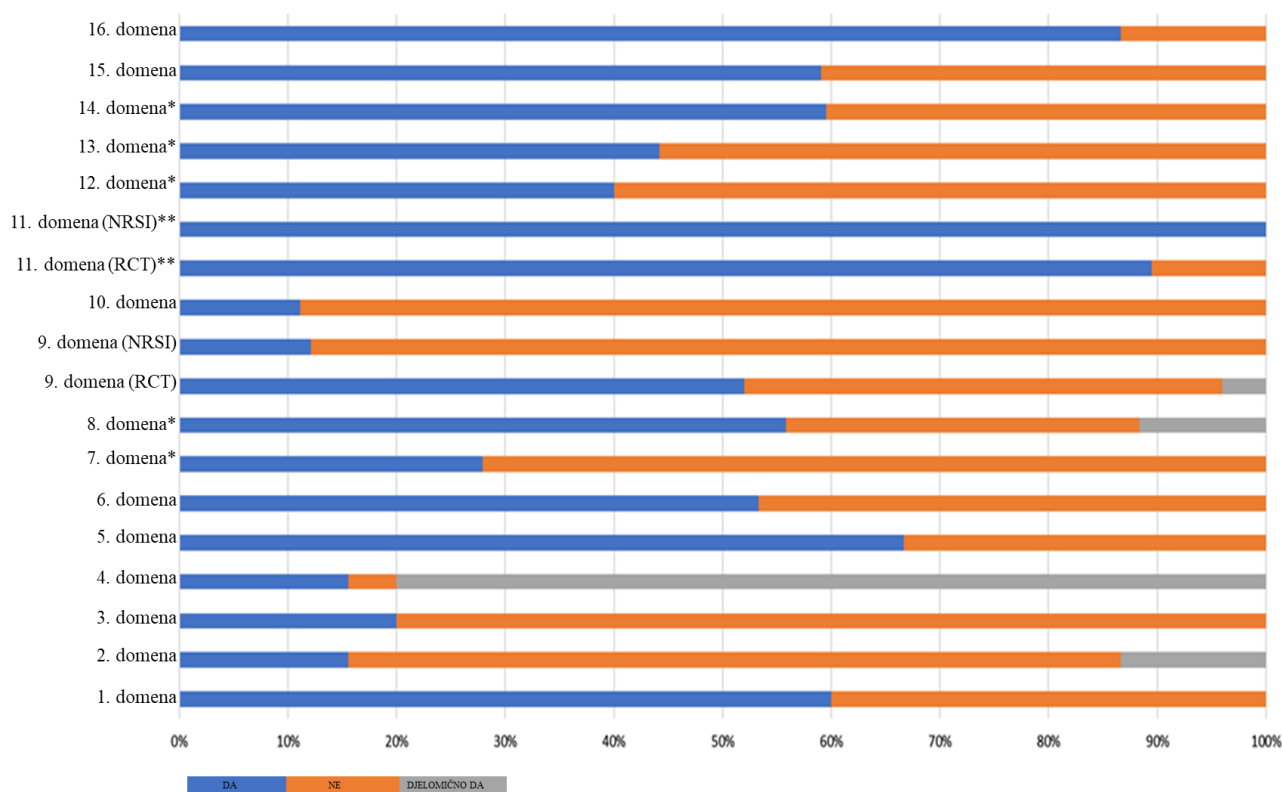
Varijabla	N (%)
Vrste sustavnih pregleda s obzirom na intervenciju	
Sustavni pregledi intervencija koji su uključili primarna istraživanja na ljudima	35 (78)
Sustavni pregledi koji su analizirali intervencije u smjernicama	2 (4)
Ne-intervencijski sustavni pregledi	8 (18)
Korištenje meta-analize	
Da	23 (51)
Ne	20 (44)
Nije primjenjivo (protokol)	2 (4)
Vrsta izvješća	
Cjeloviti sustavni pregled	43 (96)
Protokol sustavnog pregleda	2 (4)
Navođenje samoprocjene korištenjem AMSTAR-a 2	
Da	7 (16)
Ne	38 (84)
Časopis*	
International Journal of Surgery	13 (29)
Annals of Medicine and Surgery	5 (11)
Medicine	2 (4)
Surgical Endoscopy	2 (4)
Godina objave	
2018	4 (9)
2019	9 (20)
2020	13 (29)
2021	15 (33)
2022	4 (9)

* Prikazano u tablici samo za časopise u kojima su objavljena više od dva sustavna pregleda.

4.2.3. Metodološka procjena kvalitete uključenih sustavnih pregleda

Prema ukupnoj procjeni metodološke kvalitete cjelovitih izvještaja SR-ova, 35 (81 %) članaka bilo je kritično niske pouzdanosti, 7 (16 %) članaka bilo je niske pouzdanosti, a 1 (2 %) je članak bio visoke pouzdanosti. Nije bilo članaka umjerene pouzdanosti. Od 11 domena koje su primijenjene na dva uključena SR protokola, oba SR protokola pridržavala su se 7 domena (domene 1, 2, 5, 6, 10, 15 i 16). Slika 4. prikazuje sažetak procjena korištenjem AMSTAR-a 2. Detaljne procjene AMSTAR 2 s razlozima i komentarima za svaku domenu svakog uključenog članka dostupne su kao Dodatna datoteka 5 na poveznici <https://osf.io/gj28f/files/osfstorage>.

Slika 4. Sažetak procjena korištenjem AMSTAR-a 2



Sedam članaka (pet SR-ova intervencija i dva SR-a koja nisu analizirala intervencije) sadržavali su AMSTAR 2 samoprocjenu (62–68), a među njima su tri detaljno navela samoprocjenu za svaku pojedinačnu domenu u tablici ili u dodatnim materijalima (63,64,67). U dva članka (64,67) korišten je popis za procjenu SR-ova po AMSTAR 2 alatu koji se nalazi na mrežnoj stranici alata (38). Ni jedna od sedam samoprocjena ukupne pouzdanosti nije se podudarala s procjenama učinjenim za

potrebe ovog istraživanja. Tablica 7. prikazuje usporedbu procjene metodološke kvalitete korištenjem AMSTAR-a 2 sa samoprocjenama autora.

Tablica 7. Usporedba procjene metodološke kvalitete korištenjem AMSTAR-a 2 sa samoprocjenama autora.

Naziv istraživanja	Samoprocjene autora - objašnjenja	Samoprocjena autora o ukupnoj razini pouzdanosti	Naša procjena ukupne razine pouzdanosti	Kratko objašnjenje naše procjene
Imani, 2019 (62)	Autori nisu dali detaljnu samoprocjenu svake AMSTAR 2 domene ni u glavnom dijelu teksta niti u dodatnom dokumentu. U procjeni su koristili bodove, što nije u skladu s preporukama za AMSTAR 2. Nisu naveli vlastitu ukupnu ocjenu. Iz izjave " <i>Ostvaren je rezultat 15 od 16 bodova</i> " ne možemo zaključiti o ukupnoj metodološkoj kvaliteti sustavnog pregleda jer ne znamo odnosi li se stavka koja je ocijenjena s "Ne" na ključnu ili ne-ključnu domenu.	Nepoznato	Kritično niska pouzdanost	2 „Ne“ u ključnim i 3 „Ne“ u ne-ključnim domenama
Al Hammadi, 2019 (63)	Samoprocjena prema AMSTAR 2 prikazana je u tablici u glavnom dijelu teksta. Autori su svoj sustavni pregled ocijenili s "Da" za sve primjenjive domene. Nisu naveli vlastitu ukupnu ocjenu. Sudeći prema njihovoj samoprocjeni po pojedinim domenama, možemo zaključiti da su smatrali kako je njihov pregled ukupno gledano visoke razine pouzdanosti.	Visoka pouzdanost	Niska pouzdanost	1 „Ne“ u ključnim i 1 „Ne“ u ne-ključnim domenama
Shu, 2018 (64)	Samoprocjena prema AMSTAR-u 2 prikazana je u Dodatku koji je dostupan na sljedećoj poveznici: https://ars.els-cdn.com/content/image/1-s2.0-S1743919118317047-mmc4.pdf . Autori su 10. domenu ocijenili s "Ne", a 2. i 4. domenu s "Djelomično da". Sudeći prema njihovoj samoprocjeni po pojedinim domenama, možemo	Umjerena pouzdanost	Kritično niska pouzdanost	3 „Ne“ u ključnim i 5 „Ne“ u ne-ključnim domenama

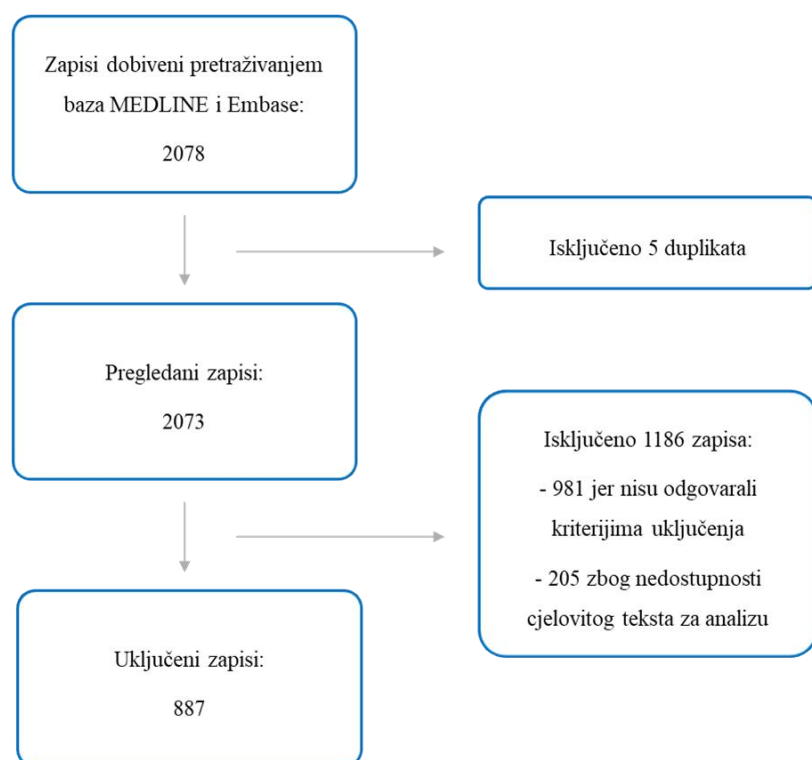
	zaključiti da bi njihova ukupna samoprocjena razine pouzdanosti mogla odgovarati umjerenoj razini pouzdanosti.			
Hailu, 2022 (65)	Autori nisu dali detaljnu samoprocjenu svake AMSTAR 2 domene ni u glavnom dijelu tekstu ni u dopunskom dokumentu. Svoju su ukupnu samoprocjenu opisali kao „visoka kvaliteta“.	Visoka pouzdanost	Kritično niska pouzdanost	2 „Ne“ u ključnim i 5 „Ne“ u neključnim domenama
Beyable, 2022 (66)	Autori nisu dali detaljnu samoprocjenu svake AMSTAR 2 domene ni u glavnom dijelu tekstu ni u dopunskom dokumentu. U članku svoju ukupnu samoprocjenu opisali su kao „umjerena kvaliteta“.	Umjerena pouzdanost	Kritično niska pouzdanost	4 „Ne“ u ključnim i 7 „Ne“ u neključnim domenama
Fernández, 2021 (67)	Samoprocjena prema AMSTAR 2 prikazana je u dodatnoj datoteci. Sve domene ocijenjene su s "Da", osim 9. domene koja je ocijenjena s "Djelomično da". Autori nisu naveli vlastitu ukupnu ocjenu. Sudeći prema njihovoj samoprocjeni po pojedinim domenama, možemo zaključiti da su smatrali kako je njihov sustavni pregled ukupno gledano visoke razine pouzdanosti.	Visoka pouzdanost	Kritično niska pouzdanost	4 „Ne“ u ključnim i 3 „Ne“ u neključnim domenama
Simegn, 2021 (68)	Autori nisu dali detaljnu samoprocjenu svake domene ni u glavnom dijelu članka ni u dopunskom dokumentu. U glavnom dijelu su svoju ukupnu samoprocjenu opisali su kao „umjerena kvaliteta“.	Umjerena pouzdanost	Kritično niska pouzdanost	4 „Ne“ u ključnim i 8 „Ne“ u neključnim domenama

4.3. Rezultati trećeg istraživanja

4.3.1. Rezultati pretraživanja

Pretraživanjem baza podataka dobiveno je ukupno 2078 zapisa. Nakon uklanjanja duplikata pregledana su 2073 zapisa. Od toga je isključeno 1186 zapisa jer nisu odgovarali kriterijima uključenja. Popis svih isključenih zapisa s pojedinačno navedenim razlogom isključenja javno je dostupan kao Dodatna datoteka 6 na poveznici <https://osf.io/gj28f/files/osfstorage>. Uključeno je 887 zapisa, odnosno istraživanja koja su koristila AMSTAR 2 za metodološku procjenu SR-ova. Popis uključenih istraživanja dostupan je javno kao Dodatna datoteka 7 na poveznici <https://osf.io/gj28f/files/osfstorage>. Slika 5. prikazuje dijagram tijeka trećeg istraživanja.

Slika 5. Dijagram tijeka trećeg istraživanja



4.3.2. Značajke analiziranih istraživanja

Ekstrakcija podataka provedena je za 50 % nasumično odabranih istraživanja s popisa uključenih istraživanja, dakle za ukupno 444 istraživanja. Najzastupljeniji ustroj istraživanja bio je OSR (N=299; 67 %). Prema godini objave, 133 (30 %) je bilo objavljeno u 2022., a 103 (23 %) u 2023.

godini. Najviše istraživanja (N=3; 3 %) bilo je objavljeno u časopisu *International Journal of Environmental Research and Public Health*. Tablica 8. prikazuje značajke analiziranih istraživanja.

Tablica 8. Značajke analiziranih istraživanja (N = 444)

Varijabla	N (%)
Ustroj istraživanja	
Pregled sustavnih pregleda (engl. <i>overview of systematic reviews</i> , OSR)	299 (67)
Metodološko istraživanje	65 (15)
Sustavni pregled (engl. <i>systematic review</i> , SR)	27 (6)
Brzi sustavni pregled (engl. <i>rapid review</i>)	13 (3)
Pretražni pregled literature (engl. <i>scoping review</i>)	13 (3)
Mapa dokaza (engl. <i>evidence map</i>)	9 (2)
Godina objave	
2018.	2 (<1)
2019.	38 (9)
2020.	75 (17)
2021.	91 (21)
2022.	133 (30)
2023.	105 (24)
Naziv časopisa	
International Journal of Environmental Research and Public Health	13 (3)
Evidence-based Complementary and Alternative Medicine	9 (2)
PLoS ONE	9 (2)
Systematic Reviews	8 (2)
BMJ Open	7 (2)
Nutrients	7 (2)

4.3.3. Pristup izvještavanju i način korištenja alata AMSTAR 2

U analiziranim istraživanjima, jedan je autor proveo procjenu prema AMSTAR-u 2 u 27 (6 %) istraživanja. U 311 (70 %) istraživanja procjene su provela dva, a u 36 (8 %) istraživanja više od dva autora. U 70 (16 %) slučajeva nije bio naveden broj autora koji su proveli procjene. Od ukupno 347 istraživanja u kojima su procjene provela dva ili više od dva autora, u 315 (91 %) je bilo navedeno da su autori procjene proveli neovisno. Od toga je u 42 (13 %) slučajeva bila navedena razina suglasja među procjeniteljima. Tablica 9. prikazuje način izvještavanja o razini suglasja među procjeniteljima.

Tablica 9. Način izvještavanja o razini suglasja među procjeniteljima (N = 42)

Način izvještavanja o razini suglasja među procjeniteljima	N (%)
Kappa koeficijent	17 (40)
Postotak	8 (19)
Intraklasni koeficijent korelacije (ICC)	7 (17)
Tablica	4 (10)
Postotak i Kappa koeficijent	3 (7)
Opisno (npr. „umjerena do visoka razina suglasja“)	2 (5)
Pearsonov koeficijent korelacije	1 (2)

Prosudbe (da, djelomično da, ne) koje su preporučili autori alata AMSTAR 2 bile su korištene u 372 (84 %) istraživanja. U 4 (1 %) istraživanja su uz preporučene prosudbe ujedno korištene i numeričke vrijednosti, dok u 68 (15 %) istraživanja nisu korištene preporučene prosudbe.

Broj istraživanja u kojima su autori koristili ključne domene bio je 365 (82 %). Od toga su u 331 (91 %) istraživanju korištene ključne domene koje su preporučili autori AMSTAR-a 2. U 21 (6 %) istraživanja autori su ključne domene definirali različito od preporučenih. U 13 (3 %) istraživanja

autori su naveli da su koristili ključne domene, ali u tekstu istraživanja nisu definirali koje su domene bile ključne. U 5 (1 %) istraživanja korištene su samo ključne domene, tj. nije korišten cjeloviti AMSTAR 2 alat.

Među 21 istraživanjem u kojima su autori ključne domene definirali različito od preporučenih, medijan broja korištenih ključnih domena bio je 6 (raspon od 2 do 14). Svaka od 16 AMSTAR 2 domena je u barem jednom od istraživanja korištena kao ključna domena. Od standardno preporučenih sedam ključnih domena, u kombinacijama domena različitim od preporučene, 4. domena bila je korištena u 19 (91 %), 9. domena u 17 (81 %), a 13. domena u 13 (62 %) istraživanja. Od standardno ne-ključnih domena, 8. domena je bila je najčešće korištena kao ključna domena i to u ukupno 8 (38 %) istraživanja, dok je 10. domena u samo jednom istraživanju korištena kao ključna. Pet (24 %) istraživanja je ključne domene definiralo na način da su od preporučenih sedam samo izostavili jednu ključnu domenu. Pritom su četiri istraživanja izostavila 2. domenu, dok je jedno istraživanje izostavilo 7. domenu. Tablica 10. prikazuje učestalost korištenja AMSTAR 2 domena kao ključnih u kombinacijama koje su različite od preporučene kombinacije ključnih domena. Definicije ključnih domena u istraživanjima koja nisu koristila preporučenu kombinaciju ključnih domena dostupne su u obliku Dodatne datoteke 9 na poveznici <https://osf.io/gj28f/files/osfstorage>.

Tablica 10. Učestalost korištenja AMSTAR 2 domena kao ključnih u kombinacijama koje su različite od preporučene kombinacije ključnih domena (N = 21)

Domena	Broj istraživanja u kojima je korištena kao ključna domena u kombinaciji različitoj od preporučene, N (%)
Preporučene ključne domene	
2. domena	9 (43 %)
4. domena	19 (91 %)
7. domena	9 (43 %)
9. domena	17 (81 %)
11. domena	11 (52 %)
13. domena	13 (62 %)
15. domena	7 (33 %)
Ostale domene alata AMSTAR 2	
1. domena	7 (33 %)
3. domena	3 (14 %)
5. domena	6 (29 %)
6. domena	5 (24 %)
8. domena	8 (38 %)
10. domena	1 (5 %)
12. domena	2 (10 %)
14. domena	7 (33 %)
16. domena	3 (14 %)

Numerički bodovi bili su dodijeljeni pojedinačnim AMSTAR 2 domenama u 51 (12 %), a ukupan numerički rezultat bio je naveden u 33 (7 %) istraživanja.

Postotak AMSTAR 2 domena koje su autori procijenili s „da“ i/ili „djelomično da“ za svaki SR bio je naveden u 262 (59 %) istraživanja. Postotak usklađenosti po domeni bio je naveden u 299 (67 %) istraživanja.

Konačna ocjena pouzdanosti za svaki uključeni SR bila je navedena u 338 (76 %) istraživanja. Od toga je u 252 (75 %) istraživanja konačna ocjena bila navedena u glavnom dijelu članka, dok je u 86 (25 %) bila navedena u dodatku. Obrazloženje konačne ocjene za svaki SR bilo je navedeno u 101 (30 %) istraživanju, od toga u glavnom dijelu članka u 94 (93 %), a u dodatku u 7 (7 %) istraživanja.

Prosudbe za svaku pojedinu AMSTAR 2 domenu za svaki procjenjivani SR bile su navedene u 321 (72 %) istraživanju. Od toga su iste bile navedene u glavnom dijelu članka u 189 (59 %), a u dodatku u 132 (41 %) istraživanja. Obrazloženje prosudbe svake domene bilo je navedeno u 9 (3 %) istraživanja, i to u 4 (44 %) u glavnom dijelu, a u 5 (55 %) u dodatku istraživanja.

5. RASPRAVA

5.1. Prvo istraživanje

Ovim smo istraživanjem utvrdili da je u člancima objavljenima od 2018. do 2020. god. nešto više od polovice autora koristilo AMSTAR 2 (52 %). Čak 44 % članaka tada je koristilo staru verziju AMSTAR alata, unatoč objavi alata AMSTAR 2 u rujnu 2017. AMSTAR je korišten u više od polovice članaka objavljenih 2018. i 2019. god., no njegova je upotreba pala na 28 % u 2020. godini. Mali broj autora koristio je R-AMSTAR, a neki su čak kombinirali korištenje dvaju alata.

Novi alati na području metodoloških istraživanja kontinuirano se razvijaju i nadograđuju. Autori AMSTAR-a 2 naveli su kako je to poboljšana verzija alata (8). Međutim, postoji mogućnost da autori OSR-ova i metodoloških istraživanja koja procjenjuju kvalitetu SR-ova smatraju da AMSTAR 2 zahtijeva više rada. Naime, AMSTAR ima 11 domena, dok AMSTAR 2 ima 16 domena. Također, korištenje novog alata zahtijeva proučavanje uputa za novi alat.

Nužno je voditi računa da je za pripremu i objavu istraživanja potrebno određeno vrijeme te da je analiza provedena u okviru ove disertacije obuhvatila i članke objavljene manje od četiri mjeseca nakon objave AMSTAR-a 2. Međutim, čak i u slučajevima kad su istraživanja bila u završnoj fazi pripreme i recenzije, autori su mogli odlučiti koristiti AMSTAR 2, odnosno urednici i recenzenti su mogli predložiti da treba koristiti AMSTAR 2. U uzorku iz ovog istraživanja, svega 11 % autora izjavilo je da su urednici ili recenzenti od njih zatražili korištenje AMSTAR-a 2. Jedno istraživanje koje je koristilo i AMSTAR i AMSTAR 2 izvijestilo je da je AMSTAR 2 dodan analizi jer je objavljen neposredno prije predaje rukopisa časopisu (54). Stoga bi se analizirano razdoblje moglo smatrati ranim razdobljem prihvatanja AMSTAR-a 2 nakon objave alata. Za očekivati je da će se upotreba prve verzije AMSTAR-a dodatno smanjivati u budućnosti.

Većina analiziranih istraživanja koristila je reference na članke o razvoju analiziranih alata kako bi potkrijepili njihovu upotrebu. U tom pogledu, uočene su određene nesukladnosti i pogreške. Naime, dio autora naveo je reference na različit alat od onoga koji su zapravo koristili. Neki autori nisu koristili reference na članke koji opisuju razvoj AMSTAR-a, AMSTAR-a 2 i R-AMSTAR-a; već su, umjesto toga, koristili reference na radove drugih autora koji su alat dodatno testirali ili ga jednostavno koristili na drugom uzorku.

Važno je naglasiti da se R-AMSTAR razlikuje od AMSTAR-a i AMSTAR-a 2. Dok je alate AMSTAR i AMSTAR 2 razvila ista istraživačka grupa (7,8), R-AMSTAR su razvili drugi istraživači (9). Kung i sur. razvili su R-AMSTAR s obrazloženjem da AMSTAR nije uključivao

kvantitativne procjene kvalitete sustavnih pregleda. Stoga je, prema autorima, cilj R-AMSTAR-a bio "kvantificirati kvalitetu" SR-ova (9). R-AMSTAR koristi 11 originalnih domena AMSTAR-a, ali svaka domena se boduje s 1 do 4 boda, čime ukupni rezultat može varirati od najmanje 11 do najviše 44 boda. SR s 11 bodova ne zadovoljava niti jedan kriterij AMSTAR-a, dok ocjena od 44 boda označava SR koji zadovoljava sve metodološke kriterije AMSTAR-a u svakoj domeni (9). Valjanost R-AMSTAR-a opisana je u literaturi kao upitna jer je teško procijeniti relativnu važnost pojedinih domena i jesu li sve domene jednako važne pri izračunu konačnog rezultata. Stoga je predloženo da se mjerni atributi R-AMSTAR-a dodatno istraže (11).

Istraživanje provedeno među autorima u okviru ove disertacije pokazalo je da su mnogi od njih znali za AMSTAR-a 2, ali ga nisu koristili jer su već definirali protokol te je prikupljanje podataka bilo u tijeku. Tek nekolicinu su urednici i recenzenti tražili da koriste novu verziju alata. Čak i oni koji nisu znali za objavu novijeg alata, mogli su ga teoretski koristiti da su ih urednici i recenzenti na to potaknuli. Malo je autora navelo prepreke za korištenje AMSTAR-a 2. Stoga se očekuje da će autori nastaviti napuštati početnu verziju AMSTAR-a i koristiti AMSTAR 2 u budućnosti.

Na temelju rezultata ovog istraživanja, vrijedi razmotriti što bi se moglo učiniti kako bi se povećala upotreba novijih, odnosno poboljšanih verzija metodoloških alata kada postanu dostupni. Na primjer, časopisi bi mogli u svojim uputama za autore navesti kako očekuju od autora korištenje novih verzija alata. Nadalje, čak i ako to ne piše u uputama za autore, urednici i recenzenti mogli bi od autora tražiti upotrebu novih verzija alata. Nadalje, edukatori u području metodologije istraživanja i sinteze dokaza trebali bi u svoju edukaciju uključiti nove alate za kritičku procjenu u.

Još jedna točka za poboljšanje je ispravno navođenje referenci na alate. Ovo istraživanje pokazalo je da dio autora navodi pogrešne reference kako bi potkrijepili korištenje odabranog alata. Na primjer, više autora koji su koristili AMSTAR navelo je reference na AMSTAR 2 i obrnuto. Nadalje, neki autori nisu naveli članke koji opisuju razvoj tih alata. Umjesto toga, citirali su druga istraživanja koja su koristila određeni alat. Problem neispravnog referenciranja je nešto čemu bi autori, recenzenti i urednici svakako trebali posvetiti više pozornosti.

Također valja spomenuti poteškoće u kontaktiranju dopisnih autora. Nije bilo moguće kontaktirati 8 % potencijalnih sudionika ankete jer adrese e-pošte nisu bile dostupne u objavljenim člancima ili e-poruka nije bila isporučena na adresu dopisnog autora. Kontakt putem e-pošte danas se smatra standardom. Časopisi bi trebali voditi računa da se navede adresa e-pošte dopisnog autora u objavljenim člancima. Nadalje, problem je prestanak valjanosti e-adresa jer to znači da te autore nije lako kontaktirati u svrhu istraživanja. Već 2006. god. uočeno je da jedna od četiri adrese e-

pošte postaje nevažeca, što negativno utječe na mogućnost istraživača da komuniciraju i razmjenjuju materijale (69).

Istraživačka zajednica kontinuirano analizira AMSTAR i njegovu novu verziju (10,70,71), a ovo je istraživanje još jedan doprinos u tom smjeru. Nadalje, budući da alate obično predlaže grupa autora, a zatim ih objavljuje, korisno je da istraživačka zajednica preispituje i prati predložene metodološke alate. To je važno zbog toga što metodološka istraživanja u konačnici mogu doprinijeti napretku medicinskih istraživanja (72).

Ograničenje ovog istraživanja uključuje korištenje *Google Forms*-a za provođenje ankete. Prednost *Google Forms*-a je što je besplatan za korištenje. Međutim, nekoliko pozvanih autora iz Kine nije moglo koristiti *Google*. Stoga je kombinirana anketa putem e-pošte s anketom putem *Google Forms*-a. Moguće je da bi se neki autori mogli ustručavati odgovoriti putem e-pošte zbog gubitka anonimnosti. Na temelju ovog iskustva, autori koji ciljaju na međunarodne ispitanike za svoje e-ankete trebali bi provjeriti ima li njihova platforma za anketiranje kakve geografske prepreke. Još jedno ograničenje ovog istraživanja je skroman postotak odgovora u anketi. Od autora kojima su poruke isporučene, 27 % je odgovorilo na anketu. Međutim, ta stopa odgovora može se smatrati zadovoljavajućom kad se uzme u obzir da se radilo o online anketi koju su autori primili od osobe koja im ranije nije bila poznata.

Nadalje, u ovom istraživanju analizirana je učestalost korištenja AMSTAR i AMSTAR 2 alata, ali nisu postavljena metodološka pitanja kao što su prednosti ili nedostaci AMSTAR-a 2 u usporedbi s AMSTAR-om. Autori koji jesu koristili AMSTAR 2 nisu anketirani kako bi se ispitalo zašto su koristili novi alat i smatraju li AMSTAR 2 boljim od AMSTAR-a u smislu temeljitosti procjene, jasnoće domena, prisutnosti smjernica za korištenje ili njihovih percepcija o ograničenjima novog alata. To bi mogla biti tema za daljnja istraživanja. Također, novim istraživanjem moglo bi se ispitati razdoblje nakon 2020. god. kako bi procijenilo daljnje prihvatanje AMSTAR-a 2.

5.2. Drugo istraživanje

Ovim je istraživanjem pokazano da je među SR-ovima u kojima je navedeno da su provedeni/pripremljeni/dizajnirani u skladu s AMSTAR-om 2, njih 81 % bilo kritično niske pouzdanosti, a 16 % niske pouzdanosti prema našim procjenama AMSTAR-om 2. Pritom su loše AMSTAR 2 ocjene pretežno bile rezultat neizvještavanja o zahtjevima AMSTAR-a 2 u ključnim domenama.

U doba pisanja ove disertacije nisu pronađena slična istraživanja u literaturi, stoga nije bilo moguće napraviti izravne usporedbe. Međutim, moguće je povući paralelu s istraživanjem Blanca i sur., objavljenom 2018. god., koje se bavilo nesukladnostima između onoga što su autori naveli u CONSORT provjernoj listi (engl. *Consolidated Standards of Reporting Trials*) koju su priložili uz članak i onoga što je zapravo bilo objavljeno u članku. Autori su nasumično odabrali 12 kliničkih pokusa iz tri časopisa koji zahtijevaju da se uz randomizirani kontrolirani pokus dostavi i ispunjena provjerna lista CONSORT. Analizirane su nesukladnosti u šest osnovnih domena CONSORT-a u odjeljcima „Metode“ i „Rezultati“. Od šest članaka u kojima su bile priložene CONSORT provjerne liste, samo je jedan rad bio u potpunosti usklađen s navedenim u priloženom CONSORT-u. Blanco i sur. zaključili su da je potrebno dodatno istražiti metode koje će osigurati dosljedno i precizno ispunjavanje CONSORT-a (73). Taj zaključak trenutno nije primjenjiv na AMSTAR 2. Naime, trenutno nijedan časopis ne zahtijeva da autori ispunjavaju AMSTAR 2 kao kontrolnu listu prilikom predaje rada. Međutim, takav zahtjev mogao bi podići svijest autora o sadržaju AMSTAR-a 2.

Većina SR-ova u ovom istraživanju analizirala je intervencije, ali ne svi. Već je ranije jedno istraživanje ispitalo primjenjivost AMSTAR-a 2 na SR-ove koji ne analiziraju intervencije (50). U tom su istraživanju autori koristili uzorak od četiri vrste SR-ova, uključujući one o točnosti dijagnostičkih testova, etiologiji i/ili rizicima, prevalenciji i/ili incidenciji te prognozi. U tom su istraživanju tri autora neovisno pokušala procijeniti svaki uključeni SR prema AMSTAR-u 2 kako bi se procijenila primjenjivost svake AMSTAR 2 domene na određeni tip SR-a. Rezultati su pokazali da je manje od polovice domena AMSTAR-a 2 bilo primjenjivo na četiri analizirane vrste SR-ova (50). S obzirom na to da nedostaje alat ekvivalentan AMSTAR-u 2 za SR-ove koji ne analiziraju intervencije, daljnji bi razvoj alata trebao uključiti proširenja, odnosno prilagodbe AMSTAR-a 2 za te druge vrste SR-ova.

U ovom istraživanju, dva SR-a koja nisu analizirala intervencije navela su samoprocjene prema AMSTAR-u 2 (62,63), ali samo je istraživanje Al Hammadija i sur. (63) navelo detaljnu procjenu, pri čemu su s „da“ ocijenili sve domene koje su bile primjenjive na njihovo istraživanje. Postavlja

se pitanje jesu li autori SR-ova svjesni da je AMSTAR 2 namijenjen za SR-ove intervencija. Naime, nijedan ne-intervencijski SR u ovom istraživanju nije naveo da koristi AMSTAR 2 zbog nedostatka prikladnijeg alata za tu vrstu istraživanja. Vrijedi spomenuti i kako su u autori dvaju ne-intervencijskih SR-ova upravo korištenje AMSTAR-a istaknuli kao jednu od prednosti svog istraživanja (63,74).

Autori uključenih istraživanja uglavnom su navodili da su se pridržavali alata AMSTAR-a 2 bez daljnjih objašnjenja. Stoga nije poznato jesu li i kako su radili samoprocjene AMSTAR-om 2. Moguće je da nisu proveli formalnu procjenu svojih SR-ova u skladu s AMSTAR 2, već su umjesto toga jednostavno naveli općenitu izjavu o pridržavanju alatu. Nadalje, ako su i proveli formalnu procjenu svojih SR-ova prema AMSTAR-u 2, ostaje nejasno na koji su način to napravili, tj. je li bodovanje provedeno ručno, primjerice, u *Excel* ili *Word* dokumentu, ili putem mrežne stranice alata AMSTAR 2 (38). Među sedam članaka koji su opisali svoje ukupne procjene prema AMSTAR-u 2, tri su članka navela podatke o detaljnim procjenama za svaki pojedini kriterij. Od spomenuta tri članka, dva su koristila AMSTAR 2 mrežnu stranicu. Moguće je da način bodovanja (ručno ili pomoću mrežne stranice) može dati različite ocjene. Nadalje, zbog nedovoljno preciznog izvještavanja, nije poznato jesu li autori uključenih istraživanja koristili originalni algoritam za dobivanje ukupne procjene pouzdanosti, a koji se temelji na sedam ključnih i devet ne-ključnih domena (75). korisnici AMSTAR-a 2 trebali bi detaljno opisati sve modifikacije originalnog algoritma kako bi se omogućila replikacija njihovih metoda za postizanje ukupne procjene pouzdanosti (76).

U ovom istraživanju nije korišteno numeričko bodovanje za procjenu prema AMSTAR-u 2 budući da je na mrežnoj stranici alata jasno navedeno da AMSTAR 2 nije namijenjen za stvaranje jedne ukupne ocjene SR-a (38). Korištenje numeričkog bodovanja, primjerice, s ocjenama poput 1 za pridržavanje, 0,5 za djelomično pridržavanje i 0 za nepridržavanje domena, moglo bi se protumačiti kao da su sve domene jednako važne, što nije nužno slučaj. Stoga je, u skladu s uputama za procjenu prema AMSTAR-u 2, analiza domena u ovom istraživanju bila opisna. Rezultati ovog istraživanja pokazali su da mnogi autori SR-ova ne koriste pravilno AMSTAR 2. Moguće je da je to posljedica nepoznavanja izvornog opisa AMSTAR-a 2 i nepoznavanje ili nerazumijevanje uputa koje su autori alata objavili.

Ranije provedena metodološka istraživanja upućuju na pretjeranu zastupljenost izrazito niskih ocjena pouzdanosti (engl. *floor effect*) prilikom korištenja AMSTAR 2 alata. To je vjerojatno posljedica načina odabira ključnih i ne-ključnih domena u AMSTAR-u 2 (77,78). Nasuprot tome, neki autori tvrde da je velika učestalost procjena koje pokazuju da su SR-ovi loši jednostavno

rezultat loše kvalitete SR-ova (79). Nadalje, dio istraživača kritizirao je pojedine domene AMSTAR-a 2 i način na koji se ocjenjuju (57,80), slično kritikama koje su objavljene o prvoj verziji AMSTAR alata (32,80).

Iako su pojedini autori u radovima analiziranim u ovom istraživanju izvještavali da je njihov SR proveden/pripremljen/osmišljen u skladu s AMSTAR ili AMSTAR 2 alatom, vrijedi istaknuti da 84 % analiziranih SR-ova nije priložilo vlastitu procjenu. Među onima koji su to učinili, samo je manjina objavila detaljnu procjenu. Moguće je da dio autora vjeruje kako samo navođenje tvrdnje da je istraživanje u skladu s AMSTAR-om 2 čitatelju ostavlja dojam ispunjenosti svih AMSTAR 2 domena, odnosno da se na taj način želi istaknuti visoka kvaliteta istraživanja. Ovo je istraživanje pokazalo da navođenje takve tvrdnje najčešće ne znači i visoku metodološku kvalitetu istraživanja, budući da je velika većina analiziranih članaka bila kritično niske pouzdanosti prema AMSTAR-u 2. Do sada je više istraživanja, koja su također koristila AMSTAR 2 za procjenu metodološke kvalitete SR-ova u različitim istraživačkim područjima, pokazala kako je većina analiziranih SR-ova imala kritično nisku ili nisku pouzdanost, što je u skladu s našim rezultatima (15–17).

U svrhu istraživanja korišteno je sedam ključnih domena koje su predložili autori AMSTAR-a 2. Oni su preporučili da korisnici prije početka korištenja alata za procjene definiraju koje su ključne domene za njihovu specifičnu situaciju, ovisno o vrstama SR-ova koje uključuju u svoje istraživanje (8). U ovom su istraživanju korištene domene koje su oni preporučili, kao standardni pristup, jer se istraživanje odnosilo na vrlo raznolike SR-ove. Ipak, to je potencijalno u određenoj mjeri utjecalo na slabe rezultate ukupne procjene prema AMSTAR-u 2..

Nadalje, u istraživanje je uključeno 18 članaka kod kojih iz ni iz cjelovitog teksta, niti iz referenci nije bilo jasno koja je verzija AMSTAR alata korištena. Međutim, prvo istraživanje iz ove disertacije pokazalo je da se upotreba originalne verzije alata brzo smanjuje nakon objave AMSTAR-a 2 (61). Stoga smo smatrali da se, s obzirom na datume objave tih 18 članaka, može pretpostaviti da su se autori referirali na AMSTAR 2. Također, postoji mogućnost pristranosti prema sličnim vrstama intervencija jer je velik broj uključenih članaka bio iz područja kirurgije.

5.3. Treće istraživanje

Ovim je istraživanjem analiziran pristup izvještavanju i način na koji se AMSTAR 2 koristi za metodološku procjenu SR-ova na uzorku od 444 istraživanja koja su bila heterogena s obzirom na ustroj i biomedicinsko područje istraživanja. Nađeno je da su procjene alatom AMSTAR 2 u 78 % uključenih istraživanja radila dva ili više od dva autora. Te procjene su u pravilu (91 %) radila dva ili više autora neovisno. Preporučeni način prosudbi (da, djelomično da, ne) korištenje u 84% istraživanja. Preporučene ključne domene koristilo je 82 % istraživanja. Konačna ocjena pouzdanosti (visoka, umjerena, niska, kritično niska pouzdanost) za svaki uključeni SR bila je navedena u 76 %, a prosudbe za svaku pojedinu domenu (da, djelomično da, ne) u 72 % istraživanja.

Rezultati istraživanja u skladu su s rezultatima istraživanja Karakasisa i sur. iz 2023. čiji je cilj bio procijeniti metode i ishode procjenjivanja pomoću AMSTAR 2 alata u OSR-ovima intervencija u području kardiovaskularne medicine i ispitati čimbenike povezane s tim ishodima (81). Njihovo istraživanje provedeno na uzorku od 34 OSR-a pokazalo je da su procjene prema AMSTAR-u 2 u 68 % slučajeva provodila dva autora, dok je ukupna ocjena pouzdanosti temeljena na sedam ključnih i devet ne-ključnih domena bila je zabilježena u 74 % OSR-ova (81). Slično tome, istraživanje Piepera i sur. iz 2020. izvijestilo je da je 37 od 53 analizirana članka iz područja psihijatrije (70 %) koristilo originalnu kategorizaciju s četiri ocjene ukupne razine pouzdanosti. Od 37 istraživanja koja su koristila originalnu kategorizaciju, samo je 18 izvijestilo kako su došli do ukupne ocjene. Većina njih (72 %) navela je da su koristili sedam ključnih domena predloženih u izvornoj publikaciji (8,76).

U našem uzorku, u 15 % istraživanja nisu korištene preporučene prosudbe (da, djelomično da, ne). Numerički bodovi bili su dodijeljeni pojedinačnim domenama u 12 %, a ukupan numerički rezultat bio je naveden u 7 % istraživanja. U skladu s tim, Karakasis i sur. su 2023. opisali navođenje ocjene izvedene na temelju ukupnog zbroja bodova u 9 % članaka (81). Naprotiv, u istraživanju Piepera i sur. iz 2020. navodi se kako je, među onima koji nisu koristili četiri kategorije pouzdanosti, upravo izračunavanje ukupnog rezultata bila česta alternativa, i to u 38 % članaka (76). Navedena razlika u potonjem istraživanju vjerojatno je posljedica toga što su analizirani članci godinom objave bili bliži objavi AMSTAR-a 2, kada je za očekivati veću učestalost njegova neprikladnog korištenja, što je najvjerojatnije bila posljedica nedovoljnog poznavanja alata (61).

U našem uzorku, ključne domene bile su korištene u 365 istraživanja, od čega se u 91 % radilo o ključnim domenama koje su izvorno definirali autori AMSTAR-a 2. U 6 % slučajeva ključne su

domene bile definirane različito. Autori AMSTAR-a 2 u originalnom članku i popratnom dokumentu s uputama za korištenje alata istaknuli su važnost definiranja ključnih domena, pri čemu je procjeniteljima ostavljeno na izbor hoće li koristiti sedam preporučenih ključnih domena ili će odrediti drugu kombinaciju ključnih domena ovisno o potrebama i specifičnostima svoga istraživanja (8). U pet istraživanja bile su korištene samo ključne domene, odnosno nije bio korišten cijeli AMSTAR 2 alat. Prednost je takvog pristupa kraće vrijeme potrebno za procjenu svakog SR-a. Autori se na to mogu odlučiti u istraživanjima koja uključuju veliki broj SR-ova, ili kad nemaju dovoljno resursa za provođenje opsežnijeg istraživanja. Upravo je jedna od kritika AMSTAR-u 2 u odnosu na stariju inačicu alata bila duže vrijeme potrebno za njegovu primjenu (49). To je opisano i kao moguća prepreka bržem prihvaćanju i široj upotrebi AMSTAR-a 2 u prvim godinama nakon njegove objave (61).

U slučajevima kada su autori odredili kombinaciju ključnih domena različitu od preporučene, među „standardnih“ sedam ključnih domena, najviše su bile zastupljene 4. domena s 91 % i 9. domena s 81 %. Navedeno upućuje da su autori prikladno pretraživanje literature i procjenu RoB-a u istraživanjima uključenima u SR smatrali čimbenicima koji u najvećoj mjeri određuju razinu pouzdanosti SR-a. Potrebno je naglasiti i da su 11. domena (prikladnost metoda u MA) i 13. domena (razmatranje RoB-a prilikom interpretacije rezultata SR-a) bile zastupljene u više od polovice istraživanja. Dakle, postojala je očigledna dominacija domena koje se odnose na RoB. Ovakav izbor ključnih domena podudara se s rezultatima istraživanja iz 2020. koje su objavili Leclercq i sur. (82), a čiji je cilj bio odrediti relativnu važnost pojedine AMSTAR 2 domene u procjeni SR-ova provođenjem tzv. analize najboljeg i najgoreg slučaja (engl. *best-worst scaling*) među uzorkom stručnjaka iz područja sinteze dokaza. Sudionici su u nizu od 15 zadataka biranja odlučivali koji je od četiri ponuđena elementa s glavnog popisa (koji je sadržavao 16 AMSTAR 2 domena) najvažniji i najmanje važan. Prema mišljenju 242 stručnjaka, najvažnije domene bile su 11. domena (primjerenost statističkih analiza) i 4. domena (prikladno pretraživanje literature), a redom su slijedile domene koje se odnose na RoB: 12., 9. i 13. domena (82).

U našem su istraživanju 2. domena (registracija protokola prije provođenja SR-a) i 7. domena (popis isključenih istraživanja s razlozima isključenja) bile zastupljene u tek nešto više od 40 % kombinacija ključnih domena različitih od preporučene. Zanimljivo je da su u ranijim istraživanjima koja su procjenjivala metodološku kvalitetu SR-ova upravo ove dvije domene često bile izdvajane kao one koje su najmanje zadovoljene. Na primjer, u metodološkom istraživanju koje je koristilo AMSTAR 2 za procjenu 679 SR-ova u području kardiovaskularnog zdravlja, 2. i 7. domeni bile su dodijeljene prosudbe „da“ u samo 18,6 % i 12,6 % SR-ova, redom (81). Nadalje,

istraživanje Siemensa i sur. iz 2021. koje je koristilo AMSTAR 2 za metodološku procjenu 261 SR-a vezanog za pacijente s uznapredovalim stadijem raka otkrilo je da je 230 (88 %) SR-ova bilo kritično niske pouzdanosti. Od toga je većina njih (209, 80 %) ovako procijenjena upravo zbog kombinacije nepostojanja prethodne registracije protokola (222, 85 %) s neizvještavanjem o isključenim istraživanjima (218, 84 %) (83).

Nameće se pitanje može li se izborom manjeg broja ključnih domena ili djelomičnim korištenjem alata AMSTAR 2 omogućiti brža i jednostavnija procjena SR-ova bez istodobnog gubitka pouzdanosti i valjanosti alata. Lorenz i sur. proveli su istraživanje kojemu je cilj bio izraditi brz i jednostavan algoritam odlučivanja (engl. *decision tree*) za procjenu metodološke kvalitete SR-ova korištenjem ograničenog broja AMSTAR 2 domena. Na temelju 118 SR-ova koji su se odnosili na psihološke intervencije kod psihijatrijskih poremećaja, a uz pomoć posebnog računalnog programa napravljen je algoritam odlučivanja koji sadrži samo 2. i 7. domenu. Prilikom njegove primjene, 14 visokokvalitetnih SR-ova bilo je ispravno prepoznato, a nijedan nije bio propušten. Istodobno su 94 niskokvalitetna SR-a bila ispravno prepoznata, a 10 je bilo pogrešno prepoznato kao visokokvalitetni. Navedeno je odgovaralo osjetljivosti od 100 % i specifičnosti od 93 % (84). Uz manje vremena koje je potrebno u odnosu na korištenje svih 16 AMSTAR 2 domena, dodatna prednost tog skraćenog pristupa je i mogućnost sudjelovanja manje iskusnih istraživača u procjenama, s obzirom da provjera postojanja unaprijed definiranog protokola istraživanja (2. domena) i navođenja isključenih istraživanja (7. domena) ne zahtijevaju napredno poznavanje metodologije SR-ova, što može biti slučaj s nekim drugim AMSTAR 2 domenama.

Nastavno, De Santis i Matthias su istraživale različite pristupe procjenjivanja uz pomoć AMSTAR-a 2 na primjeru 30 SR-ova vezanih za korištenje digitalnih intervencija za poticanje tjelesne aktivnosti (85). U prvom pristupu koristile su samo dvije ključne domene – 2. i 7. domenu, pri čemu je izbor tih domena bio temeljen na istraživanju Lorenza i suradnika (84), kao i njihovom opažanju da ove domene često nisu bile ispunjene u ranijim istraživanjima. Drugi je pristup bio standardno korištenje svih 16 AMSTAR 2 domena. Prvim je pristupom 19 od 30 SR-ova bilo kritično niske pouzdanosti jer nisu zadovoljili 2. i 7. domenu, a prosječno vrijeme potrebno za procjenu jednog SR-a bilo je 5 minuta. Preostalih 11 SR-ova nisu dobili konačne procjene pouzdanosti jer su zadovoljili ili 2. ili 7. domenu pa je u tom slučaju bilo uputno koristiti cjeloviti AMSTAR 2 alat. Drugi je pristup zahtijevao 20 minuta po SR-u, a rezultirao je s 27 SR-ova kritično niske i 2 SR-a niske pouzdanosti. Autorice su navele da bi se korištenje samo dviju ključnih domena trebalo razmotriti u slučajevima kada je za donošenje neke odluke u zdravstvenom sustavu dostupno

mnogo SR-ova na sličnu temu, a s ciljem brzog identificiranja i isključivanja SR-ova kritično niske pouzdanosti (85).

Lorenz i sur. su naknadnim istraživanjem, koje je koristilo 1519 procjena alatom AMSTAR 2 iz ranije objavljene 24 publikacije, razvili dva pojednostavljena algoritma odlučivanja (engl. *fast and frugal trees*, FFT). Svrha FFT-a za probir bila je isključiti SR-ove kritično niske pouzdanosti korištenjem 2., 7. i 13. domene alata AMSTAR 2. Osjetljivost FFT-a bila je 100 %, a specifičnost 77 %. S druge strane, svrha FFT-a za brzu procjenu je brzo prepoznati SR-ove visoke kvalitete korištenjem 7., 2. i 4. domene, redom. Njegova osjetljivost je 80 %, a specifičnost 97 % (86).

Prema rezultatima ovog istraživanja, konačna ocjena pouzdanosti za svaki uključeni SR bila je navedena u 76 % istraživanja, a od toga je obrazloženje konačne ocjene bilo navedeno u samo 30 %. Prosudbe za svaku pojedinu AMSTAR 2 domenu za svaki procjenjivani SR bile su navedene u 72 %, a od toga su obrazloženja prosudbe svake domene bila navedena u tek 3 % istraživanja. Slično tome, Karakasis i sur. su izvijestili da su u njihovom uzorku procjene pojedinih AMSTAR 2 domena bile dane su u 71 % uključenih SR-ova (81).

Potrebno je poticati korisnike alata AMSTAR 2 da navode obrazloženja svojih procjena kako pojedinačnih domena, tako i ukupnih procjena pouzdanosti jer se na taj način osigurava transparentnost, dosljednost i ponovljivost procjene SR-ova. Ujedno se jača vjerodostojnost rezultata istraživanja, omogućuje se prepoznavanje metodoloških nedostataka i olakšava donošenje informiranih odluka utemeljenih na dokazima. Ograničenja u broju riječi prilikom slanja rada u časopis ne smiju biti prepreka transparentnom izvještavanju zato što uvijek postoji mogućnost objave u obliku dodatnih datoteka ili putem poveznice.

Prednosti ovog istraživanja su smanjen rizik od pristranosti zbog unaprijed registriranog protokola i provođenja ekstrakcije podataka neovisno u duplikatu. S obzirom da u pretraživanju literature nije bilo ograničenja s obzirom na jezik, biomedicinsko područje niti ustroj istraživanja, rezultati se mogu generalizirati, odnosno primijeniti na sva područja sinteze dokaza u medicini.

Iako prema izvornoj preporuci autora AMSTAR-a 2 ovaj alat nije namijenjen za procjenu istraživanja kao što su brzi pregledi (engl. *rapid review*), pretražni SR-ovi (engl. *scoping review*) ili mape dokaza (engl. *evidence map*) (8), u ovom istraživanju navedeni ustroji nisu isključeni zbog toga što svrha istraživanja nije bila procijeniti spomenute vrste istraživanja pomoću AMSTAR-a 2, već utvrditi načine na koji su njihovi autori koristili ovaj alat. Ipak, potrebno je napomenuti da su, očekivano, dvije trećine istraživanja bila OSR-ovi.

Nismo analizirali način na koji su autori došli do ukupne ocjene, odnosno jesu li ukupnu ocjenu dodjeljivali samostalno iz vlastitih tablica ili korištenjem AMSTAR-ove mrežne stranice koja je ranije kritizirana zbog nejasnog načina rada algoritma. Pieper i sur. su pokazali da se korištenjem AMSTAR-ove mrežne stranice mogu dobiti bolje ocjene od realnih, odnosno da se tim načinom generiranja ukupne procjene manjem broju SR-ova dodijeli ukupno niska ili kritično niska pouzdanost (76).

6. ZAKLJUČCI

6.1. Zaključak prvog istraživanja

U člancima objavljenima u razdoblju od 2018. do 2020. godine, a koji su poslani u časopis nakon što je objavljen AMSTAR 2 alat, 44 % autora i dalje je koristilo AMSTAR, odnosno staru verziju alata. Time nije potvrđena hipoteza da će tijekom prve tri godine od objave alata AMSTAR 2, 50% autora i dalje koristiti AMSTAR (prvu inačicu alata) za procjenu SR-ova. Nadalje, utvrđeno je kako se korištenje izvornog alata AMSTAR smanjivalo se iz godine u godinu. Opisano je vrlo malo prepreka za korištenje AMSTAR-a 2, te se stoga očekuje da će korištenje AMSTAR-a uskoro postati opsoletno.

6.2. Zaključak drugog istraživanja

Od ukupno 45 analiziranih istraživanja, 81 % je bilo kritično niske pouzdanosti prema AMSTAR-u 2, čime je potvrđena hipoteza da će među SR-ovima koji su naveli da su koristili AMSTAR 2 za ustroj svojeg istraživanja više od polovice njih biti kritično niske pouzdanosti. Samo tri SR-a su navela detaljne samoprocjene prema AMSTAR-u 2, čime je potvrđena hipoteza da će manje od 10 % istraživanja sadržavati detaljnu AMSTAR 2 procjenu vlastitog SR-a. Kada autori SR-a navedu da je istraživanje provedeno/pripremljeno/osmišljeno u skladu s AMSTAR-om 2, to ne znači nužno da se radi o SR-u visoke pa čak ni umjerene razine pouzdanosti. Iako AMSTAR 2 izvorno nije razvijen kao provjerna lista za procjenu sustavnih pregleda, može biti koristan za tu svrhu. Naime, kad se pravilno primjenjuje, može pomoći autorima da naprave SR-ove visoke razine pouzdanosti. Kako bi povećali metodološku kvalitetu SR-ova, časopisi bi mogli zahtijevati samo-ocjenjivanje prema AMSTAR-u 2 kao dodatak SR-u, koje bi urednici/recenzenti detaljno provjeravali. Autori SR-ova trebali bi pomno proučiti i ispravno primijeniti domene AMSTAR-a 2 kako bi postigli visoku kvalitetu SR-a, posebno ako u svom članku navode da su to učinili. Recenzenti i urednici trebali bi posvetiti veću pozornost metodološkoj kvaliteti SR-ova. Takva bi praksa mogla dovesti do boljih dokaza u znanosti.

6.3. Zaključak trećeg istraživanja

Sukladno uputama o korištenju AMSTAR-a 2 u procjeni SR-ova, autori trebaju jasno definirati ključne domene i navesti način na koji se određuje konačna ocjena pouzdanosti. Potrebno je navesti

ukupne procjene pouzdanosti analiziranih SR-ova, kao i prosudbe za svaku domenu s objašnjenjima istih. Među analiziranim istraživanjima, 12 % je dodijelilo numeričke bodove AMSTAR 2 domenama, čime nije potvrđena hipoteza da će više od polovice istraživanja koristiti bodove prilikom upotrebe AMSTAR-a 2. U 72 % istraživanja bile su navedene procjene za svaku domenu svakog procijenjenog istraživanja. Obrazloženja tih procjena bila su navedena u samo 3 % istraživanja, čime je potvrđena hipoteza da će manje od polovice istraživanja objaviti svoje detaljne AMSTAR 2 procjene za svaku domenu svakog procijenjenog istraživanja. Zbog ograničenja broja riječi u časopisima, navedeno se može napraviti u obliku dodatnih datoteka. Na ovaj se način doprinosi transparentnijem izvještavanju i otvara se mogućnost lakšeg i bržeg ponavljanja ili obnavljanja postojećih istraživanja u području sinteze dokaza.

7. LITERATURA

1. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.4 (updated August 2023). Cochrane, 2023. Available from www.training.cochrane.org/handbook.
2. Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts. Treatments for myocardial infarction. JAMA. 1992 Jul 8;268(2):240–8.
3. Oxman AD, Guyatt GH. The Science of Reviewing Research. Ann N Y Acad Sci. 1993 Dec;703(1):125–34.
4. Chandler J, Hopewell S. Cochrane methods - twenty years experience in developing systematic review methods. Syst Rev. 2013 Dec;2(1):76.
5. Guyatt GH. Users' Guides to the Medical Literature: IX. A Method for Grading Health Care Recommendations. JAMA. 1995 Dec 13;274(22):1800.
6. Sackett DL, Rosenberg WMC, Gray JAM, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. BMJ. 1996 Jan 13;312(7023):71–2.
7. Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol. 2007 Dec;7(1):10.
8. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017 Sep 21;j4008.
9. Kung J. From Systematic Reviews to Clinical Recommendations for Evidence- Based Health Care: Validation of Revised Assessment of Multiple Systematic Reviews (R-AMSTAR) for Grading of Clinical Relevance~!2009-10-24~!2009-10-03~!2010-07-16~! Open Dent J. 2010 Jul 16;4(2):84–91.
10. Dosenovic S, Jelacic Kadic A, Vucic K, Markovina N, Pieper D, Puljak L. Comparison of methodological quality rating of systematic reviews on neuropathic pain using AMSTAR and R-AMSTAR. BMC Med Res Methodol. 2018 Dec;18(1):37.

11. Pieper D, Buechter RB, Li L, Prediger B, Eikermann M. Systematic review found AMSTAR, but not R(evised)-AMSTAR, to have good measurement properties. *J Clin Epidemiol*. 2015 May;68(5):574–83.
12. Bastian H, Glasziou P, Chalmers I. Seventy-Five Trials and Eleven Systematic Reviews a Day: How Will We Ever Keep Up? *PLoS Med*. 2010 Sep 21;7(9):e1000326.
13. Catalá-López F, Aleixandre-Benavent R, Caulley L, Hutton B, Tabarés-Seisdedos R, Moher D, et al. Global mapping of randomised trials related articles published in high-impact-factor medical journals: a cross-sectional analysis. *Trials*. 2020 Dec;21(1):34.
14. Mulrow CD. Systematic Reviews: Rationale for systematic reviews. *BMJ*. 1994 Sep 3;309(6954):597–9.
15. Hammel C, Pandis N, Pieper D, Faggion CM. Methodological assessment of systematic reviews of in-vitro dental studies. *BMC Med Res Methodol*. 2022 Dec;22(1):110.
16. Matterne U, Theurich MA, Pröbstl S, Pieper D, Wang J, Xu A, et al. Quality of systematic reviews on timing of complementary feeding for early childhood allergy prevention. *BMC Med Res Methodol*. 2023 Apr 4;23(1):80.
17. Matthias K, Rissling O, Pieper D, Morche J, Nocon M, Jacobs A, et al. The methodological quality of systematic reviews on the treatment of adult major depression needs improvement according to AMSTAR 2: A cross-sectional study. *Heliyon*. 2020 Sep;6(9):e04776.
18. Moher D. Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Ann Intern Med*. 2009 Aug 18;151(4):264.
19. Stroup DF. Meta-analysis of Observational Studies in EpidemiologyA Proposal for Reporting. *JAMA*. 2000 Apr 19;283(15):2008.
20. Dechartres A, Charles P, Hopewell S, Ravaud P, Altman DG. Reviews assessing the quality or the reporting of randomized controlled trials are increasing over time but raised questions about how quality is assessed. *J Clin Epidemiol*. 2011 Feb;64(2):136–44.
21. Shea B, Dubé C, Moher D. Assessing the Quality of Reports of Systematic Reviews: The QUOROM Statement Compared to Other Tools. In: Egger M, Smith GD, Altman DG, editors.

Systematic Reviews in Health Care [Internet]. 1st ed. Wiley; 2001 [cited 2024 Jun 11]. p. 122–39. Available from: <https://onlinelibrary.wiley.com/doi/10.1002/9780470693926.ch7>

22. Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, Chalmers TC. Meta-Analyses of Randomized Controlled Trials. *N Engl J Med*. 1987 Feb 19;316(8):450–5.
23. Whiting P, Savović J, Higgins JPT, Caldwell DM, Reeves BC, Shea B, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. *J Clin Epidemiol*. 2016 Jan;69:225–34.
24. Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol*. 2003 Dec;3(1):25.
25. Verhagen AP, De Vet HCW, De Bie RA, Kessels AGH, Boers M, Bouter LM, et al. The Delphi List. *J Clin Epidemiol*. 1998 Dec;51(12):1235–41.
26. Terwee CB, Mokkink LB, Knol DL, Ostelo RWJG, Bouter LM, De Vet HCW. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res*. 2012 May;21(4):651–7.
27. Thompson S, Ekelund U, Jebb S, Lindroos AK, Mander A, Sharp S, et al. A proposed method of bias adjustment for meta-analyses of published observational studies. *Int J Epidemiol*. 2011 Jun;40(3):765–77.
28. Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *J Clin Epidemiol*. 2009 Oct;62(10):1013–20.
29. Shea BJ, Bouter LM, Peterson J, Boers M, Andersson N, Ortiz Z, et al. External Validation of a Measurement Tool to Assess Systematic Reviews (AMSTAR). Gagnier J, editor. *PLoS ONE*. 2007 Dec 26;2(12):e1350.
30. Faggion CM. Critical appraisal of AMSTAR: challenges, limitations, and potential solutions from the perspective of an assessor. *BMC Med Res Methodol*. 2015 Dec;15(1):63.

31. Burda BU, Holmer HK, Norris SL. Limitations of A Measurement Tool to Assess Systematic Reviews (AMSTAR) and suggestions for improvement. *Syst Rev*. 2016 Dec;5(1):58, s13643-016-0237-1.
32. Wegewitz U, Weikert B, Fishta A, Jacobs A, Pieper D. Resuming the discussion of AMSTAR: What can (should) be made better? *BMC Med Res Methodol*. 2016 Dec;16(1):111, s12874-016-0183-6.
33. <https://www.crd.york.ac.uk/PROSPERO/>.
34. <https://osf.io/>.
35. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Medica*. 2012;22(3):276–82.
36. Higgins JPT, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011 Oct 18;343(oct18 2):d5928–d5928.
37. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016 Oct 12;i4919.
38. https://amstar.ca/Amstar_Checklist.php.
39. DeAngelis CD, Fontanarosa PB. Impugning the Integrity of Medical Science: The Adverse Effects of Industry Influence. *JAMA*. 2008 Apr 16;299(15):1833.
40. Lexchin J. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ*. 2003 May 29;326(7400):1167–70.
41. Rosenthal R. The file drawer problem and tolerance for null results. *Psychol Bull*. 1979 May;86(3):638–41.
42. Nuesch E, Trelle S, Reichenbach S, Rutjes AWS, Tschannen B, Altman DG, et al. Small study effects in meta-analyses of osteoarthritis trials: meta-epidemiological study. *BMJ*. 2010 Aug 2;341(jul16 1):c3515–c3515.
43. Bero L. What is in a name? Nonfinancial influences on the outcomes of systematic reviews and guidelines. *J Clin Epidemiol*. 2014 Nov;67(11):1239–41.

44. Greenland S. On the bias produced by quality scores in meta-analysis, and a hierarchical view of proposed solutions. *Biostatistics*. 2001 Dec 1;2(4):463–71.
45. Jüni P. The Hazards of Scoring the Quality of Clinical Trials for Meta-analysis. *JAMA*. 1999 Sep 15;282(11):1054.
46. Egger M, Schneider M, Smith GD. Meta-analysis Spurious precision? Meta-analysis of observational studies. *BMJ*. 1998 Jan 10;316(7125):140–4.
47. Page MJ, Shamseer L, Altman DG, Tetzlaff J, Sampson M, Tricco AC, et al. Epidemiology and Reporting Characteristics of Systematic Reviews of Biomedical Research: A Cross-Sectional Study. Low N, editor. *PLOS Med*. 2016 May 24;13(5):e1002028.
48. Banzi R, Cinquini M, Gonzalez-Lorenzo M, Pecoraro V, Capobussi M, Minozzi S. Quality assessment versus risk of bias in systematic reviews: AMSTAR and ROBIS had similar reliability but differed in their construct and applicability. *J Clin Epidemiol*. 2018 Jul;99:24–32.
49. Pieper D, Puljak L, González-Lorenzo M, Minozzi S. Minor differences were found between AMSTAR 2 and ROBIS in the assessment of systematic reviews including both randomized and nonrandomized studies. *J Clin Epidemiol*. 2019 Apr;108:26–33.
50. Puljak L, Bala MM, Mathes T, Poklepovic Pericic T, Wegewitz U, Faggion CM, et al. AMSTAR 2 is only partially applicable to systematic reviews of non-intervention studies: a meta-research study. *J Clin Epidemiol*. 2023 Nov;163:11–20.
51. Solmi M, Radua J, Stubbs B, Ricca V, Moretti D, Busatta D, et al. Risk factors for eating disorders: an umbrella review of published meta-analyses. *Braz J Psychiatry*. 2021 Jun;43(3):314–23.
52. Diaz C, Rezende LFM, Sabag A, Lee DH, Ferrari G, Giovannucci EL, et al. Artificially Sweetened Beverages and Health Outcomes: An Umbrella Review. *Adv Nutr*. 2023 Jul;14(4):710–7.
53. World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. *JAMA*. 2013 Nov 27;310(20):2191.
54. Sharma S, Oremus M. PRISMA and AMSTAR show systematic reviews on health literacy and cancer screening are of good quality. *J Clin Epidemiol*. 2018 Jul;99:123–31.

55. McGuire C, Samargandi OA, Corkum J, Retrouvey H, Bezuhly M. Meta-Analyses in Plastic Surgery: Can We Trust Their Results? *Plast Reconstr Surg*. 2019 Aug;144(2):519–30.
56. Kim HR, Choi CH, Jo E. A Methodological Quality Assessment of Meta-Analysis Studies in Dance Therapy Using AMSTAR and AMSTAR 2. *Healthc Basel Switz*. 2020 Nov 1;8(4):446.
57. De Santis KK, Kaplan I. Assessing the Quality of Systematic Reviews in Healthcare Using AMSTAR and AMSTAR2: A Comparison of Scores on Both Scales. *Z Für Psychol*. 2020 Mar;228(1):36–42.
58. Jeyaraman M, Muthu S, Jain R, Khanna M. Autologous bone marrow derived mesenchymal stem cell therapy for osteonecrosis of femoral head: A systematic overview of overlapping meta-analyses. *J Clin Orthop Trauma*. 2021 Feb;13:134–42.
59. Thomson K, Hillier-Brown F, Todd A, McNamara C, Huijts T, Bambra C. The effects of public health policies on health inequalities in high-income countries: an umbrella review. *BMC Public Health*. 2018 Dec;18(1):869.
60. Brouwers MC, Kho ME, Browman GP, Burgers JS, Cluzeau F, Feder G, et al. AGREE II: advancing guideline development, reporting and evaluation in health care. *Can Med Assoc J*. 2010 Dec 14;182(18):E839–42.
61. Bojcic R, Todoric M, Puljak L. Adopting AMSTAR 2 critical appraisal tool for systematic reviews: speed of the tool uptake and barriers for its adoption. *BMC Med Res Methodol*. 2022 Dec;22(1):104.
62. Imani MM, Azizi F, Bahrami K, Golshah A, Safari-Faramani R. In vitro bleaching effect of hydrogen peroxide with different time of exposition and concentration on shear bond strength of orthodontic brackets to human enamel: A meta-analysis of in vitro studies. *Int Orthod*. 2020 Mar;18(1):22–31.
63. Al Hammadi H, Reilly J. Prevalence of obesity among school-age children and adolescents in the Gulf cooperation council (GCC) states: a systematic review. *BMC Obes*. 2019 Dec;6(1):3.
64. Shu WB, Zhang X bo, Lu H ya, Wang HH, Lan GH. Comparison of effects of four treatment methods for unstable intertrochanteric fractures: A network meta-analysis. *Int J Surg*. 2018 Dec;60:173–81.

65. Hailu S, Mekonen S, Shiferaw A. Prevention and management of postoperative nausea and vomiting after cesarean section: A systematic literature review. *Ann Med Surg* [Internet]. 2022 Mar [cited 2023 Jul 30];75. Available from: <https://journals.lww.com/10.1016/j.amsu.2022.103433>
66. Beyable AA, Bayable SD, Ashebir YG. Pharmacologic and non-pharmacologic labor pain management techniques in a resource-limited setting: A systematic review. *Ann Med Surg* [Internet]. 2022 Feb [cited 2023 Jul 30];74. Available from: <https://journals.lww.com/10.1016/j.amsu.2022.103312>
67. Fernández JÁ, Alconchel F, Gómez B, Martínez J, Ramírez P. Unresectable GIST liver metastases and liver transplantation: A review and theoretical basis for a new indication. *Int J Surg*. 2021 Oct;94:106126.
68. Simegn GD, Bayable SD, Fetene MB. Prevention and management of perioperative hypothermia in adult elective surgical patients: A systematic review. *Ann Med Surg*. 2021 Dec;72:103059.
69. Wren JD, Grissom JE, Conway T. E-mail decay rates among corresponding authors in MEDLINE: The ability to communicate with and request materials from authors is being eroded by the expiration of e-mail addresses. *EMBO Rep*. 2006 Mar;7(4):455–455.
70. Bühn S, Ober P, Mathes T, Wegewitz U, Jacobs A, Pieper D. Measuring test-retest reliability (TRR) of AMSTAR provides moderate to perfect agreement – a contribution to the discussion of the importance of TRR in relation to the psychometric properties of assessment tools. *BMC Med Res Methodol*. 2021 Dec;21(1):51.
71. Pieper D, Koensgen N, Breuing J, Ge L, Wegewitz U. How is AMSTAR applied by authors – a call for better reporting. *BMC Med Res Methodol*. 2018 Dec;18(1):56.
72. Puljak L. Evidence synthesis and methodological research on evidence in medicine—Why it really is research and it really is medicine. *J Evid-Based Med*. 2020 Nov;13(4):253–4.
73. MiRoR network, Blanco D, Biggane AM, Cobo E. Are CONSORT checklists submitted by authors adequately reflecting what information is actually reported in published papers? *Trials*. 2018 Dec;19(1):80.

74. Al-Ziftawi NH, Shafie AA, Mohamed Ibrahim MI. Cost-effectiveness analyses of breast cancer medications use in developing countries: a systematic review. *Expert Rev Pharmacoecon Outcomes Res.* 2021 Jul 4;21(4):655–66.
75. Matthias K, Lorenz R, De Santis K. The application of AMSTAR 2 in overviews of systematic reviews shows a need for improvement. *Eur J Public Health.* 2021 Oct 20;31(Supplement_3):ckab164.647.
76. Pieper D, Lorenz RC, Rombey T, Jacobs A, Rissling O, Freitag S, et al. Authors should clearly report how they derived the overall rating when applying AMSTAR 2—a cross-sectional study. *J Clin Epidemiol.* 2021 Jan;129:97–103.
77. Li L, Asemota I, Liu B, Gomez-Valencia J, Lin L, Arif AW, et al. AMSTAR 2 appraisal of systematic reviews and meta-analyses in the field of heart failure from high-impact journals. *Syst Rev.* 2022 Dec;11(1):147.
78. Leclercq V, Beaudart C, Tirelli E, Bruyère O. Psychometric measurements of AMSTAR 2 in a sample of meta-analyses indexed in PsycINFO. *J Clin Epidemiol.* 2020 Mar;119:144–5.
79. Lorenz RC, Matthias K, Pieper D, Wegewitz U, Morche J, Nocon M, et al. AMSTAR 2 overall confidence rating: lacking discriminating capacity or requirement of high methodological quality? *J Clin Epidemiol.* 2020 Mar;119:142–4.
80. De Santis KK, Pieper D, Lorenz RC, Wegewitz U, Siemens W, Matthias K. User experience of applying AMSTAR 2 to appraise systematic reviews of healthcare interventions: a commentary. *BMC Med Res Methodol.* 2023 Mar 16;23(1):63.
81. Karakasis P, Bougioukas KI, Pamporis K, Fragakis N, Haidich A. Appraisal methods and outcomes of AMSTAR 2 assessments in overviews of systematic reviews of interventions in the cardiovascular field: A methodological study. *Res Synth Methods.* 2024 Mar;15(2):213–26.
82. Leclercq V, Hiligsmann M, Parisi G, Beaudart C, Tirelli E, Bruyère O. Best-worst scaling identified adequate statistical methods and literature search as the most important items of AMSTAR2 (A measurement tool to assess systematic reviews). *J Clin Epidemiol.* 2020 Dec;128:74–82.

83. Siemens W, Schwarzer G, Rohe MS, Buroh S, Meerpohl JJ, Becker G. Methodological quality was critically low in 9/10 systematic reviews in advanced cancer patients—A methodological study. *J Clin Epidemiol*. 2021 Aug;136:84–95.
84. Lorenz R, Jacobs A, Matthias K. Fast and frugal decision tree for the critical appraisal of systematic reviews in situations with limited time periods. *Deutsche Gesellschaft Für Medizinische Informatik, Biometrie Und Epidemiologie*, editor. 65th Annu Meet Ger Assoc Med Inform. 2021 Feb 26;Biometry and Epidemiology:AustroSwiss Region and Polish Region) of the International Biometric Society (IBS).
85. De Santis KK, Matthias K. Different Approaches to Appraising Systematic Reviews of Digital Interventions for Physical Activity Promotion Using AMSTAR 2 Tool: Cross-Sectional Study. *Int J Environ Res Public Health*. 2023 Mar 7;20(6):4689.
86. Lorenz RC, Jenny M, Jacobs A, Matthias K. Fast-and-frugal decision tree for the rapid critical appraisal of systematic reviews. *Res Synth Methods*. 2024 Nov;15(6):1049–59.

8. SAŽETAK

Ciljevi: Istraživanja opisana u ovoj doktorskoj disertaciji imala su za cilj utvrditi učestalost i način korištenja alata AMSTAR 2 (engl. *A MeaSurement Tool to Assess Systematic Reviews*) u procjeni sustavnih pregleda.

Metode: Prvo istraživanje koje je analiziralo učestalost korištenja alata AMSTAR 2 tijekom prve tri godine od njegovog objavljivanja i potencijalne prepreke njegovom prihvaćanju bilo je multimetodološko i sastojalo se od dva dijela. Prvi dio je bilo presječno meta-istraživanje (metodološko istraživanje objavljene literature). Drugi dio sastojao se od anketiranja autora istraživanja koji su koristili AMSTAR umjesto AMSTAR-a 2. Drugo istraživanje koje je analiziralo metodološku kvalitetu i obilježja sustavnih pregleda u kojima je navedeno da su provedeni u skladu s alatom AMSTAR 2 bilo je presječno meta-istraživanje. Treće istraživanje koje je analiziralo pristup izvještavanju i način na koji se AMSTAR 2 koristi za metodološku procjenu sustavnih pregleda bilo je također presječno meta-istraživanje. U svim istraživanjima analizirani su članci u kojima su autori koristili AMSTAR 2 alat za procjenu sustavnih pregleda.

Rezultati: Prvo istraživanje kojim je analiziran 871 članak objavljen u razdoblju od 2018. do 2020. je pokazalo da je 44 % autora i dalje koristilo staru verziju alata AMSTAR. Korištenje alata AMSTAR 2 povećalo se sa 16 % u 2018. na 44 % u 2019., i na 69 % u 2020. U drugom istraživanju analizirano je 45 članaka, od čega su 43 bila sustavni pregledi i dva protokola za sustavne preglede. Više od polovice imalo je meta-analizu. Prema neovisnim AMSTAR 2 procjenama, 35 sustavnih pregleda bilo je kritično niske razine pouzdanosti, 7 sustavnih pregleda niske razine pouzdanosti, a jedan sustavni pregled visoke razine pouzdanosti. Nije bilo sustavnih pregleda umjerene razine pouzdanosti. Treće istraživanje kojim su analizirana 444 članka pokazalo je da su procjene AMSTAR-om 2 u 70 % slučajeva radila dva autora. Numerički bodovi AMSTAR 2 domenama bili su dodijeljeni u 12 % istraživanja. U 82 % istraživanja korištene su ključne domene, od čega u 91 % one preporučene od autora AMSTAR-a 2. Među 21 istraživanjem u kojima su autori kombinaciju ključnih domena definirali različito od preporuke, najviše su bile zastupljene četvrta, deveta i trinaesta domena. Konačna ocjena pouzdanosti za svaki uključeni SR bila je navedena u 76 %, a obrazloženje konačne ocjene u 30 % istraživanja. U 72 % istraživanja bile su navedene procjene za svaku domenu svakog procijenjenog istraživanja, dok su obrazloženja tih procjena bila su navedena u 3 % istraživanja.

Zaključak: Korištenje izvornog alata AMSTAR sve je rjeđe iz godine u godinu dok je istodobno opisano vrlo malo prepreka za korištenje AMSTAR-a 2 pa je stoga za očekivati da će upotreba prve verzije AMSTAR-a uskoro postati opsoletna. Kada autori sustavnog pregleda navedu da je isti

proveden/pripremljen/osmišljen u skladu s AMSTAR-om 2, to ne znači nužno da se radi o sustavnom pregledu visoke pa čak ni umjerene razine pouzdanosti. Iako AMSTAR 2 izvorno nije razvijen kao provjerna lista za izradu sustavnih pregleda, može biti koristan za tu svrhu. Autori bi trebali pomno proučiti i ispravno primijeniti domene AMSTAR-a 2 kako bi postigli visoku kvalitetu sustavnog pregleda, osobito ako u svom članku navode da su to učinili. Recenzenti i urednici trebali bi posvetiti veću pozornost metodološkoj kvaliteti sustavnih pregleda. Takva bi praksa mogla dovesti do boljih dokaza u znanosti. Sukladno uputama o korištenju AMSTAR-a 2 u procjeni SR-ova, autori trebaju jasno definirati ključne domene i navesti način na koji se određuje konačna ocjena pouzdanosti. Potrebno je navesti ukupne procjene pouzdanosti analiziranih SR-ova, kao i prosudbe za svaku domenu s objašnjenjima istih. Zbog ograničenja broja riječi u časopisima, navedeno se može napraviti u obliku dodatnih datoteka. Na taj se način doprinosi transparentnijem izvještavanju i otvara se mogućnost lakšeg i bržeg ponavljanja ili obnavljanja postojećih istraživanja u području sinteze dokaza.

9. SUMMARY

Title: Prevalence and Application of the AMSTAR 2 Tool in the Critical Appraisal of Systematic Reviews

Aim: The studies described in this doctoral dissertation aimed to determine the prevalence and how the AMSTAR 2 (A MeaSurement Tool to Assess Systematic Reviews) tool has been used to appraise systematic reviews critically.

Methods: The first study, which investigated the prevalence of AMSTAR 2 use during the first three years following its publication as well as potential obstacles to its adoption, was designed as a multimethodological study and consisted of two parts. The first part was a cross-sectional meta-research study (a methodological investigation of published literature). The second part consisted of a survey sent to authors of studies who had opted to use AMSTAR rather than AMSTAR 2. The second study had a cross-sectional meta-research design and it analyzed the methodological quality and characteristics of systematic reviews explicitly reported as having been conducted following the AMSTAR 2 tool. The third study was also a cross-sectional meta-research study; it investigated reporting practices and the specific ways in which AMSTAR 2 was applied for methodological appraisal of systematic reviews. All three studies analyzed research articles that used AMSTAR 2 tool for appraising systematic reviews.

Results: The first study, which analyzed 871 articles published between 2018 and 2020, demonstrated that 44 % of authors continued to use the first version of the AMSTAR tool. The adoption of AMSTAR 2 increased from 16 % in 2018 to 44 % in 2019 and to 69 % in 2020. The second study assessed 45 articles, of which 43 were systematic reviews and 2 were systematic review protocols. More than half of these included a meta-analysis. According to the independent AMSTAR 2 appraisals, 35 systematic reviews were rated as being of critically low confidence, 7 as low confidence, and 1 as high confidence. Not a single systematic review was rated with moderate confidence. The third study, which analyzed 444 articles, revealed that in 70 % of cases, AMSTAR 2 appraisals were carried out by two authors. Numerical scores for AMSTAR 2 domains were reported in 12 % of the studies. Key domains were used in 82 % of the studies, and of these, 91 % used the domains recommended by the authors of AMSTAR 2. Among the 21 studies in which the authors defined a combination of key domains that differed from the official recommendations, domains 4, 9, and 13 were most frequently included. The final overall confidence rating for each included systematic review was reported in 76 % of studies, whereas the justification for the final rating was provided in 30 %. Domain-specific ratings for each assessed systematic review were

reported in 72 % of studies, yet the accompanying justifications for those domain ratings were provided in only 3%.

Conclusion: The findings of this doctoral dissertation demonstrate that the use of the original AMSTAR tool has decreased from year to year, whereas only a minority of obstacles to adopting AMSTAR 2 have been described. It is therefore to be expected that the use of AMSTAR will soon become obsolete. Nevertheless, when authors of systematic reviews report that their review was conducted, prepared, or designed following AMSTAR 2, this does not necessarily indicate that the review is of high, nor even of moderate confidence. Although AMSTAR 2 was not originally developed as a checklist for conducting systematic reviews, it may be useful for this purpose. Authors should familiarize themselves thoroughly with AMSTAR 2 and apply its domains appropriately in order to achieve a high-quality systematic review, particularly when stating that their work has been guided by this tool. Reviewers and journal editors should devote increased attention to the methodological quality of systematic reviews. Such practice could contribute to strengthening the scientific evidence base. Following the official guidance for AMSTAR 2 appraisal of systematic reviews, authors should clearly define the key domains employed and specify the procedure used for determining the final overall confidence rating. They should also provide overall ratings of confidence for the included systematic reviews, along with judgments for each individual domain, accompanied by appropriate explanations. Given the word limits imposed by journals, such details may be presented in supplementary files. This approach contributes to more transparent reporting and enables easier and faster replication or updating of existing research in the field of evidence synthesis.

10. PRILOZI

10.1. Prilog 1. Strategija pretraživanja literature

Pretraženi su izvori informacija MEDLINE i Embase putem OVID-a kako bi se dohvatile publikacije koji sadrže riječ AMSTAR. Ista strategija pretraživanja korištena je u sva tri istraživanja:

((AMSTAR) ILI (AMSTAR-2)) ILI (AMSTAR 2)) ILI (R-AMSTAR).

Pretraživanje nije bilo ograničeno na određeni dio članka (npr. samo na sažetak ili metode). Nije bilo ograničenja vrste ili jezika publikacije. Zapisi su izvezeni u računalni program EndNote X5 (Clarivate Analytics, London, UK) gdje su duplikati automatski izbrisani.

Prvo istraživanje

Uključeni su svi znanstveni članci koji su koristili AMSTAR ili AMSTAR 2 za procjenu sustavnih pregleda. Uključena istraživanja objavljena su u razdoblju od 1. siječnja 2018. do 31. prosinca 2020. godine.

Drugo istraživanje

Uključeni su svi članci koji su spominjali bilo gdje u tekstu da su provedeni/pripremljeni/ustrojeni u skladu s AMSTAR-om 2. Za istraživanja objavljena između 1. siječnja 2018. i 31. prosinca 2020. pregledani su zapisi s popisa isključenih članaka iz prvog istraživanja. Za zapise objavljene između 1. siječnja 2021. i 3. svibnja 2022. proveli smo novu pretragu baza podataka.

Treće istraživanje

Za treće istraživanje korištena je ista strategija pretraživanja. Uključeni su svi znanstveni članci koji su koristili AMSTAR 2 za metodološke procjene sustavnih pregleda ili bilo koje druge vrste sinteze dokaza. Pretraživanje literature učinjeno je 5. studenoga 2024. godine.

10.2. Prilog 2. Programski kôd korišten za nasumični odabir istraživanja za analizu s popisa uključenih istraživanja

```
import os
```

```
import pandas as pd
```

```
df = pd.read_excel("./2024_11_20_List of included studies_source.xlsx")
```

```
RANDOM_SAMPLE_SIZE_PERCENTAGE = 0.5
```

```
RANDOM_SAMPLE_SIZE = len(df) * RANDOM_SAMPLE_SIZE_PERCENTAGE + 1
```

```
RANDOM_SEED = 144
```

```
random_sample_df = df.sample(n=int(RANDOM_SAMPLE_SIZE),
```

```
random_state=RANDOM_SEED).sort_values("ORN")
```

```
random_sample_df.to_excel("./2024_11_20_List of included studies_source_random_sample.xlsx",  
index=False)
```

10.3. Prilog 3. Tekst ankete poslane e-poštom u prvom istraživanju

Title of the message:

Your opinion about the uptake of AMSTAR 2 (it will take only few minutes of your time)

Text of the message:

Dear Participant [*insert title and name here*],

We are researchers from Cochrane Croatia, and we are conducting a study about the adoption of AMSTAR 2, and barriers related to AMSTAR 2 adoption. We are contacting you because you have recently published a study in which you used AMSTAR:

[*Name of the study*]

We have several questions for you, which may require only 1-3 minutes of your time to answer. Your responses will be anonymized and analyzed in aggregate form. We would be very grateful for few minutes of your time and your responses to these questions.

Link to the survey:

<https://docs.google.com/forms/d/e/1FAIpQLSezWQ08c0A9ccpCgKMIQwmFZwFqFN689puNrZOM9i5saq95sw/viewform>

More detailed information about this study can be found at the bottom of this message.

Thank you very much for your consideration.

Sincere regards,

Name and surname

Additional information about this study:

Principal investigator: Prof. Livia Puljak, MD, PhD

Contact of the principal investigator: livia.puljak@unicath.hr

Ethics approval: The study was approved by the Ethics Committee of the Catholic University of Croatia.

Informed consent: Based on the protocol approved by the Ethics Committee, entering the survey and responding to the study questions will be considered as a consent to participate.

Intention to publish: Our research team intends to publish anonymized data collected in this study. Your name as a study participant will never be mentioned in any publicly available information about this study.

If you have any questions about this study, please do not hesitate to contact us.

11. ŽIVOTOPIS

Ružica Bojčić, dr. med.

Datum rođenja: 6. listopada 1993.

Obiteljski status: udana, majka dvojice dječaka

E-adresa: ruza1511@gmail.com

ORCID: 0000-0001-9561-6550

Radno iskustvo

- od 11./2022. Specijalizantica ginekologije i opstetricije – Klinički bolnički centar Zagreb, za potrebe Doma zdravlja Zagreb – Centar
- 2./2020. – 9./2022. Doktorica medicine u izvanbolničkoj hitnoj službi - Zavod za hitnu medicinu Karlovačke županije
- 8./2019. – 10./ 2019. Doktorica medicine u izvanbolničkoj hitnoj službi - Zavod za hitnu medicinu Šibensko-kninske županije
- 2018./2019. Doktorica medicine na pripravničkom stažu - Klinički bolnički centar Sestre Milosrdnice

Obrazovanje

- TRIBE – poslijediplomski doktorski studij – Medicinski fakultet u Splitu (2022.-2025.)
- Integrirani preddiplomski i diplomski studij Medicina - Medicinski fakultet u Splitu (2012. – 2018.)

Studentske aktivnosti

- 2013.- 2014., 2017.- 2018. Demonstratorica pri Katedri za anatomiju – Medicinski fakultet u Splitu
- 2017.- 2018. Demonstratorica pri Katedri za kliničke vještine - Medicinski fakultet u Splitu
- 2013. - 2014. Demonstratorica pri Katedri za histologiju i embriologiju - Medicinski fakultet u Splitu

Trajna medicinska edukacija

- Poslijediplomski tečajevi I. kategorije:
 - Nove spoznaje u liječenju raka vrata maternice (Zagreb, 5./2025.);
 - Nove spoznaje u liječenju raka endometrija (Zagreb, 11./2024.);
 - Kontracepcija i planiranje obitelji danas (Zagreb, 9./2024.)
- 5. hrvatski kongres menopauzalne medicine (Zagreb, 10./2024.)
- 14. hrvatski kongres o ginekološkoj endokrinologiji, humanoј reprodukciji i menopauzi (Opatija, 9./2023.)

Objavljeni znanstveni članci

1. Puljak L, **Bojčić R**, Botić MC, Odak Z, Riva N, Gold VR, et al. Trials of interventions for ovarian and testicular germ cell tumors registered in ClinicalTrials.gov: a cross-sectional study. Eur. J. Surg. Oncol., p. 108545, Jul. 2024, doi: 10.1016/j.ejso.2024.108545.
2. Erceg Ivkošić I, Fureš R, **Bojčić R**, Hrgović Z, Fureš D, Ćosić V, et al. Adolescents' Knowledge and Attitudes about Contraception, Sexuality, Sexually Transmitted Diseases, and Pelvic Health. Curr Women Health Rev [Internet]. 2024 Apr 26 [cited 2024 Jun 20];20. Available from: <https://www.eurekaselect.com/229409/article>
3. **Bojčić R**, Todoric M, Puljak L. Most systematic reviews reporting adherence to AMSTAR 2 had critically low methodological quality: a cross-sectional meta-research study. J Clin Epidemiol. 2024;165:111210.
4. Gabelica M, **Bojčić R**, Puljak L. Many researchers were not compliant with their published data sharing statement: a mixed-methods study. J Clin Epidemiol. 2022;150:33–41.
5. **Bojčić R**, Todoric M, Puljak L. Adopting AMSTAR 2 critical appraisal tool for systematic reviews: speed of the tool uptake and barriers for its adoption. BMC Med Res Methodol. 2022;22(1):104.

Nagrade

Dekanova pohvalnica za izvrsne rezultate postignute tijekom studija

Članstva

Hrvatska liječnička komora

Hrvatski liječnički zbor

Hrvatski liječnički sindikat

Hrvatsko katoličko liječničko društvo

Strani jezici

Engleski jezik

Njemački jezik