

**Szegedi Tudományegyetem, Általános Orvostudományi Kar,
Szent-Györgyi Albert Klinikai Központ,
Laboratóriumi Medicina Intézet**

**Módszertani minőség és tartalmi
validitás laboratóriumi szakmai
irányelvekben**

Ph.D. Tézis

Dr. Nagy Éva

**Témavezető:
Prof. Dr. Horváth Andrea Rita
PhD, FRCPath, EurClinChem**

2010

1. RÖVIDÍTÉSEK

AACE: American Association of Clinical Endocrinologists	ICC: Interclass Correlations
ADA: American Diabetes Association	IDF: International Diabetes Federation
AAFP: American Academy of Family Physicians	IFCC C-EBLM: Committee on Evidence-based Laboratory Medicine of International Federation of Clinical Chemistry and Laboratory Medicine
ACP: American College of Physicians	Kaiser P: Kaiser Permanente
ACCP: American College of Chest Physicians	LD: Lactate Dehydrogenase
AGREE: the Appraisal of Guidelines for Research and Evaluation for Europe Instrument	NAC: North America Conference
ALP: Alkaline Phosphatase	NACB: National Academy of Clinical Biochemistry
ANDEM: Agence Nationale pour le Développement de l'Évaluation Médicale	NICE: National Institute for Clinical Excellence
ASCO: American Society of Clinical Oncology	NHMRC: National Health and Medical Research Council
ATS-ERS: American Thoracic Society and European Respiratory Society	NHS: the National Health Service of the United Kingdom
BTS-SCG: British Thoracic Society and Society of Cardiothoracic Surgeons of Great Britain and Ireland	NSCLC: Nem kissejtes tüdőrák
CEA: Carcinoembryonic Antigen	NSE: Neuron-Specific Enolase
CDA: Canadian Diabetes Association	NZG: New-Zealand Guidelines Group
CPG: szakmai irányelv	PRODIGY: The NHS Clinical Knowledge Summaries
D: AGREE kérdőív szempont csoportja	SE: Standard Error
DM: Diabetes Mellitus	SEMDSA: Society for Endocrinology, Metabolism and Diabetes of South Africa.
DNA: DNS	SGOT: Serum Glutamic-Oxaloacetic Transaminase
DS: AGREE kérdőív szempont csoport pontja	SIGN: Scottish Intercollegiate Guidelines Network
EBM: Bizonyítékokon alapuló orvoslás	SOGC: Society of Obstetricians and Gynaecologists of Canada
EGTM: European Group on Tumour Markers	SPLF: Société de Pneumologie de Langue Française
FNCLCC: Fédération Nationale des Centres de Lutte Contre le Cancer	TPA: Tissue Polypeptide Antigen
GGT: Gamma-Glutamyl Transferase	USPSTF: U.S. Preventive Services Task Force
GPAC: Guidelines and Protocols Advisory Committee	WHO: World Health Organisation
I: AGREE kérdőívben szereplő egyes kérdések	

2. BEVEZETÉS

A bizonyítékokon alapuló orvoslás (EBM) a gyakorlatban azt jelenti, hogy a szisztematikus módszerrel megállapított, legjobb kutatási bizonyítékokat integráljuk az egyéni klinikai tapasztalattal, szakértelemmel. Így a betegek ellátásakor felmerülő döntések meghozatalában az elérhető legjobb bizonyítékok kerülhetnek alkalmazásra. Az Európa Tanács Miniszteri Bizottsága 1997-ben a bizonyítékokon alapuló irányelvek beillesztését ajánlotta a minőségjavító rendszerekbe. A szakmai irányelvek (CPGs) szisztematikus fejlesztett ajánlások, amelyek segítik az orvos és a beteg döntését specifikus klinikai körülmények között, a megfelelő egészségügyi ellátás érdekében. Az eddigi tapasztalatok alapján egyre növekszik az érdeklődés a szakmai irányelvek iránt, és számos szervezet rendelkezik már módszertani kézikönyvvel, és fejleszt irányelveket. A laboratóriumi diagnosztika területén összesen csak egy tanulmányt találtunk, ami a bizonyítékokon alapuló irányelvfejlesztés adaptációjával foglalkozik. Ebben a közleményen az IFCC C-EBLM munkacsoportja egy irányelvfejlesztési algoritmust és laboratórium specifikus közlési standardokat határoz meg. Bár a publikált nemzetközi szisztematikus irányelvfejlesztési módszerek jól harmonizáltak, különösen a terápiás ajánlások területén, az irányelvek módszertani minősége széles körben kritizált. A diagnosztikai területek szakmai irányelveinek módszertani minősége és azok hatása a gyakorlatra kevésbé vizsgált terület.

3. CÉLOK

Tanulmányainkban az alábbi kulcsfontosságú kérdésekre kerestük a választ:

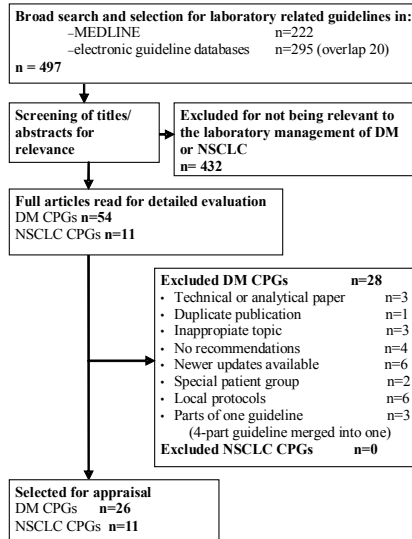
- Van-e olyan könnyen alkalmazható kérdőív, amely a diagnosztikai irányelvek módszertani minőségének értékelésére alkalmas?
- Milyen a laboratóriumi vizsgálatokkal kapcsolatos ajánlások módszertani minősége?
- Különböző-e a diagnosztikai és a terápiás irányelvek módszertani minősége?
- Van-e összefüggés az irányelvek tulajdonságai és módszertani minősége között?
- A diagnosztikai irányelvek megfelelnek-e az alapvető közlési standardoknak?
- Van-e összefüggés az irányelvek módszertani minősége és tartalmi validitása között?

4. MÓDSZEREK

4.1. Téma kiválasztása, szakmai irányelvek keresése és kiválasztása

Vizsgálataink középpontjában olyan két, laboratóriumi vizsgálatok terén is kiemelt jelentőségű terület állt, mint a diabetes mellitus (DM) és a nem kissejtes tüdőrák (NSCLC) diagnosztikájával és a kezelés monitorozásával foglalkozó irányelvek. Szisztematikus irodalmi kutatást végeztünk a PubMed, irányelv adatbázisok és szakmai honlapok felhasználásával. 1999. január 1. – 2007. december 31. között angol nyelven megjelent diagnosztikai irányelveket kerestünk. A kiválasztást két, egymástól független értékelő végezte előre megadott szempontok alapján (1. ábra), aminek eredményeképp 26 DM és 11 NSCLC irányelvet vizsgáltunk a továbbiakban.

1. ábra Irányelvek kiválasztása



4.2. Szakmai irányelvek módszertani minőségének vizsgálata

4.2.1. Kritikus értékelésre alkalmas kérdőív kiválasztása és annak alkalmazhatóságának vizsgálata a diagnosztikai irányelvek esetében

Standardizált, validált és a leggyakrabban használt kérdőívet, az AGREE kérdőívet választottuk az irányelvek módszertani minőségének megítélésére. Annak érdekében, hogy teszteljük az AGREE kérdőív alkalmazhatóságát diagnosztikai irányelvek esetében egy próbatanulmányt végeztünk. A négy leggyakrabban idézett DM irányelvet egymástól függetlenül heten értékelték és az értékelők közötti egyetértést statisztikai módszerekkel elemeztük. Az értékelők pontértékeit összevetettük az interrater korrelációs (ICC) és a Cronbach α együtthatóval szempont csoportonként (domain). A négy irányelvet egyszempontos, illetve ismétléses variancia analízis (one-way ANOVA és ANOVA) segítségével, ismételt méréssel hasonlítottuk össze. A szignifikancia szint $p < 0,05$ volt.

4.2.2. Diabetes mellitus és nem kissejtes tüdőrák irányelvek kritikus értékelésének módszere AGREE kérdőív használatával

Minden irányelvet négy képzett bíráló értékelt az AGREE kérdőív szabályai szerint. 23 szempont teljesülését pontoztunk 1-4 terjedő skálán, majd ezeket 6 szempont csoportba (domain) osztva Domain Score-t (DS) számoltunk minden irányelv esetében. Végül négy lehetséges opció alapján szöveges értékelést (overall assessment) adtunk arra vonatkozólag, hogy ajánlható-e a gyakorlatban az irányelv. Az NSCLC irányelvek esetében a módszertan és a tartalmi validitás könnyebb összehasonlíthatósága miatt az összesítő szöveges értékelés terminológiája (“strongly recommend”, “recommend with provisos or alterations”, “would

not recommend”, “unsure”) helyett a “very good”, “good”, “not so good”, or “dubious” kifejezéseket használtuk.

4.3. Statisztikai módszerek

4.3.1. Szakmai irányelvek tulajdonságai és módszertani minőségük közötti összefüggés vizsgálata

A DM irányelveket azok témája, terjedelme, hossza, eredete és kézikönyvének megléte alapján csoportokba rendeztük. Azt vizsgáltuk, hogy a minőség milyen összefüggésben van az irányelvek kiadásának idejével és kiadás típusával. A statisztikai elemzések során a csoportokat pontjaik (DS) alapján hasonlítottuk össze Kruskal-Wallis próbát alkalmazva. A szignifikancia szint $p \leq 0,01$ volt. Minden statisztikai analízist az SPSS Windows version 13 programmal készítettük. A DM irányelvek esetében azt is vizsgáltuk, hogy az irányelv tartalmazott-e 1) evidenciátáblát, 2) ajánlások erősségének meghatározási módját, valamint 3) meghatározza-e az ajánlások erősségét, 4) megadja-e frissítési vagy a felülvizsgálati időpontot? Az adatokat táblázatban foglaltuk össze és leíró statisztikát (relatív gyakoriság) használtunk az elemzéshez.

4.3.2. Diagnosztikai és kombinált irányelvek közötti különbség vizsgálata

A DM irányelveket a témájuk alapján "tiszán diagnosztikai" és a "kombinált" csoportba rendeztük, aminek DS-jait Mann-Whitney U teszt és Bonferroni korrekcióval hasonlítottuk össze. A szignifikancia szint $p \leq 0,01$ volt. A statisztikai analízist az SPSS for Windows version 13 programmal készítettük.

4.3.3. Diagnosztikai közlési standardoknak való megfelelés vizsgálata

Megvizsgáltuk a jelenlétét a 1) prevalencia, 2) diagnosztikai pontosság, 3) pre-analitikai, és 4) analitikai specifikációknak. Ezen közlési standardok meglétének gyakoriságát összehasonlítottuk a különböző irányelv csoportok között a Fisher-féle egzakt teszttel. A szignifikancia szint $p \leq 0,01$ volt. A statisztikai analízist az SPSS for Windows version 13 programmal végeztük.

4.4. Irodalomkeresés egyéb egészségügyi területen készült irányelvek módszertani értékeléséről

Szisztematikus irodalomkeresést végeztünk elektronikus úton a PubMed-en 2007 májusában, az alábbi kulcsszavak használatával: (("Guideline "[Publication Type] OR "Guidelines as Topic"[Mesh] OR "Guideline Adherence"[Mesh] OR "Practice Guideline "[Publication Type]) AND quality) AND AGREE) nyelvi limit nélkül. A tanulmányokban szereplő irányelvek adatait (téma, eredet, szám, közlési dátum), valamint az AGREE kérdőív használatával kapott DS-okat egy összefoglaló táblázatban és egy diagramban ábrázoltuk.

4.5. Irányelvek módszertani minősége és tartalmi validitása közötti összefüggés vizsgálata

A 11 NSCLC irányelvekből két bíráló kiemelte az összes laboratóriumi vizsgálatokkal kapcsolatos ajánlásokat. Egy meglévő rendszerezett összefoglaló tanulmány alapján az irányelvek ajánlásainak (tumor markerek illetve külön az egyéb tesztek használatára vonatkozó ajánlások) validitását a két bíráló szöveges kifejezésekkel értékelte. A módszertani és a tartalmi validitás eredményeit egy táblázatba foglalva összehasonlítottuk.

5. EREDMÉNYEK

5.1. Az AGREE kérdőív alkalmazhatósága a diagnosztikai irányelvekre

A statisztikai számítások alapján a bírálók közötti egyetértés megfelelő volt (1. táblázat), mivel az egyes bírálók pontjainak összehasonlításakor igen nagy eltéréseket találtunk, ezért úgy döntöttünk, hogy az ICC és a Cronbach's α számítását a későbbiekben nem végezzük el, hanem minden egyes kérdést külön megvizsgálunk és a bírálók konszenzusára törekszünk. Annak ellenére, hogy statisztikai kalkulációval a NICE és az NACB irányelvek között nem találtunk szignifikáns különbséget a módszertani minőség tekintetében, az AGREE által javasolt szöveges értékelés során az NACB rosszabb minősítést kapott a bírálók által. Ezért a bírálók megállapodtak abban, hogy a továbbiakban csak az AGREE kérdőív által ajánlott szöveges értékelést használjuk a statisztikai rangsor képzése helyett. A bírálók egyértelműen megállapították, hogy az AGREE kérdőív alkalmazható a diagnosztikai irányelvek módszertani minőségének megítélésére. Meglehető módon a legismertebb DM irányelvekben is találtunk súlyos módszertani hiányosságokat, ami bizonyítja, hogy szükség van a módszertani minőség vizsgálatára.

1. táblázat Négy DM irányelv módszertani minőségét jellemző DS-ok és ezen alapuló rangsor számítás eredményei, valamint a bírálók közötti egyetértést jelző statisztikai együtthatók.

CPG and date of issue (ref)	Domain Score (%)						Overall assessment	Pair-wise comparison (P value)	Rank
	Scope and purpose	Stakeholder involvement	Rigor of development	Clarity and presentation	Applicability	Editorial independence			
WHO 1999	41	11	8	38	13	24	Not recommend	0,03*	3
NICE 2001	95	94	82	90	36	52	Strongly recommend	-	1
NACB 2002	52	31	43	76	28	19	Recommend with alteration	0,097	1⁺
ADA 2003	70	30	32	73	22	7	Recommend with alteration	0,060**	2
Mean Domain Score (%)	70	39	45	63	28	35			
ICC	0.78	0.91	0.86	0.76	0.02†	0.68			
Cronbach's α	0.78	0.88	0.90	0.77	0.23†	0.63			

* significant ($p \leq 0,05$); ** notable difference (but not significant); †: not acceptable;
⁺ difference is not statistically significant

5.2. Szakmai irányelvek módszertani minősége

5.2.1. Diabetes Mellitus

22 irányelvet tekintettek a bírálók ajánlhatónak, amiből csak 11 volt változtatás nélkül ajánlható, a többi csak módosításokkal, kiegészítésekkel volt elfogadható (2. táblázat). A legjobb pontszámot az AGREE kérdőív első szempont csoportja érte el, ami az irányelvek által megjelölt téma és cél megadásával foglalkozik. Bár az áttekinthetőséget és

dokumentációt vizsgáló 4. szempont csoport is jó pontszámokat ért el, mégis csak az irányelvek 38%-a tartalmazott az irányelv használathoz, bevezetéséhez szükséges segédanyagokat. Az irányelvfejlesztés folyamatát vizsgáló 3. szempont csoport alacsonyabb pontjai a szisztematikus fejlesztési folyamat hiányosságára utalnak, hasonlóan a 2. csoport gyenge eredményeivel, ami jelzi, hogy a fejlesztési munkába sok érintett felet nem vontak be. A legrosszabb eredmények az alkalmazhatóság (5. szempont csoport 34%) és szerzői függetlenség (6. szempont csoport 39%) területén születtek. A szempont csoportokban tapasztalt pontok nagy szórása az irányelvek minősége közötti nagy heterogenitásra utal.

2. táblázat DM irányelvek AGREE kérdőív alapján történt kritikus értékelésének eredményei

CPG and date of issue (ref)	Domain Score (%)						Overall assessment
	Scope and Purpose	Stakeholder involvement	Rigour of development	Clarity and Presentation	Applicability	Editorial indep.	
AAFP 1999 (30)	89	40	69	35	17	21	Recommend with alteration
SIGN 2001 (31)	56	75	74	71	8	71	Strongly recommend
NAC 2002 (32)	47	6	17	33	14	4	Would not recommend
NACB 2002 (33)	53	23	31	67	11	17	Recommend with alteration
NICE BG 2002 (34)	92	85	87	98	33	42	Strongly recommend
NICE L 2002 (35)	92	88	90	98	33	42	Strongly recommend
SEMDSA 2002 (36)	14	23	6	56	0	0	Would not recommend
SOGC 2002 (37)	81	21	40	73	19	13	Recommend with alteration
CDA 2003 (38)	86	33	60	90	25	42	Recommend with alteration
NZG 2003 (39)	86	83	76	96	56	100	Strongly recommend
USPSTF T2 2003 (40)	97	21	77	90	39	88	Strongly recommend
USPSTF GDM 2003 (41)	94	23	74	81	42	83	Strongly recommend
WHO T2 2003 (42)	100	33	29	52	58	42	Recommend with alteration
ACP 2004 (43)	97	27	64	79	6	75	Recommend with alteration
Kaiser P 2004 (44)	42	27	6	65	0	0	Would not recommend
NICE T1 2004 (45)	97	88	92	98	72	92	Strongly recommend
GPAC 2005 (46)	72	35	13	85	69	29	Recommend with alteration
IDF T2 2005 (47)	58	46	55	79	44	96	Recommend with alteration
NHMRC 2005 (48)	97	73	90	81	39	21	Strongly recommend
WHO DG 2006 (49)	78	15	26	69	25	21	Wouldn't recommend
AACE 2007 (50)	64	42	39	69	17	46	Recommend with alteration
ADA 2007 (51)	61	31	39	92	39	0	Recommend with alteration
IDF BG 2007 (52)	86	27	55	60	28	9	Recommend with alteration
PRODIGY L 2007 (53)	97	71	64	90	56	21	Strongly recommend
PRODIGY R 2007 (54)	97	69	67	88	56	21	Strongly recommend
PRODIGY BG 2007 (55)	75	71	67	81	72	29	Strongly recommend
Mean Domain Score (%)	77	45	54	76	34	39	
Range (%)	14-100	6-88	6-92	33-98	0-72	0-100	
No of CPGs with DS more than 60%	20	9	15	22	3	7	
Percentage of CPGs with DS more than 60%	77	35	58	85	11	27	

5.2.2. Nem kissejtes tüdő rák

Csak 5 NSCLC irányelvet értékelték a bírálók elfogadhatónak, amelyek közül egy sem volt módosítások nélkül ajánlható (3. táblázat). Három irányelv egyáltalán nem volt ajánlható a módszertani hiányosságok miatt, másik három módszertani minőségének megítélését pedig bizonytalanak találták a bírálók. A legjobb értékelést az 1. és 4. szempont

csoport kapta, amik az irányelv célját és téma meghatározását, valamint az áttekinthetőséget és dokumentáltságot jellemzi. Jelentős hiányosságokat találtunk a többi szempont csoportban, amik a fejlesztési folyamatot, az érintettek bevonását a fejlesztésbe és a szerzői függetlenséget vizsgálják. A leggyengébb eredményt az alkalmazhatóság (5. szempont csoport) területén találtunk.

3. táblázat NSCLC irányelvek AGREE kérdőív alapján történt kritikus értékelésének eredményei

CPG (ref)	Domain score (%)						Overall assessments
	Scope and purpose	Stakeholder involvement	Rigour of development	Clarity and presentation	Applicability	Editorial dependence	
ACCP (56)	61	46	60	46	6	75	Recommend with alterations
ANDEM (57)	89	25	10	71	0	25	Not recommend
ASCO (58)	94	50	71	67	17	75	Recommend with alterations
ATS-ERS (59)	44	4	5	29	0	8	Unsure
BTS-SCG (60)	100	33	60	79	6	83	Recommend with alterations
CIGNA (61)	67	13	12	54	11	8	Unsure
EGTM (62)	44	4	2	29	0	0	Unsure
FNCLCC (63)	94	54	57	79	17	33	Recommend with alterations
NACB (64)	50	17	29	54	11	25	Not recommend
SIGN (65)	89	75	76	75	33	25	Recommend with alterations
SPLF (66)	61	46	48	38	17	8	Not recommend
Mean Domain Score (%)	72	33	39	56	11	33	
Range (%)	44-100	4-75	2-76	29-79	0-33	0-83	
No of CPGs with DS more than 60%	8	1	4	5	0	3	
Percentage of CPGs with DS more than 60%	73	9	36	45	0	27	

5.3. A diabetes mellitus irányelvek módszertani hiányosságának okai

5.3.1. Az irányelvek jellemzői és módszertani minősége közötti összefüggések

A kiadás időpontja

A legtöbb irányelvet 2002 után, 2 irányelvet pedig 1999 és 2001 között adtak ki. Csak a legmagasabb pontszámokat elérő 1. és 4. szempont csoport mutatott marginális fejlődést az idő előrehaladtával (2. táblázat). A gyenge teljesítményű 6. szempont csoport inkább további romlás mutatott 2005-től, mivel a legtöbb irányelv nem nyilatkozott a szerzői függetlenségekről és az érdekellentétekről.

Közlés típusa

A fejlesztők által megadott közlési típusok igen sokszínűek voltak: 19 esetben nevezték irányelvnek, ajánlásnak, ebből 7 nevezte bizonyítékokon alapulóknak a kiadványt, 4 állásfoglalásnak vagy jelentésnek, és 3 útmutatónak nevezte a kiadott dokumentumot. A 7 irányelvből, ami tényeken alapulóknak sorolta be a dokumentumot, csak 5 tartalmazott evidencia táblát és 6 adta meg az ajánlások erősségét. Három irányelv azonban, amik tartalmaztak evidencia táblázatot, nem nevezte bizonyítékokon alapuló irányelvnek a kiadott dokumentumot. Az irányelvek több mint kétharmada adta meg az ajánlások erősségének meghatározási módját, de csak 16 határozta meg az ajánlásainak erősségét.

Frissítés időpontja, módja

Tizenöt irányelv (58%) adott meg lejárati időt. A leggyakoribb felülvizsgálati időpont 3 és 4 év volt. Csak 10 irányelv (38%) adott megfelelő információkat a frissítési folyamatról.

4. táblázat A DM irányelvek jellemzői

CPG (ref)	Date of issue	Source	Scope	Length (pages)	Guideline manual*	Origin	Type of publication as described by authors	Evidence table	Description of grading system	Graded recommendations	Review date (year)
AAFP (30)	1999	Both	Diagnostic	>100	no	USA	review of the evidence and recommendations	+	-	-	-
SIGN (31)	2001	Database	Combined	51-100	yes	UK	national clinical guidelines	-	+	+	3
NAC (32)	2002	Journal	Combined	1-10	no	North America	consensus report	-	-	-	1
NACB (33)	2002	Both	Diagnostic	11-50	no	USA	guidelines and recommendations	-	+	+	-
NICE BG (34)	2002	Database	Combined	>100	yes	UK	clinical guidelines and evidence review	+	+	+	4
NICE L (35)	2002	Database	Combined	>100	yes	UK	clinical guidelines and evidence review	+	+	+	4
SEMDSA (36)	2002	Database	Combined	1-10	no	South Africa	guideline	-	-	-	-
SOGC (37)	2002	Both	Diagnostic	1-10	no	Canada	clinical practice guidelines	-	+	+	-
CDA (38)	2003	Database	Combined	>100	no	Canada	clinical practice guidelines	-	+	+	-
NZG (39)	2003	Database	Combined	>100	yes	New Zealand	evidence-based best practice guidelines	-	+	+	3
USPSTF T2 (40)	2003	Database	Diagnostic	51-100	yes	USA	recommendation and rationale statement	+	+	+	-
USPSTF GDM (41)	2003	Database	Diagnostic	>100	yes	USA	recommendation and rationale statement	+	+	+	-
WHO T2 (42)	2003	Database	Diagnostic	51-100	yes	International	report	-	-	-	-
ACP (43)	2004	Both	Combined	1-10	yes	USA	clinical practice guidelines	+	-	-	5
Kaiser P (44)	2004	Database	Combined	1-10	no	USA, Canada	guidelines	-	-	-	-
NICE T1 (45)	2004	Database	Combined	>100	yes	UK	clinical guidelines and evidence review	+	+	+	4
GPAC (46)	2005	Database	Diagnostic	11-50	yes	Canada	guidelines and protocols	-	-	-	3
IDF T2 (47)	2005	Database	Combined	51-100	yes	International	global guideline	-	+	-	3-5
NHMRC (48)	2005	Database	Diagnostic	>100	yes	Australia	evidence based guidelines	+	+	+	3
WHO DG (49)	2006	Database	Diagnostic	51-100	yes	International	report	-	-	-	-
AACE (50)	2007	Both	Combined	>100	yes	USA	medical guidelines (evidence based)	-	+	+	-
ADA (51)	2007	Both	Combined	11-50	yes	USA	position statement	-	+	+	1**
IDF BG (52)	2007	Database	Diagnostic	0-50	yes	International	guideline	-	+	+	3
PRODIGY L (53)	2007	Database	Combined	51-100	yes	UK	guidance	-	+	+	Con.
PRODIGY R (54)	2007	Database	Combined	51-100	yes	UK	guidance	-	+	-	Con.
PRODIGY BG (55)	2007	Database	Combined	51-100	yes	UK	guidance	-	+	+	Con.
Percentage of CPGs fulfilling criteria								31	69	62	58

*: Guideline development manual or technical document was available before CPG publication.

** : Information on updating is provided in a separate guideline development manual.; Con.: Continuos

5. táblázat Csoportképzés

Scope and purpose		Domains																		Strongly recommended	Recommended with alteration	Weakly recommended
		Stakeholder involvement			Rigor of development			Clarity and presentation			Applicability			Editorial independence								
		DS (%)	SE	Range	DS (%)	SE	Range	DS (%)	SE	Range	DS (%)	SE	Range	DS (%)	SE	Range	No (%)	No (%)	No (%)			
Source	Guideline database (n=19)	80	5.2	14-100	52	6.1	15-88	58	6.5	6-92	80	3.3	32-98	40	5.1	0-72	45	7.6	0-100	11 (58)	5 (26)	3 (16)
	Journal and GL database (n=7) ^a	70	7.1	47-97	27	4.6	6-42	43	6.8	17-69	64	8.3	33-92	18	3.9	6-39	25	10.0	0-75	0 (0)	6 (86)	1 (14)
		<i>P</i> = 0.209			0.055			0.169			0.083			0.018**			0.152					
Scope	Diagnostic (n=10)	85	4.5	53-100	31	5.2	15-73	50	8.2	13-90	69	5.3	35-90	35	5.8	11-69	34	8.9	9-88	3 (30)	6 (60)	1 (10)
	Combined (n=16)	73	6.2	14-97	54	6.8	6-88	56	6.9	6-92	80	4.5	33-98	33	6.1	0-72	43	8.8	0-100	8 (50)	5 (31)	3 (19)
		<i>P</i> = 0.286			0.023**			0.660			0.097			0.776			0.551					
Length	1-50 pages	61	8.5	14-97	24	2.7	6-35	30	7.0	6-64	68	5.8	33-92	21	7.4	0-69	16	8.0	0-75	0 (0)	6 (33)	3 (33)
	>50 pages (n=17)	86	3.5	56-100	56	6.2	15-88	67	4.8	26-92	80	4.1	35-98	41	4.6	8-72	52	7.2	21-100	11 (65)	5 (29)	1 (6)
		<i>P</i> = 0.009**			0.003*			0.001*			0.051			0.018**			0.001*					
Origin	North America (n=12)	74	5.7	42-97	27	2.8	6-42	44	7.1	6-77	72	5.7	33-92	25	5.5	0-69	35	9.3	0-88	2 (17)	8 (66)	2 (17)
	British (n=7)	87	5.9	56-97	62	3.2	69-88	67	4.5	64-92	82	3.8	71-98	42	8.9	0-72	44	10.1	0-92	7 (100)	0 (0)	0 (0)
	Other (n=7)	74	11.3	14-100	43	9.8	15-83	48	11.2	6-90	70	5.9	52-96	36	7.6	0-58	41	15.4	0-100	2 (28.5)	3 (43)	2 (28.5)
		<i>P</i> = 0.355			0.001*			0.028**			0.037**			0.112			0.606					
Manual	yes (n=19)	84	3.5	56-100	53	5.9	15-88	62	5.3	52-98	82	3.0	50-98	42	4.5	6-72	49	7.4	0-100	11 (58)	7 (37)	1 (5)
	no (n=7)	59	10.5	14-89	25	4.0	6-40	33	9.5	6-69	60	7.7	33-90	12	3.6	0-25	14	5.6	0-42	0 (0)	4 (57)	3 (43)
		<i>P</i> = 0.013**			0.015**			0.022**			0.010*			0.001*			0.004*					

^a One guideline was published in journal only; * $p \leq 0.01$; ** $p \leq 0.05$

Közlés forrása

Egy irányelvet publikáltak csak lektorált folyóiratban, 19-et csak elektronikus adatbázisban, 6 pedig folyóiratban és elektronikus adatbázisban is szerepelt. Az az irányelv, amelyik csak folyóiratban jelent meg, módszertanilag nem volt ajánlható. A mindkét helyen megjelent irányelvek közül egy sem volt változtatások nélkül ajánlható. A csak elektronikus adatbázisban közzétett irányelvek érték el a legjobb végső minősítést (5. táblázat).

Az irányelv hossza

Egyértelmű összefüggés mutatható ki az irányelv hossza és módszertani minősége között (5. táblázat). A legtöbb nem ajánlható irányelv rövidebb volt, mint 50 oldal és minden ajánlott irányelv hosszabb volt, mint 50 oldal. Az AGREE szempont csoportok összehasonlításával jelentős különbségeket találtunk a legtöbb esetben. Meglepő módon az alkalmazhatóság szempontjából legjobb irányelvek hosszabbak voltak 50 oldalnál és elektronikus adatbázisokban voltak elérhetőek.

Irányelvek eredete

Kilenc irányelvet az USA-ban, 3-at Kanadában, 7-et az Egyesült Királyságban, egyet-egy Ausztráliában, Új-Zélandon, Dél-Afrikában fejlesztettek, 4 pedig nemzetközi eredetű. A legtöbb módosítások nélkül ajánlható irányelv az Egyesült Királyságból származott. Szignifikáns különbség ($p \leq 0,01$) volt megfigyelhető 2. szempont csoport esetében a brit irányelvek magasabb pontszáma miatt.

Módszertani útmutató megléte

Az irányelvek kétharmadának volt módszertani útmutatója valamilyen formában. Minden módosítások nélkül ajánlható irányelvnek volt módszertani útmutatója. Az útmutatóval rendelkező irányelvek minden AGREE szempont csoportban jobbnak bizonyultak.

6. táblázat A diagnosztikai és kombinált diabetes mellitus irányelvek AGREE pontjainak összehasonlítása

	AGREE ITEMS	Diagnostic CPG's n=10		Combined CPG's n=16		P
		Mean Score	SE	Mean Score	SE	
Scope and purpose	1 The overall objective of the guideline is specifically described.	3.75	0.2	3.08	0.2	0.023**
	2 The clinical questions covered by the guideline are specifically described.	3.43	0.2	3.00	0.2	0.421
	3 The patients to whom the guideline is meant to apply are specifically described.	3.45	0.2	3.45	0.2	0.897
Stakeholder involvement	4 The guideline development team involves all relevant professional groups.	2.30	0.2	2.67	0.2	0.241
	5 The patients' views and preferences have been sought.	1.58	0.3	2.45	0.3	0.060
	6 The target users of the guideline are clearly defined.	2.63	0.2	3.28	0.2	0.047**
	7 The guideline has been piloted among target users.	1.23	0.2	2.08	0.2	0.220
Rigor of development	8 Systematic methods were used to search for evidence.	2.63	0.2	2.42	0.2	0.551
	9 The criteria for selecting the evidence are clearly described.	2.48	0.3	2.39	0.3	0.660
	10 The methods used for formulating the recommendations are clearly defined.	2.08	0.2	2.44	0.2	0.391
	11 The health benefits, side effects and risks have been considered.	3.28	0.2	3.06	0.2	0.391
	12 There is an explicit link between the recommendations and the supporting evidence.	3.18	0.3	3.08	0.3	0.938
	13 The guideline has been externally reviewed by experts prior to its publication.	2.43	0.2	2.67	0.2	0.856
14 A procedure for updating the guideline is provided.	1.55	0.3	2.78	0.3	0.041**	
Clarity and presentation	15 The recommendations are specific and unambiguous.	3.38	0.1	3.69	0.1	0.363
	16 The different options for management of the condition are clearly presented.	3.23	0.1	3.45	0.1	0.182
	17 The recommendations are easily identifiable.	3.78	0.1	3.73	0.1	0.150
	18 The guideline is supported with tools for application.	1.95	0.3	2.73	0.3	0.220
Applicability	19 The potential barriers in applying the recommendations have been discussed.	1.95	0.2	1.83	0.2	0.452
	20 The potential cost implications of applying the recommendations have been considered.	2.75	0.2	1.88	0.2	0.041**
	21 The guideline presents key review criteria for monitoring and/or audit purposes.	1.43	0.3	2.28	0.3	0.623
Bias and conflicts of interest	22 The guideline is editorially independent from the funding body.	2.60	0.3	2.59	0.3	1.000
	23 Conflicts of interest of guideline development members have been recorded.	1.95	0.3	1.95	0.3	0.979

** $p \leq 0.05$

5.3.2. Diagnosztikai és kombinált szakmai irányelvek módszertani minősége

A terápiás és diagnosztikai ajánlásokat egyaránt tartalmazó irányelvek között nagyobb arányban volt mind a módosítások nélkül elfogadható, mind az egyáltalán nem ajánlható irányelvek aránya a tisztán diagnosztikai irányelvekhez képest (5. táblázat). Az AGREE szempont csoportok tekintetében viszont nem volt szignifikáns különbség ezen két csoport között (6. táblázat).

5.3.3. Diagnosztikai közlési standardoknak való megfelelés

A vizsgált alapvető laboratórium-specifikus információkat, közlési standardokat csak az irányelvek 60%-a említette meg valamilyen formában (7. táblázat). Ezeket az információkat a diagnosztikai irányelvek gyakrabban említették, mint a kombinált irányelvek, de a különbség nem volt statisztikailag szignifikáns a különböző alcsoportok között (8. táblázat).

7. táblázat Laboratórium-specifikus közlési standardok előfordulása a diabetes mellitus irányelvekben

CPG and date of issue (ref)	Prevalence / Pre-test probability	Diagnostic accuracy	Preanalytical information	Analytical information
AAFP 1999 (30)	-	+	-	-
SIGN 2001 (31)	+	+	-	+
NAC 2002 (32)	-	-	-	-
NACB 2002 (33)	+	+	+	+
NICE BG 2002 (34)	-	+	-	+
NICE L 2002 (35)	-	-	-	-
SEMDSA 2002 (36)	-	-	-	-
SOGC 2002 (37)	+	+	+	+
CDA 2003 (38)	+	+	+	+
NZG 2003 (39)	+	+	+	+
USPSTF T2 2003 (40)	+	+	+	+
USPSTF GDM 2003 (41)	+	+	+	+
WHO T2 2003 (42)	+	+	-	-
ACIP 2004 (43)	-	-	-	-
Kaiser P 2004 (44)	-	-	-	-
NICE T1 2004 (45)	-	+	+	+
GPAC 2005 (46)	-	-	+	-
IDF T2 2005 (47)	+	+	+	+
NHMRC 2005 (48)	+	+	+	+
WHO DG 2006 (49)	+	+	+	-
AACE 2007 (50)	+	-	+	-
ADA 2007 (51)	+	+	+	+
IDF BG 2007 (52)	-	-	-	-
PRODIGY L 2007 (53)	-	-	+	+
PRODIGY R 2007 (54)	+	+	+	+
PRODIGY BG 2007 (55)	+	-	+	+
Percentage of CPGs fulfilling criteria	58	58	62	58

8. táblázat Laboratórium-specifikus közlési standardok előfordulása különböző irányelv csoportokban

		Percentage of guidelines fulfilling criteria				
		Prevalence	Diagnostic accuracy	Preanalytical information	Analytical information	
Source	Guideline database (n=19)	58	58	63	63	
	Journal and database (n=7) ^a	57	57	57	43	
		<i>P</i> =	<i>0.973</i>	<i>0.973</i>	<i>0.780</i>	<i>0.407</i>
Scope	Diagnostic (n=10)	70	80	70	50	
	Combined (n=16)	50	44	56	63	
		<i>P</i> =	<i>0.428</i>	<i>0.109</i>	<i>0.683</i>	<i>0.689</i>
Length	0-50 pages (n=9)	33	33	44	33	
	>50 pages (n=17)	71	71	71	71	
		<i>P</i> =	<i>0.103</i>	<i>0.103</i>	<i>0.234</i>	<i>0.103</i>
Manual	Yes (n=19)	63	58	68	63	
	No (n=7)	43	57	43	43	
		<i>P</i> =	<i>0.407</i>	<i>0.973</i>	<i>0.369</i>	<i>0.407</i>
Origin	North America (n=12)	58	58	67	50	
	British (n=7)	43	57	57	86	
	Other (n=7)	71	57	57	43	
		<i>P</i> =	<i>0.556</i>	<i>0.998</i>	<i>0.884</i>	<i>0.205</i>

^a= one guideline published in journal only

9. táblázat Tanulmányok, amelyek az irányelvek módszertanát vizsgálták AGREE kérdőívvel

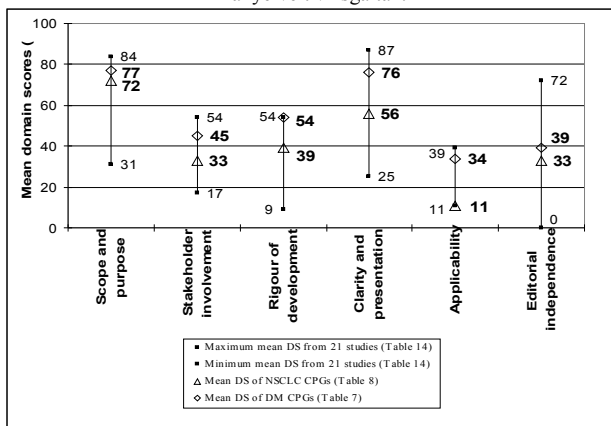
Study (ref)	Topics covered by the CPGs	Publication date of CPGs	Origin of CPGs	No. of CPG	Mean Domain Scores (%)						Comments
					D1	D2	D3	D4	D5	D6	
AGREE 2003 (9)	mixed	1992-1999	international*	33	69	36	41	66	37	30	*CPG origin : 10 from European countries and 1 from Canada
Burgers <i>et al.</i> 2003 (67)	mixed	1992-1999	international*	86	66	34	37	57	31	48	* CPG origin :10 European countries and Canada (62 different agencies and organisations)
Harpole <i>et al.</i> 2003 (17)	lung cancer	1989-2001	international	51	72	35	52	57	20	24	
Brosseau <i>et al.</i> 2004 (68)	musculoskeletal physiotherapy	1998-2002	French or English	9	64	54	49	60	29	24	
Burgers <i>et al.</i> 2004 (18)	non-oncology	1992-1999	international	68	65	30	29	52	30	41	
van Tulder <i>et al.</i> 2004 (19)	acute low back pain in primary care	1987-2001	international*	17	79	50	52	76	28	28	*CPG origin.: 4 USA, 3 Canada, 1 UK, 1 Israel, 2 Netherlands, 1 Germany, 1 Sweden, 1 New Zealand, 1 Finland, 1 Switzerland, 1 Denmark,
Burgers <i>et al.</i> 2004 (18) Fervers <i>et al.</i> 2005 (69)	oncology	1992-1999	international*	32	63	34	42	57	26	47	*CPG origin: 13 countries
Boluyt <i>et al.</i> 2005 (70)	pediatrics	1990-2005	mostly North American *	17	84	42	54	78	19	40	*CPG origin: 13 US A, 3 Canada, 1 Scotland
Horvath <i>et al.</i> 2005 (Paper1)	diagnosis of DM	1990-2003	international*	4	64	41	41	69	25	25	*CPG origin: 2 USA, 1 UK, 1 WHO
Lindberg <i>et al.</i> 2005 (71)	Swedish CPG on diabetes mellitus	not stated	local*	1	78 (73)	30 (67)	14 (37)	61 (72)	31 (60)	72 (60)	* Ostergötland county; percentages in parentheses represent evaluation by lay persons
Lindberg <i>et al.</i> 2005 (71)	Swedish CPG on asthma/allergy	not stated	local*	1	50 (67)	25 (25)	9 (13)	71 (56)	26 (30)	8 (0)	* Ostergötland county; percentages in parentheses represent evaluation by lay persons
Navarro Puerto <i>et al.</i> 2005 (72)	Spanish CPGs	1999-2002	national	61	31	18	18	25	13	38	Domain scores are an approximation of the true mean, reconstructed from the available original data.
Presztoczki <i>et al.</i> 2005 (73)	Hungarian CPGs on management of DM	1993-2004	national	9	78	17	12	54	39	0	Thesis, unpublished
Stiegler <i>et al.</i> 2005 (74)	psychiatric treatment	1998-2003	international*	61	33	31	48	71	23	20	*CPGs origin: 14 European countries
Arnau <i>et al.</i> 2006 (75)	diagnosis and treatment of low back pain	1994-2002	international	17*	63	38	32	53	21	22	*11 guidelines common with the study by van Tulder <i>et al.</i> (19) 2004
Cates <i>et al.</i> 2006 (76)	occupational health	2004	USA	1	80	46	27	87	31	29	
Ministry of Health, Hungary 2006 (77)	Hungarian care pathway protocols	2005-2006	national	180	72	30	28	74	37	8	Internal evaluation for the Ministry of Health; unpublished
Verveij <i>et al.</i> 2006 (78)	suicide attempts	1995-2005*	local in the Netherlands	27	43	22	12	65	15	**	*starting date is not precisely defined; **not used in this study for being considered irrelevant
Watine <i>et al.</i> 2006 (Paper III)	laboratory tests in lung cancer	1997-2003	international*	11	72	33	39	56	11	33	*CPG origin: 5 USA, 3 France, 2 UK, 1 EU
Nagy <i>et al.</i> 2007 (Lecture VII)	diagnosis and monitoring of DM	1999-2005	international*	26	74	41	50	70	27	35	*CPG origin: 13 USA, 3 Canada 6 UK, 1 Australia, 1 New Zealand, 1 South Africa, 1 WHO (unpublished)
TOTAL/MEAN				712	65	34	34	63	26	30	

D1: Scope and purpose; D2: Stakeholder involvement; D3: Rigour of development; D4: Clarity and presentation; D5: Application, D6: Editorial independence.

5.4. Szakmai irányelvek módszertani minősége egyéb egészségügyi területen

21 tanulmányt találtunk 2007-ig, amelyek AGREE kérdőívet használtak az irányelvek módszertani értékeléséhez (9. táblázat). Az összesen 712, különböző orvosi területekről származó, túlnyomórészt terápiás irányelv értékelése történt meg ezekben a tanulmányokban. A tanulmányok maximális és minimális DS értékeit a saját eredményeinkkel vetettük össze (2. ábra). Nagy heterogenitás látható a 21 tanulmányon belül. Eredményeink hasonlóak voltak az egyéb szakterületen megjelent irányelvek eredményeivel. Heterogenitás ellenére hasonló hiányosságok voltak minden orvosi szakterületen. Egyedüli eltérés az NSCLC irányelvek esetében volt, mivel ezek átlagos DS-jai alacsonyabb pontszámokat értek el minden területen, mint a DM irányelvek.

2. ábra Az AGREE kritériumok teljesülése 21 tanulmányban, ahol különböző egészségügyi terület irányelveit vizsgálták.



5.5. Irányelvek módszertani minősége és tartalmi validitása közötti összefüggés

Kigyűjtöttük a laboratóriumi vizsgálatok használatával kapcsolatos ajánlásokat a nem kissejtes tüdőrák irányelvekből (10. táblázat). A közölt szisztematikus irodalmi áttekintés eredményei, amelyek a kezelés előtti laboratóriumi vizsgálatokat elemzi (11. táblázat) nem utalnak arra, hogy a tumor markerek mérése a rutin gyakorlatban javítaná ezen betegség kimenetelét. Ezért csak 4 irányelv tartalmi validitását ítéltük meg „jó”-nak a tumor markerek használatával kapcsolatban. A többi laboratóriumi vizsgálatokat, amelyeket a szisztematikus irodalmi áttekintés hasznosnak ítél, csak 5 irányelv ajánlotta egyértelműen, ami alapján a tartalmi validitást jónak ítéltük. A módszertani minőség és a tartalmi validitás között semmilyen összefüggést nem találtunk (12. táblázat). Eredményeink nem mutattak összefüggést az irányelvek közlési időpontja, a módszertani minősége és tartalmi validitása között.

10. táblázat Kezelés előtti laboratóriumi vizsgálatokra vonatkozó ajánlások az NSCLC irányelvekben

CPG (ref)	Recommended	Unclear recommendation	Not recommended
ACCP (56)	Hematocrit, ALP, calcium, electrolytes, glucose, GGT, SGOT	Other routine laboratory tests	None
ANDEM (57)	Leucocyte count, albumin, SR, calcium, ALP, LD	None	Tumor markers
ASCO (58)	Hemoglobin, leucocyte counts, LD, ALP, calcium	Other routine chemistries, liver function tests	LASA, CA 19-9, DNA index, DNA flow cytometric proliferation analysis, p53 tumor suppressor gene, as oncogene
ATS-ERS (59)	Blood counts, electrolytes, albumin, calcium, ALP, transaminases, bilirubin, creatinine	None	Tumor markers
BTS-SCG (60)	Albumin, creatinine, glucose	None	None
CIGNA (61) *	None	None	CEA, NSE, cyfra 21-1
EGTM (62) *	cyfra 21-1, CEA**	CA 125, TPA	None
FNCLCC (63)	Hemoglobin, leucocyte counts with differential, LD, albumin, calcium	None	Tumor markers
NACB (64) *	None	cyfra 21-1, CEA, NSE	None
SIGN (65)	ALP, calcium	Other biochemistry and hematology tests, liver function tests	None
SPLF (66) *	None	cyfra 21-1	CEA

*: CPGs intended for tumor markers only

**.: Only in cases of adenocarcinoma or large cell carcinoma

11. táblázat Szisztematikus irodalmi áttekintés alapján szükségesnek tartott kezelés előtti laboratóriumi vizsgálatok nem kissejtes tüdő tumoros betegek esetében

Purpose of test	Variables to be measured
Evaluation of toxicity (or tolerance) to treatments	In all patients: hemoglobin, leucocyte counts with differential, platelets, electrolytes, glucose, creatinine, transaminases, bilirubin, albumin
Pretreatment prognostic evaluation	In all patients: hemoglobin (if radiation therapy), leucocyte counts with differential, lactate dehydrogenase, albumin, calcium In patients participating in therapeutic trials: hemoglobin, leukocyte counts with differential, lactate dehydrogenase, albumin, calcium, NSE

12. táblázat Módszertani minőség és tartalmi validitás közötti összefüggés NSCLC irányelvekben

CPGs (ref)	Methodological quality	Validity of content of recommendation	
		Tumor markers	Other laboratory tests
ACCP (56)	Good	Not so good	Good
ANDEM (57)*	Not so good	Good	Good
ASCO (58)	Good	Not so good	Good
ATS-ERS (59)*	Dubious	Good	Good
BTS-SCG (60)*	Good	Not so good	Not so good
CIGNA (61)*	Dubious	Good	
EGTM (62)*	Dubious	Not so good	
FNCLCC (63)	Good	Good	Good
NACB (64)*	Not so good	Not so good	
SIGN (65)	Good	Not so good	Not so good
SPLF (66)*	Not so good	Not so good	

*: diagnostic CPG only

6. MEGBESZÉLÉS

6.1. Az AGREE kérdőív, mint alkalmas eszköz a diagnosztikai szakmai irányelvek módszertani értékelésére

Vizsgálataink során megállapítottuk, hogy az AGREE kérdőív használata könnyen megtanulható, és a laboratóriumi diagnosztikai irányelvek esetében is megfelelően alkalmazható módszertani minőséget értékelő kérdőív. Megfigyeléseink alapján szükségesnek tartjuk, hogy legalább 2, de inkább 3 vagy 4 független bíráló értékelje ugyanazt az irányelvet, akiknek minden kérdésben konszenzusra kell jutni az előítéletek és szubjektivitás torzító hatásának elkerülése miatt.

6.2. Laboratóriumi diagnosztikai irányelvek módszertani minősége

Függetlenül az irányelvek témájától, nagy eltéréseket találtunk a diagnosztikai ajánlások fejlesztésének módjában. Súlyos hiányosságok láthatóak mind a DM és a NSCLC irányelvek módszertani minőségében, mivel a legtöbb irányelvfejlesztő csoport nem használ szisztematikus fejlesztési kritériumokat, nem vonja be a fejlesztési folyamatokba a felhasználókat és a betegcsoportokat. Komoly hiányosságok tapasztalhatóak az irányelvek alkalmazhatósága területén, illetve a szerzői függetlenségről való nyilatkozatok tekintetében.

Figyelmeztető eredmény, hogy számos diagnosztikai irányelv (DM 15%-a és NSCLC 27%-a), köztük a világszerte leginkább elfogadott és hivatkozott irányelvek, mint pl. WHO és NACB; az AGREE kritériumok alapján nem tekinthető felhasználhatónak az alapvető módszertani hiányosságok miatt. A tapasztalt heterogenitás miatt szükségesnek tartjuk minden irányelv kritikus módszertani értékelését, mielőtt az ajánlásait a klinikai gyakorlatba ültetnénk.

Az irodalmi áttekintésünk alapján kimutatható volt, hogy számos más tudományág is hasonló módszertani hiányosságoktól szenved. A viszonylag nagyszámú (n = 712) módszertani értékelés adatai alapján, melynek túlnyomó részét a terápiás ajánlások teszik ki, arra a következtetésre juthatunk, hogy a laboratóriumi diagnosztikai irányelvek nem különböznek a terápiás irányelvektől.

6.3. Módszertani hiányosságok okai a diabetes mellitus diagnosztikai irányelvekben

Néhány korábbi tanulmány már vizsgálta az irányelvek módszertani hiányosságainak lehetséges okait, de csak a diagnosztikai irányelvek esetében még senki nem vizsgálta a módszertani gyengeség okait, valamint a laboratóriumi vizsgálatokkal kapcsolatos speciális információk közlését. Az eredményeink alapján megállapíthattuk, hogy a hosszabb, elektronikus formában az interneten közölt, illetve a fejlesztés módját leíró kézikönyvvel rendelkező irányelvek a legtöbb AGREE kérdőív által vizsgált szempont szerint jobbnak bizonyultak. Ennek egyik kézenfekvő magyarázata, hogy a folyóiratok limitált terjedelme nem alkalmas módszertanilag megkövetelt részletes közlésre. Pedig a kapott eredményeink ellentmondanak annak a szemléletnek, hogy a hosszú irányelvek nem alkalmazhatóak a napi gyakorlatban, mivel a jó módszertani minőségű hosszú irányelvek tartalmaznak rövid, folyamatokat leíró algoritmusokat, oktatási anyagokat, betegtájékoztató anyagokat, szervezési és költség kalkulációra alkalmas adatokat, és olyan indikátorokat, amik a használat ellenőrzésére szolgálhatnak. A „Conference on Guideline Standardization” már meghatározta az irányelvek közlési standardjait épp abból a célból, hogy a javítson a módszertani minőségen és az irányelv implementációját segítse, de a legtöbb folyóirat nem alkalmazza ezeket és a hivatalos bírálók sem értékelik kritikusan a közlésre benyújtott irányelveket AGREE vagy egyéb kérdőívvel.

Az eredményeink alapján nem találtunk szignifikáns különbséget a diagnosztikai és terápiás irányelvek módszertani minősége között. Viszont a diagnosztikai irányelvek közel fele nem közölt preanalitikai, analitikai, a diagnosztikai tesztet jellemző adatokat, amelyek a tesztek megfelelő bevezetéséhez, alkalmazásához, az eredmények interpretálásához elengedhetetlenek. Ezek a hiányosságok abból eredhetnek, hogy az irányelvfejlesztés folyamatában kevés, vagy egyáltalán nem vett részt laboratóriumi szakember, amelyet alátámaszt az az adatunk is, miszerint a 2. „domain” –ben, ahol az összes érintett bevonását vizsgáltuk az irányelvfejlesztési folyamatba, csak 42%-os megfelelést találtunk.

Minden olyan irányelv magasabb pontszámokat ért el, amelyet olyan szervezet fejlesztett, amelynek a fejlesztést részletesen leíró kézikönyve volt. A kézikönyv, azonban nem mindig garantálja a módszertani minőséget, mivel több esetben tapasztalhattuk, hogy a kézikönyvek nem elég részletesek, nem az adott irányelvfejlesztést írják le, illetve a munkacsoportok nem következetesen használják ezeket a standardokat, így gyakran nem egyértelmű, hogyan az adott ajánlás esetében hogyan születtek a döntések. Találtunk, például több olyan irányelvet, amelyekben definiálták, hogyan történt az ajánlások erősségének meghatározása, de magában az irányelvben az ajánlások mellett nem volt feltüntetve az elvileg meghatározott erősségi fokozat. Az evidencia táblázat hiánya is megkérdőjelezi, hogy a kézikönyvben leírt bizonyítékokon alapuló irányelvfejlesztés folyamatát követték volna.

A rendkívül változatos elnevezése az általunk vizsgált, ajánlásokat tartalmazó dokumentumoknak és az a példa, hogy a bizonyítékokon alapuló irányelvfejlesztésnek teljesen megfelelő dokumentumot a szerzők nem nevezték irányelvnek, azt tükrözi, hogy nem egységes módon készültek ezek az ajánlások, és mind a fejlesztők, mind a felhasználók számára nem egyértelműek ezek a szakkifejezések.

Annak ellenére, hogy az irányelvfejlesztési módszerek fogalma sokat tisztult és harmonizálódott az elmúlt évtizedben, nem tudtunk kimutatni jelentős javulást a legtöbb értékelési szempont alapján, sőt a szerzői függetlenség kinyilvánítása szerint romlás volt megfigyelhető az évek múlásával.

Tanulmányunknak volt néhány limitációja. Értékelésünket csak angol kiadványokra szűkítettük. Ez a nyelvi elfoglaltság azonban nem okozhatott számottevő torzulást a következtetéseinkben, mivel számos publikáció, köztük a saját irodalmi áttekintésünk is,

megerősítette, hogy nincs lényeges különbség az angol, illetve nem angol nyelvű irányelv publikációk minőségében. Másrészt, mivel a legtöbb nemzeti DM irányelvek a nemzetközi ajánlásokon alapulnak, amelyek angol nyelven jelentek meg. Az AGREE kérdőívnek és az összes kritikus értékelési eszköznek hiányossága, hogy nem tesz különbséget aközött, hogy a gyenge minőség mögött közlési hiányosság van, vagy ténylegesen rossz módszertani folyamat. Ezért eredményeinket nem úgy kell értelmezni, mint a tudományos adatok, ajánlások valódiságának, érvényességének megkérdőjelezését.

6.4. Módszertani minőség és a tartalmi validitás közötti összefüggés a nem kissejtes tüdő rák irányelvek esetében

Számos tanulmány megerősíti azt a feltételezést, hogy a rossz módszertani minőségű irányelvek esetlegesen torzíthatják az ajánlásokat, és ezáltal szükségtelen terhetek és költségeket okozhat a betegek és a társadalom számára. Más tanulmányok azonban kimutatták, hogy az ajánlások kialakításának módszerében látható különbségek, illetve a hivatkozott bizonyítékok különbözősége ellenére, az ajánlások tartalma egyforma volt. Eredményeink pedig azt mutatták, hogy a rossz minőségű irányelv nem feltétlenül vezet rossz ajánláshoz, és fordítva; kiváló minőségű irányelv sem feltétlenül a legjobb ajánlásokat tartalmazza.

A módszertani minősége és az ajánlás validitása közötti ellentétet az magyarázhatja, hogy a fejlesztők különböző bizonyítékokat használtak fel, vagy eltérő módon hoztak döntéseket az ajánlás kialakításakor. Ennek oka lehet, hogy: nem szisztematikusan az irodalom felkutatása; figyelmen kívül hagyják azokat a bizonyítékokat, amik nem egyeznek a saját tapasztalatokkal, vagy feltételezésekkel; anyagi érdekeket helyez előtérbe; figyelembe vesz egyéb befolyásoló tényezőket, mint a költségek, a szervezeti akadályok, a betegek preferenciáit, az etika, vagy a betegbiztonság. El kell ismernünk, hogy a bizonyíték csak egy elem az ajánlások kialakításának folyamatában. Az irányelvfejlesztők le- vagy felfelé pontozhatják az ajánlás erősségét, ha más ok (pl. a társadalmi, gazdasági, szervezeti, etikai, a beteg kilátásai, biztonsági vagy jogi) erősen indokoltá teszi. Ugyanakkor ennek a folyamatnak jól dokumentálnak és átláthatónak kell lennie, hogy az irányelv felhasználók megértsék, hogy milyen logikai érvelés alapján születtek meg a végső ajánlások az adott kutatási eredményekből.

Másrészt az ajánlások alapját adó tudományos bizonyítékok minősége is okozhatja a módszertan és tartalom közötti ellentétet. Számos vizsgálat igazolta, hogy kevés a jó minőségű bizonyíték számos egészségügyi területen, különösen az onkológia területén.

A tanulmányunk korlátai közé tartozik, hogy az onkológiai terület csak egy kis szeletéből vettünk mintát, ezért az adatainkból nem lehet általános következtetést levonni. Sem az AGREE kérdőív, sem más módszertani kritikus értékelési eszköz nem vizsgálja az ajánlások tartalmának validitását, és azok hatása a betegség kimenetelére. Ennek ellenére az eredményeink alapján talált ellentmondások a minőség és a tartalom között felhívja a figyelmet, hogy az ajánlások használata előtt nem csak módszertani kritikus értékelés szükséges, hanem a tartalmi validitás vizsgálatára is. Egymásnak ellentmondó ajánlások a laboratóriumi vizsgálatok használatára vonatkozóan pénzkidobást jelent és kárt okozhat a betegeknek is. A hatékony kezelés attól függ, hogy a hatékonyan használjuk-e a diagnosztikai tesztek, és ha a diagnosztikai ajánlások nem tényeken alapulnak, indokolatlan terápiás beavatkozást indikálhat, illetve nem megfelelő terápiás monitorozást okozhat.

7. ÖSSZEGLZÉS

Tanulmányaink alapján az alábbi megállapításokat tehetjük:

- Jelentős különbségek vannak a diagnosztikai irányelvek módszertani minősége és fejlesztési módjuk között.
- A diabetes mellitusra és a nem kissejtes tüdőrákra vonatkozó irányelvek hasonló módszertani problémával rendelkeznek, mint egyéb egészségügyi területek irányelvei.
- Jelentős módszertani hiányosságok láthatóak az irányelvfejlesztés folyamatában, az érintettek bevonásában, az alkalmazhatóság és a szerzői függetlenségről szóló nyilatkozatok területén, amelyek az irányelvek belső és külső validitását is meghatározzák.
- A csak diagnosztikai ajánlásokat tartalmazó irányelvek nem különböztek szignifikánsan a kombinált irányelvektől.
- A diabetes mellitus irányelvek esetében a hosszabb, az interneten közölt és a fejlesztés módját leíró kézikönyvvel rendelkezők jobb módszertani minőségűek, az AGREE kérdőív alapján magasabb pontokat értek el.
- A diabetes mellitus irányelvek közel fele nem közöl pre-analitikai, analitikai és a diagnosztikai tesztek robusztusságára vonatkozó adatokat, amely a tesztek megfelelő alkalmazásához, illetve az eredmények interpretációjához feltétlenül szükségesek lennének.
- A tumor markerekre vonatkozó ajánlások ellentmondóak a nem kissejtes tüdőrák ellátásáról szóló irányelvekben.
- Nem találtunk egyértelmű összefüggést a módszertani minőség és az ajánlások tartalmának validitása között a nem kissejtes tüdőrák irányelvekben.
- Az eredményeinkből kiderül, hogy minden irányelv esetében szükség van a módszertani minőség és a tartalmi validitás kritikus értékelésére, mielőtt a klinikai gyakorlatba bevezetnénk azok ajánlásait.

A következő ajánlásokat adjuk a jövőre vonatkozólag:

- A laboratóriumi diagnosztika területén is szisztematikus, explicit módon fejlesztett bizonyítékokon alapuló irányelvekre van szükség, amit a közlési standardoknak megfelelően publikálnak.
- A laboratóriumi diagnosztika területén is az irányelvfejlesztés bizonyítékokon alapuló módszertanának hatásosabb terjesztésére van szükség.
- A bizonyítékok értékelésére és az ajánlások kialakításakor szoros együttműködésre van szükség a laboratóriumi szakemberekkel, és a folyamatot explicit, átlátható módon dokumentálni kell.
- Szükség van az irányelvfejlesztés terminológiájának egyszerűsítésére és standardizációjára.
- Egy egységes rendszer az ajánlások erősségének megítélésére javíthatja az ajánlások

tartalmának validitását.

- További, részletesebb tanulmányok szükségesek, amik a módszertan minősége és az ajánlások tudományos validitása közötti összefüggést vizsgálja.
- Minden felhasználónak kritikusan értékelni kell az alkalmazásra kiválasztott irányelv módszertanát és tartalmát is.

8. KÖSZÖNETNYÍLVÁNÍTÁS

Hálásan köszönöm témavezetőmnek, **Prof. Dr. Horváth Andrea Ritának**, hogy lehetőséget biztosított az IFCC C-EBLM munkacsoport kutatásához való csatlakozáshoz és egyéb nemzetközi és országos rendezvényen való részvételhez, és biztosította a Laboratóriumi Medicina (korábban Klinikai Kémia) Intézetben belül a munkakörülményeket.

Köszönettel tartozom a cikkeink társszerzőinek; **Dr. Jean-Christophe Charet, Dr. Bruno Friedberg, Dr. Rita Ónody, Dr. Wytze Oosterhuis, Dr. Dunja Rogic, Dr. Swere Sandberg**; a kitaró és segítőkész munkájukért, különösen **Dr. Joseph Watine** és **Dr. Peter S Bunting**-nek, akiktől ezen idő alatt sok szakmai tanácsot kaptam.

Külön köszönöm **Dr. Boda Krisztinának** és **Dr. Nyári Tibornak** a statisztikai feladatokban nyújtott szakmai segítséget, valamint **Dr. Kís Erikának** az eredményes együttműködést más bizonyítékokon alapuló medicina területeken.

Végül, de nem utolsó sorban, hálás vagyok **szüleimnek**, akik lehetővé tették egyetemi és az utáni tanulmányaimat, és az **egész családomnak**, akik sok szeretettel és toleranciával támogattak ebben a nehéz időszakban.

PUBLIKÁCIÓS LISTA

Ph.D. alapjául szolgáló közlemények

Cikkek

- I. Horvath AR, **Nagy E**, Watine J. Quality of Guidelines for the Laboratory Management of Diabetes Mellitus Scand J Clin Lab Invest 2005;65 (Suppl 240):41-50 (IF:1.235)
- II. Watine J, Friedberg B, **Nagy E**, Onody R, Oosterhuis W, Bunting P, Charet J-C, Horvath AR. Conflict between guideline methodological quality and recommendation validity: a potential problem for practitioners Clin Chem 2006;1(52): 65-72 (IF:4.803)
- III. **Nagy E**, Watine J, Bunting PS, Onody R, Oosterhuis WP, Rogic D, Sandberg S, Boda K, and Horvath AR. Do Guidelines for the Diagnosis and Monitoring of Diabetes Mellitus Fulfill the Criteria of Evidence-Based Guideline Development? Clin Chem 2008;54(11):1872-1882 (IF:4.803)

Cumulative impact factor (ISI JCR 2007): 10,841

Könyv fejezet

- I. Horvath AR, **Nagy E**, Watine J. Critical appraisal of guidelines. In: Evidence-based Laboratory Medicine: From Principles to Practice. Price CP, Christenson RH (eds) AACC Press, Washington. 2nd ed. 2007. 295-319 pp.

Idézhető absztraktok

- I. **Nagy É**, Ónody R, Nyári T, Horváth A. Bizonyítékokon alapul-e a diabetes mellitus diagnosztikája? Klinikai és Kísérletes Laboratóriumi Medicina 2004, 31 (1): 26.
- II. **Nagy E**, Onody R, Horvath AR. Are diabetes mellitus guidelines evidence-based? Clinica Chimica Acta 2005; 355, S340 (IF:1.633)
- III. Horvath AR, **Nagy E**, Watine J, Bunting P, Onody R, Oosterhuis W, Rogic D, Sandberg S, Boda K. Quality of Guidelines for Diabetes Mellitus: Can We Do Better? Clinical Chemistry and Laboratory Medicine 2007; 45, Special Supplement, ppS1-s473 (IF:1.725)

Előadások

- I. **Nagy É**, Ónody R, Nyári T, Horváth A. Bizonyítékokon alapul-e a diabetes mellitus diagnosztikája? Magyar Laboratóriumi Diagnosztikai Társaság 52. kongresszusa 2004.09.04. Sopron
- II. Horvath A R, **Nagy E**. Quality of Guidelines for the Diagnosis of Diabetes Mellitus 10th Bergmeyer Conference: Diabetes Mellitus and Cardiovascular Disease. 14-16 March 2005, Grainau Garmisch-Partenkirchen, Germany
- III. Horvath AR, **Nagy E**. Quality of Guidelines in Laboratory Medicine XIXth International Congress of Clinical Chemistry, 24-28 July 2005, Orlando, US
- IV. Horvath AR, **Nagy E**. Quality of Guidelines for the Laboratory Management of Diabetes Mellitus. 3rd G-I-N Conference, 5-7 December 2005, Lyon, France
- V. **Nagy É**, Ónody R, Horváth A Megbízható bizonyítékokon alapulnak-e a nemzetközi diabetes mellitus irányelvek? Magyar Laboratóriumi Diagnosztikai Társaság 53. kongresszusa 2006. 09. 02. Szeged,
- VI. **Nagy E**, Watine J, Bunting PS, Onody R, Oosterhuis WP, Rogic D, Sandberg S, Boda K, and Horvath AR. A nemzetközi diabetes mellitus irányelvek módszertani minősége. Magyar Tudomány Ünnepe, 2007. 11. 13. Szeged

- VII. Horvath AR, *Nagy E*, Watine J, Bunting P, Onody R, Oosterhuis W, Rogic D, Sandberg S, Boda K. Quality of Guidelines for Diabetes Mellitus: Can We Do Better? 17th IFCC-FESCC European Congress of Clinical Chemistry and Laboratory Medicine 3-7 June 2007, Amsterdam, The Netherlands

Ph.D. témájához köthető egyéb publikációk

Cikkek

- I. Kulier R, Hadley J, Weinbrenner S, Meyerrose B, Decsi T, Horvath AR, *Nagy E*, Empananza JI, Coppus S FPJ, Arvanitis TN, Burls A, Cabello JB, Kaczor M, Zanrei G, Pierer K, Stawiarz K, Kunz R, Mol BWJ and Khan KS. Harmonising Evidence-based medicine teaching: a study of the outcomes of e-learning in five European countries. BMC Medical Education 2008, 8:27 <http://www.biomedcentral.com/1472-6920/8/27>
- II. Kulier R, Hadley J, Coppus S FPJ, Zamora J, Hadley J, Malick S, Das K, Weinbrenner S, Meyerrose B, Decsi T, Horvath AR, *Nagy E*, Empananza JI, Arvanitis TN, Burls A, Cabello JB, Kaczor M, Zanrei G, Pierer K, Stawiarz K, Kunz R, Mol WJ B and Khan SK. The effectiveness of a clinically integrated e-learning course in evidence-based medicine: A cluster randomised controlled trial. BMC Medical Education 2009; 9:21 <http://www.biomedcentral.com/1472-6920/9/21>
- III. Kunz R, *Nagy E*, Coppus SFPJ, Empananza JI, Hadley J, Kulier R, Weinbrenner S, Arvanitis TNRT, Burls A, Cabello JB, Decsi T, Horvath AR, Walzak J, P. Kaczor PM, Zanrei GE, Pierer K, Schaffler R, Suter K, Mol WJB, Khan SK. How far did we get? How far to go? A European survey on postgraduate courses in evidence-based medicine Journal of Evaluation in Clinical Practice 2009; 15(6):1196-1204 (IF:1.843)
- IV. Hadley J, Kulier R, Zamora J, Coppus SFPJ, Weinbrenner S, Meyerrose B, Decsi T, Horvath AR, *Nagy E*, Empananza JI, Arvanitis TN, Burls A, Cabello JB, Kaczor M, Zanrei G, Pierer K, Kunz R, Wilkie V, Wall D, Mol BWJ, Khan KS. Effectiveness of an e-learning course in evidence-based medicine for foundation (internship) training. J R Soc Med 2010; 103: 288-294 (IF:1.402)
- V. Kis E, Szegesdi I, Dobos E, *Nagy E*, Boda K, Kemény L, Horvath AR. Quality assessment of clinical practice guidelines for adaptation in burn injury. Burns. 2010; 36 (5): 606-15. (IF:1.950)
- VI. Kis E, Dobos E, *Nagy E*, Kemény L, Horvath AR. A bizonyítékokon alapuló orvoslás gyakorlata. Bőrgyógyászati és Venerológiai Szemle. 2010; 86(4):103-107.
- Könyvfejezetek
- I. Watine J, Oosterhuis WP, *Nagy E*, Bunting PS, Horvath AR. Formulating and using evidence-based guidelines. In: Evidence-based Laboratory Medicine: From Principles to Practice. Price CP, Christenson RH (eds) AACC Press, Washington. 2nd edition, 2007. 275-294 pp.
- II. Dobos Éva, *Nagy Éva*, Horváth Andrea 8. fejezet A szakmai irányelvek fejlesztése In: A klinikai hatékonyság fejlesztése az egészségügyben. Szerk.: Gődény Sándor Budapest 2007. 427-464. old.
- III. *Nagy Éva*, Horváth Andrea 3. fejezet Bizonyítékokon alapuló laboratóriumi medicina In: Gyakorlati Laboratóriumi Medicina Szerk.: Debreczeni Lóránd, Kovács L. Gábor 2008. 2. (átdolgozott) kiadás 127-138. old.

Idézhető absztraktok

- I. Horvath AR, **Nagy E**. Evidence for monitoring hypothyroid patients on thyroxine replacement. *Clinica Chimica Acta* 2005; 355, S25 (IF:1.633)
- II. Seres E, **Nagy E**, Miko TL, Watine J, Horvath AR. The role of evidence based laboratory medicine in meeting ISO 15189:2003 standards for accreditation. *Clinica Chimica Acta* 2005; 355, S412 (IF:1.633)

Előadások

- I. Horvath AR., **Nagy E**. Guidelines in Central-Eastern European countries. AGREE/G-I-N International Meeting, Edinburgh, Scotland, Nov14-15, 2003
- II. Horvath AR, **Nagy E**. Evidence Based Healthcare Central European Perspective International Symposium of the German Agency for Quality in Medicine: 10 Jahre evidenzbasierte Gesundheitsversorgung (ÄZQ) 3 March 2005, Berlin
- III. Watine J, Friedberg B, Charet J-C, **Nagy E**, Onody R, Horvath AR. Conflicting Guideline Recommendations: a Practitioner's Dilemma. XIX International Congress on Clinical Chemistry/2005 AACC Annual Meeting. 24-28 July 2005, Orlando, USA
- IV. Horvath AR, **Nagy E**. Evidence for monitoring hypothyroid patients on thyroxine replacement. 16th IFCC-FESCC European Congress of Clinical Biochemistry and Laboratory Medicine 8-12 May 2005, Glasgow
- V. Horvath AR, **Nagy E**, Onody R. What is the Evidence Base for Biochemistry Testing? 28th Annual Conference of the Association of Clinical Biochemists in Ireland. 14-15 October 2005, Limerick, Ireland
- VI. **Nagy E**, Watine J, Horvath AR. Are Guidelines in Laboratory Medicine Evidence Based? How to Assess Quality of Guidelines? Evidence-based Laboratory Medicine Conference of Polish Laboratory Medicine College 2627 June 2009, Kraków, Poland