

概要

この項目では4項目においてウェブサイトにて記述される場合に各3点を加算する。

病因

この項目では10項目においてウェブサイトにて記述される場合に各3点を加算する。ただしアレルギーの項目は4項目のうち1つでも記述される場合に3点を加算する。

診断

この項目では10項目においてウェブサイトにて記述される場合に各1点を加算する。

治療

この項目では11項目においてウェブサイトにて記述される場合2~12点を加算する。ただし、発作治療薬のうちステロイド薬、ロイコトリエン拮抗薬、抗菌薬については1つでも記述があれば2点を加算する。

表3 医学的知識から作成した評価ツール

		チェック項目		点數	
観察	気道の慢性疾患である 炎症により気道が狭くなり、呼吸が苦しくなる 自然もしくは治療により、初期のうちにはもとにもどる 早期の発作時の炎症を抑える治療が有効		気道の慢性疾患である 炎症により気道が狭くなり、呼吸が苦しくなる 自然もしくは治療により、初期のうちにはもとにもどる 早期の発作時の炎症を抑える治療が有効	3 3 3 3	
		危険因子 (誘発原因)	遺伝	祖父母、父母、兄弟姉妹にアレルギーがあるか	3
			アレルギー	ダニ ネコ、イヌ、ハムスター、モルモットなどの動物 真菌類 花粉	3
		病態	無症状 (発作原因)	アレルギー	ダニ ネコ、イヌ、ハムスター、モルモットなどの動物 真菌類 花粉
運動	運動をしたときに発作が起こりやすい			3	
気候変化	急激な気候の変化で発作が起こりやすい			3	
食料、食品添加物	アレルギーとなっている食品や何らかの強い食品で発作が起こりやすい			3	
アルコール	飲酒によりぜんそく症状が悪化する			3	
薬物	アスピリン、非ステロイド性消炎薬で発作が起こる			3	
心理的ストレス 月経、過労	激しい感情表現や、ストレスでぜんそく症状が悪化する 月経や過労でぜんそくが悪化する			3 3	
診断	重症度	重症度	喘息自体の重症度 発作の頻度	1 1	
		検査法	血液検査	採血により、好酸球数の増加、アレルギー検査	1
	痰		一服前に少なく、水筒またはゼリー状。重症発作では粘りが増し、排出しにくくなる。気道の感染などがある場合には黄色い	1	
	胸部X線		診断では診断の決め手にはならず、他の疾患の除外診断に用いる	1	
	肺機能		診断では非発作時に正常で、発作が中重以上になると肺活量も減少する	1	
	アレルギーの特定	血液ガス	中程度の発作でpO ₂ 、pCO ₂ の両方が低下、重症化するでpCO ₂ は上昇	1	
		皮内反応テスト	少量のアレルゲンを皮膚内に注射し、アレルギーが何であるかを調べる	1	
		ブリックテスト	アレルギー薬を一薄皮膜にたらしめて薄をつけ、皮膚によってアレルギーを知る	1	
		RAST/MAST	血液を用いてアレルギーが何であるかを調べる	1	
	治療	薬物療法	長期管理薬 (コントローラー)	β2刺激薬の吸入、貼付、経口投与	2
テオフィリン徐放錠剤				2	
吸入ステロイド薬				2	
抗アレルギー薬				2	
11β2サイトカイン阻害薬				2	
発作治療薬 (リリーバー)			漢方薬	2	
			アミノフィリン	5	
			β2刺激薬 ステロイド薬	5	
ステップアップ、ステップダウン			ロイコトリエン拮抗薬 抗菌薬	2	
			現在の治療薬で発作を起こさない期間が3ヶ月以上続いて状態が安定していたら治療薬を1段階へるしてもよい。逆に現在の治療薬でも発作を起こしてしまうときには、もう1段階上の治療薬に進む。	12	
	薬物の副作用		12		

4. 評価の手順

評価の手順の目的は喘息を解説するウェブサイトの評価指標、正確さをもとめる正確さの手順で作成した評価ツールを用いて評価し、喘息を解説するウェブサイトの正確さと評価指標の各項目との相関を求めるためにカイ2乗検定を行い、P値を算出することである。

喘息を解説するウェブサイトの収集手順で収集したウェブサイト 92 件を医学的知識から作成された評価ツールによって評価を行い、正確さを求める。そのうち 25 点以上をより正確な情報を提供するウェブサイトと定義した。25 点以上であるウェブサイトは 24 件であり、25 点未満のウェブサイトは 68 件であった。25 点以上であるもののうち、評価指標の収集手順で収集された各評価項目に合致するかどうかチェックする。また 25 点未満のものについても同様に各評価項目に合致するかどうかチェックする。25 点以上か未満か、評価項目と合致するかどうかの 2 要素から 2×2 の分割表 (表 4) を作成することができ、カイ2乗検定によって P 値を算出する (図 3)。

表 4 2×2 分割表 (評価項目 “サイトの目的が書かれているか” の例)

	25 点以上 (n=24)	25 点未満 (n=68)
評価項目に合致	22	36
評価項目に不合致	2	32

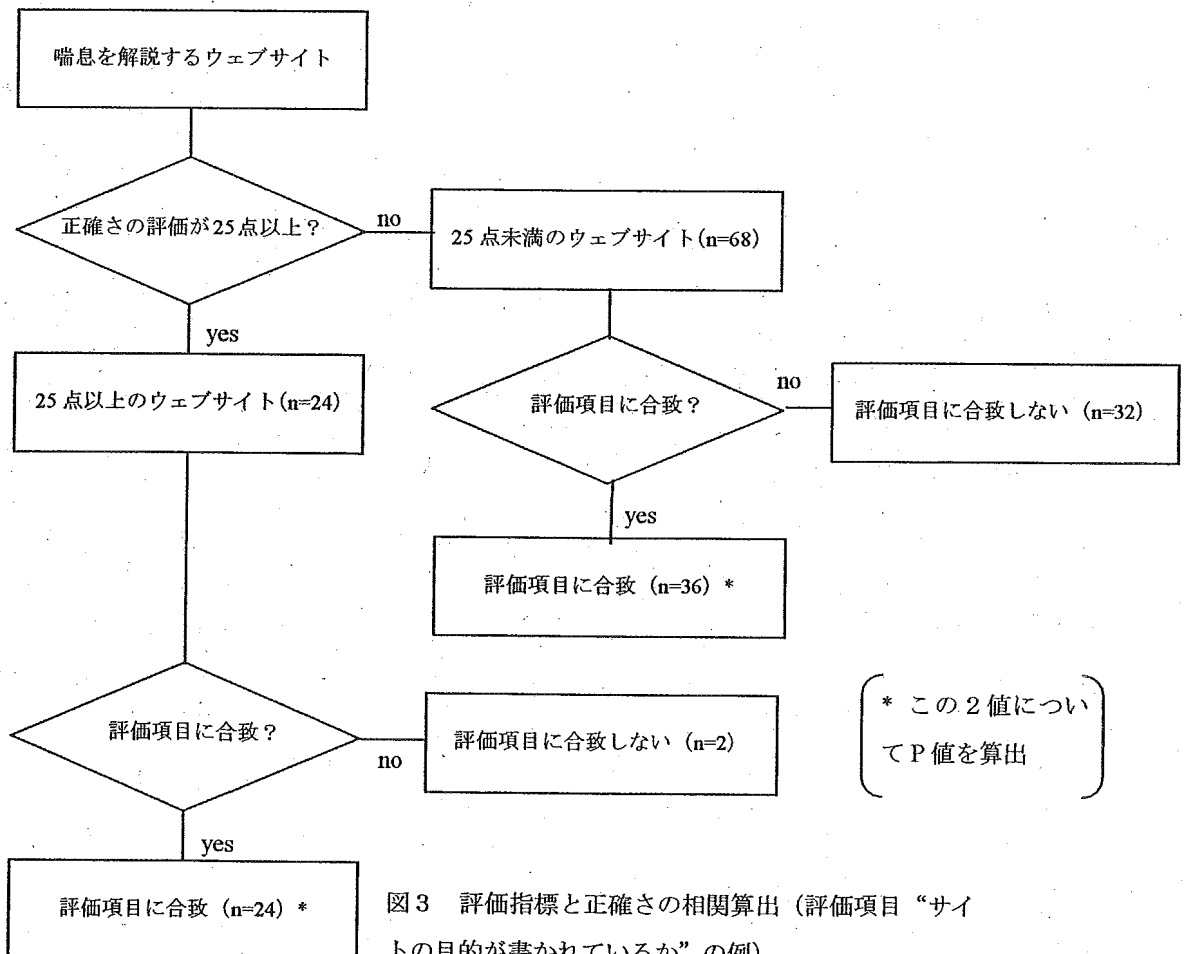


図 3 評価指標と正確さの相関算出 (評価項目 “サイトの目的が書かれているか” の例)

C. 研究結果

1. 喘息を解説するウェブサイト

検索語“ぜんそく OR 喘息”によって収集された喘息を解説するウェブサイトは 172 件の重複、336 件の除外項目との合致により 92 件が調査対象となった。92 件の喘息を解説するウェブサイトを 5 種類の URL のドメイン（表 5）による割合についてみると、「.com」が 40%、「.net」が 23%、「.org」が 16%、「.edu」が 5%、「.gov」が 7%を占めていた（図 4）。ただし現在では各ドメインを使用しているのがドメインを表す機関であるとは限らず、ある程度誰もが自由にドメインを取得している。

表 5 ドメインの種類と件数

ドメイン	説明	件数 (n=92)	割合 (%)
.com	商用ドメイン	37	40
.net	ネットワークスインフラ関連	21	23
.org	非営利団体	15	16
.edu	高等教育機関	5	5
.gov	政府機関	6	7

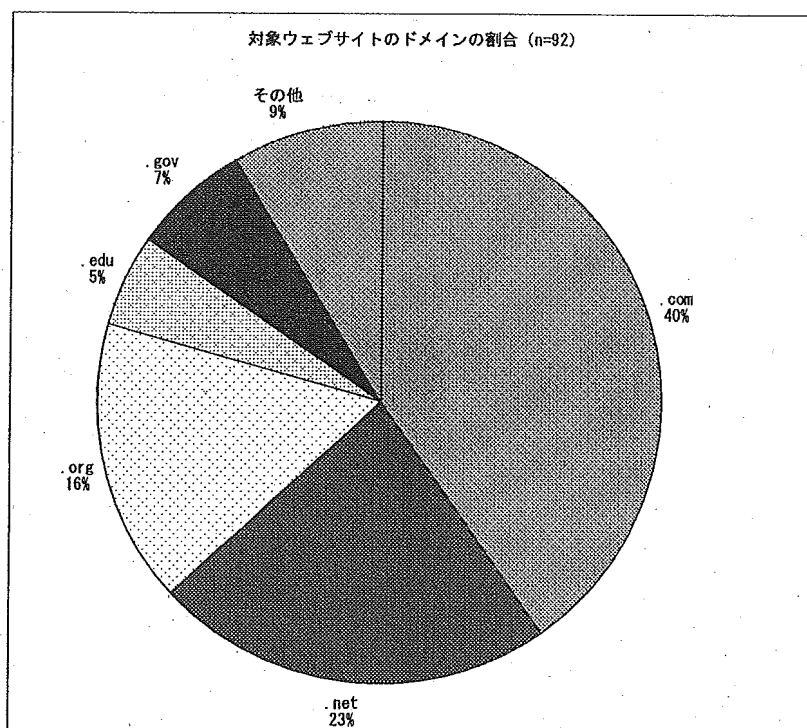


図 4 調査対象喘息ウェブサイトのドメイン別割合

92 件の喘息を解説するウェブサイトの内容において、著者が明らかにわかるものについて8つに区分した(表6)。(1)国、地方公共団体は著者が国や地方公共団体の機関である。

(2) 商業製品は製品を販売している商業機関である。(3) 商業は特定の商品販売しているのではない商業機関である。(4) 学術は大学や調査研究機関である。(5) 医師は個人、もしくは医療機関である。(6) 医師以外の医療従事者は、看護師などの医師以外の医療従事者である。(7) 一般個人は医療従事者ではなく、また機関としてウェブサイトを作成していない個人である。(8) 不明は、上記7区分によって分けることができないか、不明な著者である。以上8の区分における著者の属性については、34%が医師により作成され、続いて商業22%、商業製品14%、一般個人が10%と続いている(図5)

表6 著者の区分

著者のタイプ		件数
国、地方公共団体	国や地方公共団体	6
商業製品	製品を販売している商業機関	13
商業	特定の商品販売しているのではない商業機関	20
学術	大学や調査機関	4
医師	医師や医療機関	32
医師以外の医療従事者	医師以外の医療従事者	1
一般個人	医療従事者以外の一般人	9
不明	不明	7

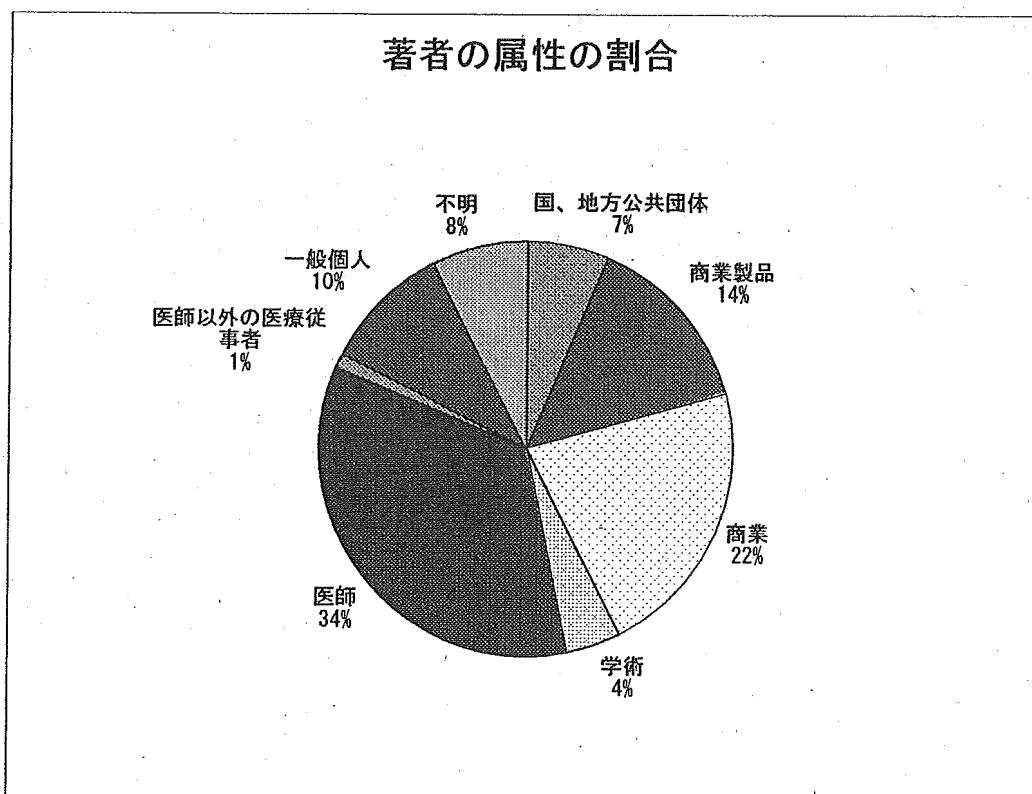


図5 喘息ウェブサイトの著者の属性

2. 収集した評価指標

31件の収集したウェブサイトに関する評価指標に含まれる評価項目は454項目であった。ひとつの評価指標あたりの平均は14.6項目（標準偏差9.5）で最も少ない4項目から最大39項目であった。評価指標には米国のNational Center for Complementary and Alternative Medicine (NCCAM) や National Cancer Institute (NCI)、NLM などの国家機関や米国医学図書館協会 (Medical Library Association: MLA) のような機能の団体が含まれている(表7)。評価指標の作成機関を表5と同様の区分のドメインによる割合については、学術機関「.edu」30%、非営利団体「.org」26%、政府機関「.gov」16%と続いている(図7)。

表7 収集したウェブサイト評価指標と項目数

評価指標	URL	項目数
NCCAM	http://nccam.nih.gov/health/webresources/	10
NCI	http://www.cancer.gov	12
NLM	http://nlm.gov/publib/criteria.html	9
MedlinePlus	http://www.nlm.nih.gov/medlineplus/	7
MLANET	http://www.mlanet.org/resources/userguide.html	9
National Institute on Aging	http://www.nia.gov/publications	7
JMA	http://www.jma.or.jp	10
DISCERN	http://www.discern.org.uk	15
ESOME	http://home.ac.uk/guests/eval/ewrta.html	6
VUW Department of library and information studies, New Zealand	http://www.vuw.ac.nz/staff/astair_smith/ebain/index.html	20
	http://jhsa.wrsb.on.ca/library/html/ebstate/ebainfo.htm	9
virtual salt	http://www.virtualsalt.com/evalufit.htm	4
Infopopple	http://infopopple.org/resources/bkwb/select.html	7
	http://www.cmsnet.org/publicdoc.cfm/60/0/gene1/99	7
to the council, the european parliament, the economic and social committee and the committee of the regions.		12
University of illinois at Chicago	http://www.usc.edu	10
HONcode	http://www.hon.ch/	8
University of texas at austin school of nursing	http://www.utexas.edu/nursing/nancy/webcriteria.html	14
Johns Hopkins University	http://www.library.jhu.edu	8
Discovery school	http://school.discovery.com	34
Dorison memorial library	http://dorison.uchec.edu	57
Center	http://www.usmc.edu	18
Health in Action	http://www.health-in-action.org	50
Betty C. Jung's website	http://www.bettysjung.net	22
	http://genome.qsc.nion.nu.jp	6
UC Berkeley	http://www.lib.berkeley.edu	15
UChC	http://library.uhc.edu	19
network	http://info.med.yale.edu	24
	http://www.library.dal.ca	18
Net scoring	http://www.chu.covent.uk	39
QUICK	http://quick.org.uk	8
#		454

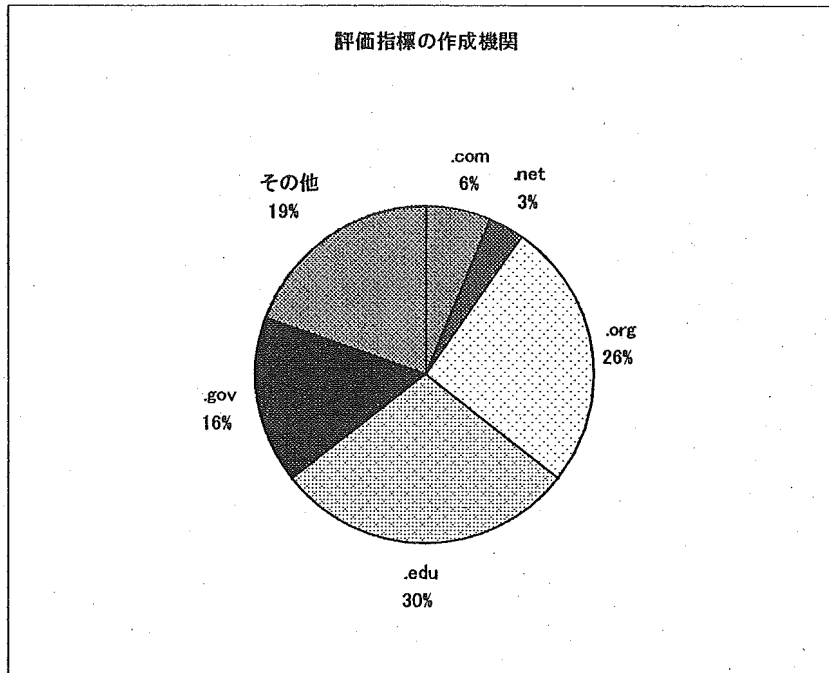


図 6 評価指標の作成機関

3. 喘息を解説するウェブサイトの正確さ

92 の喘息を解説するウェブサイトの正確さをはかるための手順において作成した医学的根拠から作成した評価ツールによって評価をおこなった。92 のウェブサイト全体での得点の平均は 100 点満点のうち 17.5 点（標準偏差 1.45）であり、最低点 0 点から最高点 77 点までであった（図 8）。評価ツールの項目のうち“概要”、“病因”、“診断”、“治療”について各概念別の平均点（標準偏差）は、概要 4.08 (3.23)、病因 8.22 (7.06)、診断 0.85 (1.94)、4.38 (7.98)であった（表 8）。

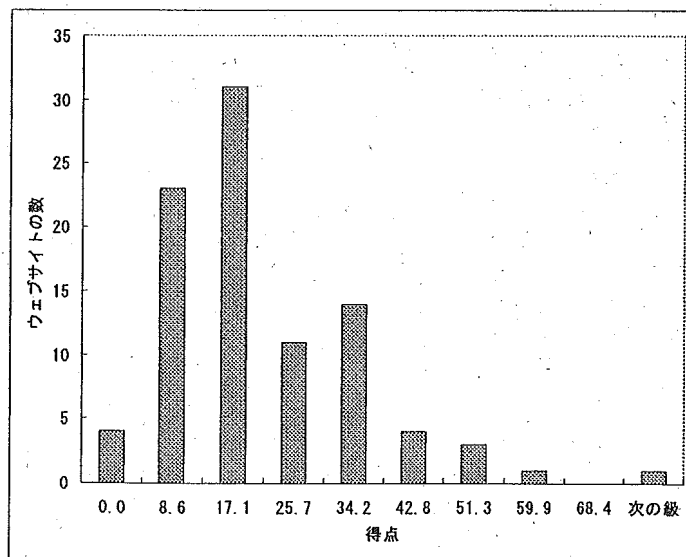


図 7 喘息を解説するウェブサイトの正確さの得点分布

表8 正確さ4概念の得点

	平均	標準偏差
概要	4.08	3.23
病因	8.22	7.06
診断	0.85	1.94
治療	4.38	7.98

正確さの評価の平均点をドメインによる区分けについてみると、「.com」が19.8点、「.net」15.2点、「.org」13.7点、「.gov」29.7点、「.edu」17.2点、その他のドメイン11.4点であった(表9)。正確さの評価の平均点を著者のタイプ別についてみると、国、地方公共団体が25.5点、一般個人が21.6点、学術17.8点と続いている(表10)。

表9 ドメインによつての平均点(標準偏差)

	平均点	標準偏差
.com	19.8	14.2
.net	15.2	12.5
.org	13.7	11.5
.gov	29.7	24.9
.edu	17.4	3.2
その他	11.4	7.6

表10 著者のタイプ別平均点(標準偏差)

	平均点	標準偏差
国、地方公共団体	25.7	26.6
商業製品	16.3	14.5
商業	15.4	12.5
学術	17.8	1.0
医師	17.4	13.5
医師以外の医療従事者	12.0	
一般個人	21.6	11.2
不明	14.7	13.5

4. 喘息を解説するウェブサイトの正確さと評価指標の相関関係

喘息を解説するウェブサイトの正確さと評価指標の相関関係については、“サイトの目

的が書かれているか”、“情報はどこからきているか”、“情報の基礎は何か”、“資料はその領域で包括的であるか”、“プライバシーポリシーが明示されているか”、“治療の便益とリスクに言及がある”、“他のサイトへのリンクの選択がよい”、“さらなる情報の詳細を提供している”、“適切な量のリンクがある”、“バランスがあり偏りが無い”、“対象がだれかはっきりわかる”の11項目であった(表11)。ドメインによる正確さの優位さにおいては「.com」が0.04となり0.05を下回った(表12)。所属による有為差はP値が0.091~0.900となかった(表13)。

表11 ウェブサイト評価項目と医学知識の正確さの関連

チェック項目		正確(n=24)	正確でない(n=68)	p 値
サイトの内容	サイトの目的がかかれているか	22	36	0.001
	情報はどこからきているか	12	15	0.010
	情報の基礎はなにか	12	15	0.010
	情報がどう選択されているか	0	1	0.550
	エディトリアルボードを持ち、情報はレビューされている	0	0	N/V
	サイトが健康要求を誇張してよくしていない。素早く、劇的に、奇跡的などを約束していない	24	61	0.102
	不確実な領域に言及がある	3	2	0.076
	治療の益とリスクが説明されている	6	5	0.022
	資料がよく知られているか、もしくは頻繁にリようされている	1	0	0.09
	資料はその領域で包括的である	6	1	0.0002
	スペル、文法に間違いがない	22	64	0.675
デザイン	訪れた人との相互作用がある	1	3	0.959
	大きな画像をのせたり、不必要な画像を使用していない	24	63	0.172
	資料から効果的に検索が行える	1	3	0.959
著者、スポンサー、開発者の公開	誰がサイトを運営しているかわかる	21	53	0.31
	誰がサイトに資金を出しているかわかる	18	48	0.680
	受け取った電子メールの正確さを人々はどうか確認するか	0	0	N/V

	ドメインのチェック(edu.gov.com.etc)	23	62	0.459
情報の流通	情報がどのくらい新しいかわかる	8	12	0.109
ソースの権威	インターネットに投函された健康要求の誤解やミスから政府は消費者をどう守っているか	0	0	N/V
	誰が内容に対して責任をもっているかわかる	17	53	0.483
利用の簡単さ	読みやすい(readability)	23	61	0.360
	サイトマップがある	6	9	0.180
接続性と利用可能性	ミラーサイトがある	0	0	N/V
	地理的なアクセス制限がない	24	66	0.396
	特別なハードウェアやソフトウェアが必要でない	22	65	0.466
	資料を利用するために登録を必要としない	24	65	0.295
	利用料金がかからない	23	66	0.771
	サイトの読み込みに時間がかからない	23	64	0.750
リンク	他のサイトへのリンクの選択はがよい	5	4	0.034
	さらなる情報の詳細を提供している	4	2	0.019
	適切な量のリンクがある	8	8	0.017
	リンクがアクティブである	8	11	0.074
帰属性と保管	バランスがあり、偏りがない	24	50	0.050
意図された利用者	対象は誰かはっきりわかる	24	58	0.047
コンタクトアドレスやフィードバック	連絡先があるか	20	47	0.178
利用者サポート	なんの個人情報を集め、それはなぜか(プライバシーポリシー)明示されている	8	3	0.000
	ヘルプがある	0	2	0.396

表 12 ドメインによる正確さ

	25点以上 (n=24)	25点未満 (n=68)	p 値
.com	14	23	0.04
.net	4	17	0.40
.org	2	13	0.22
.edu	0	5	0.17
.gov	3	3	0.17
その他	1	7	0.36

表 13 著者のタイプによる正確さ

著者	25点以上 (n=24)	25点未満 (n=68)	p 値
国、地方公共団体	2	4	0.676
商業製品	5	8	0.273
商業	5	15	0.900
学術	0	4	0.224
医師	8	24	0.862
医師以外の医療従事者	1	0	0.091
一般個人	3	6	0.602
不明	1	6	0.459

D. まとめ

喘息を解説するウェブサイトのドメイン別件数については、商業サイト「.com」がもっともおおく、次いでネットワークインフラ関連を表す「.net」、非営利団体「.org」が続いて多かった。教育的な「.edu」は5%と少なかった。著者のタイプについてみると医師がもっとも多く続いて商業や商業製品が多かった。

評価指標を提示しているウェブサイトのドメイン別件数については、「.edu」、「.org」、「.gov」と多く、ウェブサイトのドメイン別件数とは反対の結果となった。

喘息を解説するウェブサイトの正確さについては100点満点で平均点が17.5点と低かった。ドメイン別のウェブサイトの正確さの平均点は「.gov」で最も高く、「.com」、「.edu」と高かった。

喘息を解説するウェブサイトの正確さと評価指標の相関関係は、評価項目の「サイトの内容」、「リンク」、「帰属性と保管」、「意図された利用者」、「利用者サポート」(表11)にお

いて多くみられた。

本研究において、喘息を解説するウェブサイトの正確さ、そして、広く利用されている評価指標の評価項目のうち機能する項目と機能しない項目があることが明らかとなった。

参考文献

- (1) PEW/INTERNET. Health information online. Pew internet & American life project Washington.D. C. 2005.
- (2) 平成17年度版情報通信白書. 総務省. 東京. 2005. 322p.
- (3) 山口直比古[他]. 日本における EBM のためのデータベース構築および提供利用に関する調査研究. 平成12年度総括・分担研究報告書 厚生科学研究費補助金医療技術評価総合研究事業. 東京. 2001, p.19-35
- (4) 奈良岡功[ほか]. 健康・医学情報を市民へ. 日本医学図書館教会出版委員会. 東京. 2004.130p.
- (5) Eysenbach G. consumer health informatics. BMJ. vol.320, 2000, p.1713-1716
- (6) 平成16年度版情報通信白書. 総務省. 東京. 2004 p.183
- (7) 三谷博明[他]. インターネット上の医療情報の質をめぐる課題. 医学図書館. 2001, vol.48. no.2. p.155-161.
- (8) Eysenbach G[et al]. How do consumers search for and appraise health information on the world wide web? Qualitative study using focus groups, usability tests, and in-depth interviews. BMJ. 2002, vol.324, p.573-577.
- (9) American medical association. "Guidelines for Medical and Health Information Sites on the Internet" available from < <http://www.ama-assn.org/ama/pub/category/1905.html#GUIDE> > , (accessed 2006-1-25).
- (10) Medical library association. "Resources" available from < <http://www.mlanet.org/resources/userguide.html> > , (accessed 2006-1-25)
- (11) Health information technology institute. Available from < <http://hitiweb.mitrettek.org/iq/default.asp> > (accessed 2005-4-26)
- (12) DISCERN. "The DISCERN Instrument" available from < http://www.discern.org.uk/discern_instrument.php > , (accessed 2006-1-25)
- (13) 丁元鎮[他]. 消費者向け医療情報の評価ガイドライン. 臨床と薬物治療.2001,vol.20, no.6 p.696-701.
- (14) Health on the net foundation. Available form <<http://www.hon.ch/>> , (accessed 2006-1-25)
- (15) Kim P [et al]. published criteria for evaluating health related web sites : review.BMJ. 1999, 318, p.647-649.
- (16) Eysenbach G[et al]. Empirical studies assessing the quality of health information for consumers

- on the world wide web. JAMA. 2002, vol.287, no.20, p.2691-2700.
- (17) Alejandro R[et al]. Rating health information on the internet. JAMA. 1998, vol.79, no.8, p.611-614.
- (18) Sandvic H. Health information and interaction on the internet : a survey of female urinary incontinence. BMJ. 1999, vol.319, p.29-32.
- (19) Coulter A. Sharing decision with patients : is the information good enough. BMJ. 1999, vol.318, p.318-322.
- (20) Fagerlin, A et al . Patient education materials about the treatment of early-stage prostate cancer : A critical review. Annals of Internal Medicine. vol. 140, no. 9, 2004, p. 721-728.
- (21) Fricke M[他]. Consumer health information on the internet about carpal tunnel syndrome: indicators of accuracy. The American journal of medicine. 2005, no.2, vol.118, p.168-174.
- (22) Fallis D. Indicators of accuracy of consumer health information on the internet. JAMIA. 2002, no.9, p.73-79.
- (23) 石原享介. 喘息教室. 新興医学出版社. 東京. 2000. 144p.
- (24) National Institute of Health. "MedlinePlus" . available from <
<http://www.nlm.nih.gov/medlineplus/>> , (accessed 2006-1-25).
- (25) 福島雅典[訳]. メルクマニュアル医学情報 : 家庭版. 日経 BP 社. 東京.1999.3001p.
- (26) 日本クリニカル・エビデンス編集委員会. クリニカル・エビデンス. 日経 BP 社. 東京.2001.1749p.
- (27) 橋本信也[編]. 薬と病気の本 : 各科・専門のお医者さんによる. 保健同人社. 東京.2004.1279p.
- (28) 福井次矢.正しい治療がわかる本.法研.東京.2003.1079.
- (29) 中山健夫. EBM を用いた診療ガイドライン 作成・活用ガイド. 金原出版. 東京. 2004.106p
- (30) 宮本昭正[監修]. EBM に基づいた喘息治療ガイドライン 2004. 共和企画. 東京. 2004. 281p.
- (31) 宮本昭正[監修]. EBM に基づいた患者と医療スタッフのパートナーシップのための喘息診療ガイドライン 2004(成人編). 共和企画. 東京. 2004. 22p.
- (32) 前野昌弘[他]. 図解でわかる統計解析. 日本実業出版社. 東京. 2000. 227p.

メタ・アナリシスにおける2つの回帰直線の傾きの比の信頼区間
(医療技術評価総合研究事業) 分担研究報告書

研究協力者 高橋 邦彦 国立保健医療科学院技術評価部
主任研究者 丹後 俊郎 国立保健医療科学院技術評価部

研究要旨：ある2つのタイプそれぞれの回帰直線の傾きの比によって、2つのタイプのリスクの差などが観察できるが、メタ・アナリシスにおいてその指標を用いた研究の統合を行う際に、通常の公開されている研究結果・論文等には全てのデータが記載されることはほとんどなく、データをグループ化した要約統計量のみが表記されていることが多い。本研究では、線形単回帰モデルにおいて横軸の doze にある分布を仮定し、グループ化された各範囲の doze の値を最尤法により推定し回帰直線を推定する方法について検討を行った。実際には、横軸 X をいくつかの範囲に分けたグループごとの要約統計量が与えられている場合に、横軸に正規分布および対数正規分布の2つのモデルを仮定し、それぞれについて最尤推定法に基づく回帰係数の推定を行い、その推定精度をシミュレーションにより確かめた。

A. 研究目的

メタ・アナリシスのための解析手法の理論の多くは、各研究から十分なデータの情報が得られていることが前提となっているが、実際に統合を行うための各研究結果の論文などでは、必要な情報が得られないことがしばしば起こる。例えば、アルコール摂取量と血圧の関連をみる場合、最も簡単なモデルとして線形単回帰モデルを用いた分析が行われる。このような回帰直線を用いた研究結果を統合することは大変重要であるが、通常明記されているのは結果の回帰係数やその標準誤差が一般であり、研究結果・論文等に全てのデータが記載されることはほとんどない。このような回帰モデルによる解析を行う多くの研究論文等では一方の変量のあるカテゴリーのグループごとにまとめ、その基本統計量などの情報を各グループごとに記述した表として掲載されている。そのため、例えば2つの回帰直線の傾きの比などを求める際、その比の標準誤差の推定など、本来の全データを用いた場合に行うことが可能ないくつかの値の推定が再現不可能になる。

本研究では、上記のようなグループ化された範囲ごとのデータから全データによる回帰係数の推定法とその精度について考察を行った。

B. 研究方法

まず、アルコール摂取量 X と血圧 Y について、オリジナルの解析データ $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$

の間の関係として線形回帰モデル

$$y_i = \alpha + \beta x_i + \varepsilon_i \quad (i = 1, 2, \dots, N)$$

を想定するものとする。ただし、 ε_i は互いに独立にいずれも正規分布 $N(0, \sigma^2)$ に従うものとする。今、実際の研究結果として、Saito et al(2003) に下記のようなデータが記述されている。

表 1. アルコール摂取量と血圧 (SBP*) の関係 (Saito et al(2003))

Alcohol (g/week)	データ数	平均値 (mmHg)	SE
0 - 20.4	58	119.320	2.387
20.4 - 183.8	57	125.683	2.470
183.8 - 367.7	18	127.011	4.336
367.7 -	4	125.994	9.160

* adjusted for some factors

このデータから回帰直線の傾き β を推定する際、全データの十分な情報は得られていないので、各範囲の assigned dose とその範囲での平均値、標準誤差を用いた重み付き回帰による推定を行うことが考えられる。例えば、各範囲 $0 < x < 20.4$, $20.4 \leq x < 183.8$, $183.8 \leq x < 367.7$, $x \geq 367.7$ の assigned dose をそれぞれ 10, 100, 270, 450 ととると、 $\hat{\beta} = 0.0243$ が得られる。しかし、もし assigned dose として、0, 92, 266, 483 ととると $\hat{\beta} = 0.0226$ となる。このとき、論文によれば全データから推定された傾きは $\hat{\beta} = 0.0273$ である。このように、assigned dose の決め方によって傾きの推定値が異なってくるのがわかる。

そこで、本研究では assigned dose を、上記のグループ化されたデータを用いた最尤推定によって定めることを提案する。まず変量 X がある密度関数 $f(x, \theta)$ に従うと仮定し、 n_j ($j = 1, 2, \dots, m$) を範囲 $(u_j, v_j]$ 内のデータ数とする。このとき対数尤度関数

$$\sum_{j=1}^m n_j \log \Pr\{x_j \in (u_j, v_j]\}$$

を最大にするパラメータ θ と端点 v_m を求める。次に、この推定されたパラメータと端点を用いて、範囲ごとに切断分布を仮定する。つまり密度関数として

$$f_j(x, \hat{\theta}) = f(x, \hat{\theta}) \Big/ \int_{u_j}^{v_j} f(t, \hat{\theta}) dt, \quad u_j \leq x \leq v_j. \quad (1)$$

を考える。さらにその範囲での X の期待値

$$d_j = \int_{u_j}^{v_j} x f_j(x, \hat{\theta}) dx = \frac{\int_{u_j}^{v_j} x f(x, \hat{\theta}) dx}{\int_{u_j}^{v_j} f(x, \hat{\theta}) dx}, \quad (2)$$

を求め、この d_j を j 番目の範囲の assigned dose とする。

この d_j を用いて、回帰係数の推定を行う。 j 番目の範囲の y の平均 \bar{y}_j とその標準誤差 se_j を用いた重み付き回帰より傾き β は

$$\hat{\beta} = \frac{1}{D} \left(\sum_{j=1}^m \frac{1}{se_j^2} \sum_{j=1}^m \frac{d_j \bar{y}_j}{se_j^2} - \sum_{j=1}^m \frac{d_j}{se_j^2} \sum_{j=1}^m \frac{\bar{y}_j}{se_j^2} \right), \quad (3)$$

ただし

$$D = \sum_{j=1}^m \frac{1}{se_j^2} \sum_{j=1}^m \frac{d_j^2}{se_j^2} - \left(\sum_{j=1}^m \frac{d_j}{se_j^2} \right)^2$$

と推定できる。さらに切片 α の推定値は

$$\hat{\alpha} = \frac{1}{D} \left(\sum_{j=1}^m \frac{d_j^2}{se_j^2} \sum_{j=1}^m \frac{\bar{y}_j}{se_j^2} - \sum_{j=1}^m \frac{d_j}{se_j^2} \sum_{j=1}^m \frac{d_j \bar{y}_j}{se_j^2} \right) \quad (4)$$

となる。

C. 結果

表 1 のデータについて上記の推定法を当てはめて検討する。まず、アルコール摂取量 X の分布として正規分布を仮定し、各範囲に切断正規分布を仮定すると、各範囲の assigned dose は $d_1 = 10.493$, $d_2 = 98.919$,

$d_3 = 232.176$, $d_4 = 395.286$ と推定された。この値を用いて、重み付き回帰より

$$\hat{\beta} = 0.0292, \quad \hat{\alpha} = 120.567$$

と求められた。また、 X の分布として対数正規分布と各範囲に切断対数正規分布を仮定した場合、 $d_1 = 16.862$, $d_2 = 84.821$, $d_3 = 236.744$, $d_4 = 447.928$ が推定され、これを用いて

$$\hat{\beta} = 0.0247, \quad \hat{\alpha} = 120.980$$

となった。どちらの場合にも、オリジナルの全データによる値 $\hat{\beta} = 0.0273$, $\hat{\alpha} = 120.720$ のよい推定値を求めることができた。

さらに、この推定法の精度を調べるためシミュレーションを行った。実際の X の分布のモデルとして model I: 正規分布

$$Y = 88.0 + 0.25X + \varepsilon, \quad X \sim N(50.0, 16.0)$$

model II: 対数正規分布

$$Y = 88.0 + 0.25X + \varepsilon, \quad X \sim LN(3.0, 0.5)$$

の 2 つを設定した。ただし、 $\varepsilon \sim N(0.0, 10.0)$ とする。それぞれのモデルを真のモデルとして、そこから乱数を発生させ、そのデータをグループ化した上で上記の正規分布を当てはめた推定法、対数正規分布を当てはめた推定法によって回帰係数の推定を行った。これを 1000 回繰り返し、その推定誤差を relative error, MSE, se によって検討した。この結果、どちらもよい推定精度をもち、真のモデルと推定の際のモデルが違った場合でも、ある程度よい推定ができることが観察された (表 2)。

D. 考察

上記の観察では、表 1 の実際のデータでの適用は、一般の単純な重み付き回帰による推定値よりも精度よく推定できることが観察された。しかし、上端点 v_m の値を推定する際に、その推定精度が落ちたり不安定になる、つまり実際の数値計算で解が求まらないことが起こる場合があった。特にシミュレーションにおいては、真のモデルと当てはめたモデルが異なる場合にそのようなことが起こり、またより複雑である切断対数正規分布の場合に解が求まらない頻度があった。一方で正規分布の当てはめた場合には比較的安定して推定を行えることが観察された。

一般に医学的な観点などから、観察者が上端点を定めることができる場合には、 v_m を与えて推定するこ

とで推定が安定するといえる。しかし一方で、常識的な範囲であっても上端点を定めることが難しい場合には、いくつかの値で推定を行ってみて目的の回帰係数の推定値に影響がどの程度あるかを観察する必要があると考えられる。このような問題について更なる研究、検討が必要であろう。

E. 結論

一般に、グループ化された不十分なデータから本来の全データによる回帰係数の推定値の再現を行うことは不可能である。そのため重み付き回帰による方法 (Greenland and Longnecker, 1992) などが利用されることが多いが、その際、各グループの dose の決め方によって推定結果が変化してしまう。特に回帰係数の推定そのものだけでなく、その比を統合したい場合など、係数のわずかな推定結果の違いが比の場合に大きな違いを生み出すことがある。そのため、より正確な推定方法を用いることが重要である。そこで、本研究では、各範囲の dose を最尤法に基づき切断分布による期待値をもとに定める方法を提案した。実際のデータに当てはめた場合、従来の適当な dose を定める場合より精度がよく、さらに dose の決め方も一意的に定まるため有用な方法であることが確かめられた。さらに、その精度を調べるためシミュレーションにより検討を行い、正規分布の場合、対数正規分布の場合ともよい推定精度をもっていることが観察された。しかし一方で、この推定の問題点も明らかになった。そのひとつが安定性の問題である。分布

やパラメータの形や数が複雑になった場合、うまく解が得られない場合があることがわかった。また、1つの変量に分布を当てはめる際、異なる分布を当てはめてしまった場合にも不安定になることがわかった。本研究のシミュレーションでは正規分布を当てはめる方法が比較的安定している様子が観察されているが、今後更なる研究をすすめて、上記の問題点について検討を重ねることが重要である。そしてメタ・アナリシスの実際の回帰係数の比の統合への適用についても研究を進めていくことが望まれる。

F. 研究発表

特になし。

参考文献

1. Saito, K. and Yokoyama, T. and Yoshiike, N. and Date, C. and Yamamoto, A. and Muramatsu M. and Tanaka, H. (2003). Do the ethanol metabolizing enzymes modify the relationship between alcohol consumption and blood pressure?, *Journal of Hypertension*, **21**, 1097-1105.
2. Greenland S. and Longnecker P. (1992). Methods for trend estimation from summarized dose-response data, with applications to meta-analysis, *American Journal of Epidemiology* **135**(11), 1301-1309.

表 2. シミュレーション結果

model I with estimation of v_m

model	β			α			no. of runs
	relative error	mse	se	relative error	mse	se	
normal	0.00559	0.0083	0.0083	-0.00073	20.769	20.786	1000/1003
lognormal	0.22373	0.0094	0.0063	0.03285	23.930	15.586	1000/2777

model I without estimation of v_m ($v_m = 60$)

model	β			α			no. of runs
	relative error	mse	se	relative error	mse	se	
normal	0.02923	0.0086	0.0086	-0.00399	21.640	21.538	1000/1000
lognormal	0.10365	0.0171	0.0164	-0.01410	42.797	41.299	1000/1125

model II with estimation of v_m

model	β			α			no. of runs
	relative error	mse	se	relative error	mse	se	
normal	0.11228	0.0018	0.0010	-0.00547	0.7999	0.5688	1000/1009
lognormal	0.05841	0.0010	0.0008	-0.00351	0.5923	0.4973	1000/1108

model II without estimation of v_m ($v_m = 100$)

model	β			α			no. of runs
	relative error	mse	se	relative error	mse	se	
normal	0.05989	1.2456	1.1981	-0.00251	0.0028	0.0026	1000/1021
lognormal	0.05919	0.0010	0.0008	-0.00358	0.5923	0.4933	1000/1000

メタ・アナリシス研究のチェックリストに関する調査

分担研究者 折笠 秀樹 富山大学医学部教授
研究協力者 酒井 弘憲 富山大学医学部協力研究員

研究要旨

メタ・アナリシス研究の品質を評価することは必須になってきている。そのためチェックリストはこれまで発表されてきたが、本研究ではそれらを網羅的に調査することにした。その結果、従来からよく知られている QUOROM (ランダム化比較試験によるメタ・アナリシスのチェックリスト)、MOOSE (観察研究によるメタ・アナリシスのチェックリスト) に加えていくつものチェックリストが明らかになった。QUOROM は 18 問、MOOSE は 35 問とその評価は困難すぎており、OQAQ は 10 問、Fletcher 式は 5 問と多少抽象的すぎると思われた。3 件の同様のテーマに関するメタ・アナリシス研究へ MOOSE を適用した結果、2 名の評価者間で評価結果に相違を認める項目が散見された。さらに良いチェックリストの開発が望まれた。

A. 研究目的

メタ・アナリシスを実施した研究論文は数多く出版されるようになったが、その中にはその品質が怪しいものも少なくない。そうした品質をチェックするツールとしてどのようなものが現在までに使用可能かを文献調査することが本研究の目的である。

B. 研究方法

メタ・アナリシスのチェックリストとして使用可能なものを検索するために、まず Medline を用いた。MeSH キーワードである 4 語 (Meta-Analysis, Systematic review, Review literature, Overview) のいずれかであ

り、しかも Checklist.mp (Checklist という単語) を含む文献で検索した。さらに、マニュアルによりメタ・アナリシスのチェックリスト (通読ガイド) を検索した。

C. 研究結果

最初に Medline データベースによるメタ・アナリシスのためのチェックリストとしては、よく知られている QUOROM (18 問からなる) 及び MOOSE (35 問からなる) のほかに、経済評価のメタ・アナリシスに関する 6 問からなるチェックリスト (Jefferson T, Demicheli V, Vale L: Quality of systematic reviews of economic

evaluations in health care. JAMA. 2002; 287(21): 2809-2812.)が挙げられた。ランダム化比較試験によるメタ・アナリシスのチェックリストである QUOROM に関する文献リストを表 1 にまとめた。また、観察研究に関するチェックリストである MOOSE に関する文献リストを表 2 にまとめた。QUOROM についてはその紹介、及び適用した事例など広く知られているが、MOOSE に関する文献の基本的に 1 つだけであり、その適用事例もなかった。

続いて、マニュアルサーチにより調べたところ、3 つのチェックリストが挙げられた。それらを表 3 に示したが、Oxman-Guyatt による OQAQ (Overview Quality Assessment Questionnaire) という 10 問からなるチェックリスト (表 4 参照) が第一である。第二は著名な臨床疫学の教科書の中に見られたチェックリスト (表 5 参照) である。これはさらに単純なものであり、全部で 5 問からなるチェックリストである。第三番目は EBM の世界で著名な教科書の中に見られるチェックリストである。表 6 に示したような内容であるが、読者向けのガイドという色彩を呈している。従って、チェックリストは表 6 の第一項が主と思われる。

最後に、MOOSE という観察研究によるメタ・アナリシスのチェックリストを適用した事例が見られなかったため、3 つの同様のテーマに関するメタ・アナリシス研究に対して適用してみた。2 名の専門家 (酒井、折笠) が独立に評価し、不整合の場合には協議の上決定した。その結果を表 7 に示した。チェックシートの日本語訳としては、表 2 に示した中山らのものを用いた。その中で誤解を招きやすいチェック項目がいく

つかあった。第一は、「3. 研究結果の説明」であった。これまでの研究について結果が説明されたかと理解してしまうが、これは研究の Outcome(s)と英語では書かれており、しかも次のチェック項目として曝露・介入があるので、これはエンドポイントが何かを説明したかの設問であることに気づく。従って、「エンドポイントの説明」と理解すべきであると思われる。第二は、「6. 研究対象集団」であった。これはメタ・アナリシスの対象となった研究と誤解されやすかった。英語を見ると Study Population とあり、これが背景に関するチェック項目になっていることも勘案すると、この研究成果が適用される集団のことと思われた。従って、「研究ターゲット集団」と取るべきであろう。第三は、「11. 調査のために使用されたソフトウェア・・・」であった。これを統計ソフトと理解されることがあるようであった。よく読むとこれは、たとえば Explosion などと書かれている。チェック項目#10 が使用されたデータベースの種類、チェック項目#11 はサーチの仕方 (Publication Type などを含む) と理解したほうが正しいのではないかと思われた。また、チェック項目#8 には検索方式があるので、検索語とその論理集合などはそちらに含めるべきなのであろう。従って、「サーチに使用されたソフトウェア及びサーチの手法」としたほうが誤解を招かないように思われた。第四は、「17. 検証すべき仮説を評価するために集められた研究の関連性及び適切性」であった。これは意味がよくわからないという意見や、集められた研究の異質性の評価と誤って取られることもあった。しかしよく英語を読むと、仮説との関連性・適切性と読

める。従って、「統合の対象とした研究と仮説との関連性・適切性についての記述」というように、仮説との関連性・適切性を強調すべきであろうと思われた。

第五は、チェック項目#18, #19 に関することである。「データ」というのが少しわかりにくかったが、これは論文中で質の評価に使う項目といった意味と思われる。たとえば、追跡年数や割付法の隠蔽化などである。すなわち、論文からどのような項目を抽出し、それらをどのように分類し、コード化したかというチェック項目と思われた。項目#18 が抽出・コード化の意味づけ、項目#19 がその具体的方法と思われた。

D. 考察

メタ・アナリシス論文の品質評価は言うまでもなく大切である。そのチェックリストを網羅的に調査した結果、著名である QUOROM, MOOSE 以外にもいくつか挙がってきた。MOOSE を 3 つのメタ・アナリシスに適用した結果、あまりにも詳細すぎており、しかもその意味するところを誤解すると思われるチェック項目も散見された。もう少し簡略なチェックリストが望まれる。そうは言うものの、表 4~6 のようにあまりに抽象的だと評価者間で相違が生じる原因になる。従って、評価者間で相違を生じさせないような、より具体的なチェックリストを開発する必要性が認識された。現在、そのための新規チェックリストを考案中である。

E. 結論

メタ・アナリシスのチェックリストとして現在入手可能なものとして 6 種類を検索

した。MOOSE チェックリストを適用してみた結果、評価者間で相違をもたらすと思われる項目が散見された。今後はもっと評価者間信頼性が高いチェックリストの考案が求められる。

F. 健康危険情報

特になし。

G. 研究発表

- 1) Origasa H, Ikeda Y, Shimada K, Shigematsu H: Oral beraprost sodium as a prostaglandin I₂ analogue for vascular events in patients with peripheral arterial disease: meta-analysis of two placebo-controlled randomized trials. Jap J Pharmacoepidemiol, 2004; 9(2): 45-51.
- 2) Shimada S, Yokoyama N, Origasa H, Tsuneki H, Kimura I: Progressive bone loss due to androgen deprivation therapy for prostate cancer: a meta-analysis. Jpn J Pharm Health Care Sci, 2005; 31(3): 203-210.
- 3) 宮原英夫、折笠秀樹 (監訳): 実践医学統計学 (Using and Understanding Medical Statistics, 3rd edition). 東京: 朝倉書店, March, 2005.
- 4) 折笠秀樹: 「治療後の生存率」, p. 54-55; 「有効率」, p. 56-57; 「QOL」, p. 66-67; 「メガトリアル」, p. 74-75. In: 「統計学 100 のキーワード (松原望、編)」 東京: 弘文堂, April, 2005.
- 5) 田崎美弥子, 石井八重子, 海老原良典, 折笠秀樹, 高山美智代, 広瀬信義, 角間

- 辰之, 加藤芳朗, 国吉緑, LeeJung Won, 鈴木千智, 長谷川恵美子, 藤井美和, 畑田けい子, 松田正巳, WHOQOL-OLD 調査票日本語版開発グループ: 高齢者の Quality of Life (QOL) 調査票開発プロジェクトにおける予備調査結果. 老年精神医学雑誌, 2005; 16(2): 221-227.
- 6) Hirashima Y, Hamada H, Kurimoto M, Origasa H, Endo S: Decrease in platelet count is an independent risk factor for symptomatic vasospasm following aneurismal subarachnoid hemorrhage. J Neurosurg, 2005; 102: 882-887.
- 7) Gotoh M, Kamihira O, Kinukawa T, Ono Y, Ohshima S, Origasa H, on behalf of the Tokai Urological Clinical Trial Group: Comparison of α_{1a} -selective adrenoceptor antagonist, tamsulosin, and α_{1d} -selective adrenoceptor antagonist, naftopidil, for efficacy and safety in the treatment of benign prostatic hyperplasia; a randomized controlled trial. Br J Urol Int, 2005; 96: 581-586.
- 8) 折笠秀樹: 臨床試験のための生物統計学. 臨床薬理, 2005; 36(3): 141-145.
- 9) 折笠秀樹: メタボリックシンドロームの介入試験. 分子脳血管病, 2006; 5(1): 55-60.
- 10) 折笠秀樹: 創薬、臨床試験、そして統計学. 日本動脈硬化学会・シンポジウム, 2005. 7, 東京. (日本動脈硬化学会総会プログラム・抄録集, p.133)
- 11) Yokoyama M, Origasa H, Matsuzki M, Matsuzawa Y, Saito Y, Ishikawa Y, Oikawa S, Sasaki J, Hishida H, Itakura H, Kita T, Kitabatake A, Nakaya N, Sakata T, Shimada K, Shirato K, for the JELIS Investigators: Effects of eicosapentaenoic acid (EPA) on major cardiovascular events in hypercholesterolemic patients: The Japan EPA Lipid Intervention Study (JELIS). 2005 American Heart Association (AHA) Late Breaking Session, Dallas, 2005. 11.
- 12) 折笠秀樹: ガイドラインの策定とその評価. 第40回糖尿病学の進歩, 2006. 2, 金沢. (プログラム・講演要旨集, p.129)