

**Table 3. Algorithmic Mean Scores of the 5 BeautyQoL Dimensions and Index per Target Language**

	No. of Subjects	BeautyQoL Dimension, % of Subjects <sup>a</sup>					BeautyQoL Index, % of Subjects
		Social Life	Self-confidence	Mood	Energy	Attractiveness	
UK	200	42.4	45.5	44.3	40.7	41.3	43.3
Swe	200	35.0	40.8	36.7	35.2	37.5	37.1
Frn	200	40.4	46.9	43.8	41.2	44.9	43.2
Itl	200	44.4	50.0	46.6	46.8	46.8	46.6
Spn	200	41.2	47.3	44.9	42.8	43.4	43.7
Ger	200	39.0	44.9	41.9	39.1	41.0	41.1
Rus	200	58.1	65.1	60.1	60.2	62.5	60.9
Brz	213	50.3	58.4	50.2	47.4	53.3	52.2
Chn	200	65.6	66.3	64.4	61.5	64.8	65.3
US	200	46.3	49.6	48.6	45.2	45.4	47.4
Ind/Eng	206	70.5	72.4	70.2	68.9	70.8	70.9
Ind/Hindi	206	69.9	71.1	66.0	67.9	66.7	69.1
SA/Eng	202	67.6	71.0	69.2	66.3	66.5	68.7
SA/Sotho	201	78.6	79.9	78.9	76.9	77.7	78.9
SA/Zulu	203	75.0	76.3	75.0	73.8	71.8	75.0
Jpn	200	43.8	48.5	49.0	46.4	47.1	46.6

Abbreviations: See Table 1, QoL, quality of life.

<sup>a</sup> Overall statistical significance between countries was calculated using the Kruskal-Wallis test.

high-coverage foundation was significantly higher than the global score of the group receiving the high-coverage loose powder at days 7 and 21 (Mann-Whitney test,  $P < .05$ ).

General acceptability was high according to the very low proportion of missing data (<1%). The average time of completion ranged from 3 minutes in Japan to 16 minutes in Hindi-speaking India, which is fully compatible with clinical practice (completion time in China was not taken into consideration because of local survey briefing issues).

### SCORING PROCEDURES

The scoring procedure of the BeautyQoL instrument had 3 levels of extensive analyses. Level 1 consisted of a descriptive analysis on a question-by-question basis. Level 2 consisted of a classic “algorithmic scoring” that led to 1 score per dimension (profile) and 1 overall score (index). The scoring procedure was based on the mean score per dimension linearly transformed to a scale of 0 to 100, with 100 indicating the best possible level of QoL and 0 indicating the worst. **Table 3** presents the algorithmic scores of the 5 dimensions in each of the 16 participating languages, underlying various differences.

For example, subjects in India and especially in South Africa have, on average, much more positive views of their social life and mood, as well as energy and attractiveness (about 70%-80%), compared with people in other countries. In addition, a global index score was computed as the mean of the dimension scores. Negatively worded item scores were reversed so that higher scores indicated a higher level of QoL. Algorithmic scores of the 5 BeautyQoL dimensions are presented in **Table 4** by sex, age group (scores decreased significantly when age increased [Kruskal-Wallis test,  $P < .05$ ]), family status (single people presented significantly higher BeautyQoL scores than people living as part of a couple or with relatives [Kruskal-Wallis test,  $P < .05$ ]), labor force status (retired/pensioner category had significantly lower BeautyQoL scores than other

categories, and students had significantly higher scores than other categories [Kruskal-Wallis test,  $P < .05$ ]), housing status (the renting category had significantly lower BeautyQoL scores than other categories [Kruskal-Wallis test,  $P < .05$ ]), level of education achieved, and primary residence. **Table 5** presents the BeautyQoL score by facial skin type and skin tone. Mann-Whitney tests confirmed significant statistical differences ( $P < .05$ ) in all dimensions of BeautyQoL scores according to facial skin type and skin tone. Level 3 analysis was based on vector projections using the PCA. This innovative approach led to vectorial scores that complemented the classic algorithmic scores and was developed specifically in the frame of the BeautyQoL initiative to significantly improve the sensitivity of the scoring procedure. For example, the classic algorithmic scoring analysis (as proposed in level 2 analysis) could not discriminate 2 opposite profiles because the mean scores were similar. By extension, several profiles of answers could generate the same score. Level 3 analysis proposed to use the data set of the 3231 subjects as a reference corpus for the PCA and to analyze subjects from any other studies as additional subjects. The projection of answers on the main axis defined vectors, which could be linearly transformed from 0 to 100 to generate vectorial scores. After testing various answering profiles leading to similar algorithmic scores of 20, 40, and 60 against the vectorial scoring approach, we established that this new approach was able to discriminate all profiles by different vectorial scores, leading to similar algorithmic scores and significantly improving the sensitivity of the instrument. The Pearson correlation between algorithmic scores and vectorial scores was 1.00, confirming an affine relationship between the 2 indicators.

### COMMENT

One of the unique strengths of the BeautyQoL research initiative is the global approach involving 13 countries rep-

**Table 4. Algorithmic Scores of the 5 BeautyQoL Dimensions and Index According to Key Sociodemographic Variables**

	BeautyQoL Dimension, % of Subjects					BeautyQoL Index, % of Subjects
	Social Life	Self-confidence	Mood	Energy	Attractiveness	
Sex						
Male	54.3	57.5	55.1	53.4	53.8	55.0
Female	55.1	59.5	56.6	54.2	56.5	56.4
Age group, y						
18-28	60.9	65.2	62.3	59.4	62.1	62.1
29-38	56.8	60.8	57.6	55.9	57.5	57.8
39-48	52.3	55.9	52.8	51.5	52.1	53.1
49-58	48.8	51.8	50.2	47.8	48.8	49.6
59-68	41.1	44.9	44.6	43.0	41.2	42.8
Family status						
Couple	53.7	57.7	54.7	52.6	54.1	54.7
Single	57.0	60.5	58.6	56.5	57.4	58.1
Live with relatives	53.8	57.8	55.2	53.4	54.7	55.0
Labor status						
Employed	56.4	60.2	57.5	55.3	56.8	57.4
Unemployed	48.0	51.8	50.3	48.6	49.2	49.5
Retired/pensioner	42.6	46.3	45.6	43.1	43.1	44.0
Student	58.9	63.6	60.7	57.1	60.9	60.3
Homemaker	55.7	59.6	55.1	54.7	55.2	56.3
Housing						
Own home	55.7	59.1	56.3	54.7	56.1	56.5
Renting	50.2	55.0	52.5	49.3	51.2	51.7
Cohabiting	61.1	64.0	62.4	60.3	60.4	61.8
Other	52.4	57.4	53.5	53.0	53.5	53.9
Educational level						
Primary <sup>a</sup>	52.8	55.4	54.7	52.5	52.2	53.6
Secondary <sup>b</sup>	57.8	60.6	59.0	56.8	57.1	58.5
Tertiary <sup>c</sup>	53.4	57.7	54.6	52.7	54.4	54.6
Primary residence						
Urban	57.1	60.7	58.1	56.1	57.5	58.0
Rural	40.5	45.6	42.8	40.2	41.5	42.1

Abbreviation: QoL, quality of life.

<sup>a</sup> Usually begins at 5, 6, or 7 years of age and lasts for 4 to 6 years.

<sup>b</sup> Generally continues the basic programs of the primary level, although teaching is typically more subject focused.

<sup>c</sup> Largely theory based and designed to provide sufficient qualifications for entry to advanced research programs and professions with a high skill requirement.

**Table 5. Algorithmic Scores of the 5 BeautyQoL Dimensions and Index According to Facial Skin Properties**

	BeautyQoL Dimension, % of Subjects					BeautyQoL Index, % of Subjects
	Social Life	Self-confidence	Mood	Energy	Attractiveness	
Facial skin type						
Very oily	60.1	62.7	61.1	57.8	59.5	60.5
Oily	54.7	58.2	55.8	54.2	54.6	55.6
Normal	56.4	59.8	57.2	55.4	56.7	57.2
Dry	52.8	57.6	53.6	52.3	53.7	54.1
Facial skin tone						
Very fair	51.6	56.7	53.8	51.1	53.4	53.3
Fair	53.7	58.0	54.8	52.8	54.3	54.8
Medium	55.3	58.8	56.6	54.4	55.6	56.2
Dark	56.1	59.6	57.0	55.4	56.3	57.0
Very dark	67.4	68.6	65.1	70.7	65.4	67.3

Abbreviation: QoL, quality of life.

representing 16 languages worldwide. Such breadth allowed us to generate a fully validated instrument, available internationally, that captures the cultural variability between highly dissimilar countries and differences between people's attitudes within a single country. A multistep iterative process was used to generate subjects' responses

and to test their validity. The initial semidirected interviews conducted in 10 countries established potentially relevant categories, and an acceptability study completed with more than 800 subjects in 13 countries representing 16 languages guided the development of the categories for a full-scale assessment. After more than 3200 men and

women contributed to the full-scale validation assessment, their responses were analyzed and a subset of categories was created that captured most of the variability in subjects' BeautyQoL attitudes. A final follow-up assessment was used on a subset of the subjects to confirm their reproducibility. The 1-week interval was selected for the test-retest reliability to minimize the likelihood of changes in the subjects' mood while allowing sufficient time between testing and retesting to avoid the subjects recalling their previous answers. Most of the reviewed studies addressing cosmetic QoL improvement measures are limited by sample size. A small sample size prevents investigators from establishing valid QoL measures owing to a lack of variability in subjects' responses. As a result, few relevant QoL categories can be established, and the resultant instrument has limited interest for potential cross-validation use in a broader international audience. We are confident that the resulting BeautyQoL instrument reflects universal categories of cosmetic QoL attitudes and a range of unique cultural characteristics. Future assessments using the BeautyQoL instrument can be deployed across a large spectrum of potential scientific and market research applications.

Another notable strength of the BeautyQoL instrument is its simultaneous codevelopment in 16 languages. This simultaneous development of the instrument permitted the creation of a standard assessment tool in a relatively short time compared with the sequential strategy. When an instrument is created in one country and later adapted to another within the frame of sequential cross-cultural validations, the process takes a very long time and the subsequent versions may not have the structural equivalence and thus will lose validity in comparisons across multiple studies and cultures. Reliability in subsequent studies also becomes an important problem. The simultaneous construction of a multicountry assessment tool, the BeautyQoL questionnaire, solves these structural problems and allows for reliable studies across cultures.

Particular efforts have been deployed to generate relevant items compared with a number of existing QoL instruments. Most recent QoL instruments are based on expert opinions or adapted from other questionnaires, an approach that leads to a limited set of assessment criteria that experts or previous investigators found most relevant. We elected to create a comprehensive instrument that reflected a thorough understanding of the general public's opinion and QoL attitudes.

The international codevelopment approach has a number of advantages but also creates some limitations. First, international standardization of the selection of relevant questions excludes a few important aspects of physical appearance according to some cultures. Second, the 1-week interval between testing and retesting can be considered too short compared with the time frame necessary to observe final changes in physical appearance after some cosmetic interventions or product use in the long term. Third, sensitivity to change is expected to vary between countries and can be considered a limitation for conducting international studies if not further investigated in the frame of comparative surveys.

Nonetheless, we are confident that the resulting instrument contains the most relevant and cross-culturally valid

categories that reflect the general population's attitudes toward physical appearance. For example, subjects in highly industrialized European countries (and the United States and Japan) seem to be much less satisfied with their QoL, as established by their BeautyQoL scores (about 40%). Such discrepancy may appear paradoxical, given that these countries have greater access to consumer goods and health care. This finding may be explained by behavioral economics demonstrating that, when people have many choices in consumer products, they are less like to make a purchase and their satisfaction is lower than when they have fewer options.<sup>19</sup> Some sociological studies found significant ethnic differences in positive body image, with African American women scoring higher than white American women.<sup>20</sup> Given that weight has been traditionally regarded as a symbol of prosperity and status in rural Africa, we should expect the cultural differences found in our study, with South African women having the most positive body image. This interpretation is consistent with the fact that subjects with very fair skin had lower BeautyQoL scores, and the scores increased with darkness of the skin (Table 5). This finding confirms the role of racial differences when assessing QoL.<sup>21</sup> Calvert et al<sup>22</sup> assessed the QoL of 5408 subjects from the South Asian and African Caribbean communities in the United Kingdom and have reported significantly higher QoL in these minority ethnic groups compared with UK normative data, confirming a difference of perception of QoL according to ethnic communities. An additional factor in positive self-image is lack of exposure to popular media beauty standards, that is, women's attitudes toward their self-image tend to worsen during the introduction of television and popular culture beauty standards. However, this explanation is not entirely satisfactory because we expected that people from rural populations would score higher in their assessment of social life, mood, energy, and attractiveness compared with urban populations; however, urban populations scored higher than rural. In addition, having a secondary rather than a primary or a tertiary education appears to be correlated with higher QoL scores as measured by the BeautyQoL instrument. This argument can be supported given that the urban population is more highly educated in general according to a report from the United Nations Educational, Scientific, and Cultural Organization<sup>23</sup> (Table 4). Another counterintuitive finding is that subjects with oily skin appear to have the highest QoL scores (Table 5). One interpretation is that dry skin conditions are often not considered important by the subjects and even by health care professionals. As a result, problems with untreated dry skin can lead to a variety of issues that affect QoL, such as pruritus or ichthyosis vulgaris, which can be distressing, whereas eczema and psoriasis can lead to more serious consequences, such as fissures and infections.<sup>24,25</sup>

In conclusion, although the results of this study confirm the interest and international validity of the BeautyQoL instrument, the next step in our research agenda is to stimulate further studies that will use the BeautyQoL instrument to assess the QoL of subjects, such as in clinical trials, case-control studies, and cohort studies. Another step will be to perform additional validation studies in more countries, because further data collection would allow a more nuanced analysis of cultural differences.

Accepted for Publication: May 25, 2012.

**Author Affiliations:** Data Mining International, Geneva, Switzerland (Dr Beresniak); Laboratory for Applied Research in Health Economics, Paris-Descartes University, Paris, France (Dr Beresniak); L'Oréal Research, Clichy, France (Ms de Linares); Department of Dermatology, University of Utah, Salt Lake City (Dr Krueger); Department of Dermatology, Federal University of São Paulo, São Paulo, Brazil (Dr Talarico); Department of Drug Policy and Management, University of Tokyo, Tokyo, Japan (Dr Tsutani); Laboratory of Health Systems Analyses, Centre National de Recherche Scientifique, Claude Bernard University, Lyon, France (Dr Duru); and Department of Biophysics, Pierre-Marie Curie University, Paris, France (Dr Berger).

**Correspondence:** Ariel Beresniak, MD, MPH, PhD, Data Mining International, Route de l'Aéroport, 29-31, CP221, CH-1215, Geneva 15, Switzerland (aberesniak@datamining-international.com).

**Author Contributions:** Dr Beresniak has full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. *Study concept and design:* Beresniak, de Linares, Krueger, Talarico, Duru, and Berger. *Acquisition of data:* Beresniak, de Linares, and Krueger. *Analysis and interpretation of data:* Beresniak, Krueger, Talarico, and Tsutani. *Drafting of the manuscript:* Beresniak, Tsutani, and Berger. *Critical revision of the manuscript for important intellectual content:* Beresniak, de Linares, Krueger, Talarico, Duru, and Berger. *Statistical analysis:* Beresniak. *Obtained funding:* Beresniak and de Linares. *Administrative, technical, and material support:* de Linares, Talarico, and Tsutani. *Study supervision:* Beresniak, de Linares, Krueger, Duru, and Berger.

**Conflict of Interest Disclosures:** Ms de Linares is employed by L'Oréal Research.

**Funding/Support:** This study was supported by an unrestricted grant from L'Oréal Research.

**Additional Contributions:** The international BeautyQoL Study Group contributed to the data collection.

## REFERENCES

1. Spilker B. *Quality of Life and Pharmacoeconomics in Clinical Trials*. 2nd ed. Philadelphia, PA: Lippincott-Raven; 1996:1-11.
2. Ganz PA. Impact of quality of life outcomes on clinical practice. *Oncology (Williston Park)*. 1995;9(11)(suppl):61-65.
3. Patzer GL. Improving self-esteem by improving physical attractiveness. *J Esthet Dent*. 1997;9(1):44-46.
4. Sadick NS. The impact of cosmetic interventions on quality of life. *Dermatol Online J*. 2008;14(8):2. doi:10.1016/j.bjps.2007.01.071.
5. Korichi R, Pelle-de-Queral D, Gazano G, Aubert A. Why women use makeup: implication of psychological traits in makeup functions. *J Cosmet Sci*. 2008;59(2):127-137.
6. Korichi R, Pelle-de-Queral D, Gazano G, Aubert A. Relation between facial morphology, personality and the functions of facial make-up in women. *Int J Cosmet Sci*. 2011;33(4):338-345.
7. Tanioka M, Miyachi Y. Camouflage for vitiligo. *Dermatol Ther*. 2009;22(1):90-93.
8. Wang KY, Wang KH, Zhang ZP. Health-related quality of life and marital quality of vitiligo patients in China. *J Eur Acad Dermatol Venereol*. 2011;25(4):429-435.
9. Zip C. The impact of acne on quality of life. *Skin Therapy Lett*. 2007;12(10):7-9.
10. Hayashi N, Imori M, Yanagisawa M, Seto Y, Nagata O, Kawashima M. Make-up improves the quality of life of acne patients without aggravating acne eruptions during treatments. *Eur J Dermatol*. 2005;15(4):284-287.
11. Dunn LK, O'Neill JL, Feldman SR. Acne in adolescents: quality of life, self-esteem, mood, and psychological disorders. *Dermatol Online J*. 2011;17(1):1. [http://dermatology.cdlib.org/1701/1\\_review/1\\_10-00256/feldman.html](http://dermatology.cdlib.org/1701/1_review/1_10-00256/feldman.html). Accessed September 12 2012.
12. Holme SA, Beattie PE, Fleming CJ. Cosmetic camouflage advice improves quality of life. *Br J Dermatol*. 2002;147(5):946-949.
13. Deshayes P. Cosmetic camouflage for a better quality of life. *Ann Dermatol Venereol*. 2008;135(suppl 3):S208-S210.
14. Ware JE Jr, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36), I: conceptual framework and item selection. *Med Care*. 1992;30(6):473-483.
15. Juniper E, Guyatt G, Jaeschke R. How to develop and validate a new health-related quality of life instrument. In: Spilker B, ed. *Quality of Life and Pharmacoeconomics in Clinical Trials*. 2nd ed. Philadelphia, PA: Lippincott-Raven; 1996:49-56.
16. Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16(3):297-334.
17. March HW, Balla JR, McDonald RP. Goodness of fit indexes in confirmatory factor analysis: the effect of sample size. *Psychol Bull*. 1988;103(3):391-410.
18. Nunnally J, Bernstein I. Statistical foundations. In: Nunnally JC, Bernstein IH, eds. *Psychometric Theory*. 3rd ed. New York, NY: McGraw-Hill; 1994:31-208.
19. Iyengar SS, Lepper MR. When choice is demotivating: can one desire too much of a good thing? *J Pers Soc Psychol*. 2000;79(6):995-1006.
20. Altabe M. Ethnicity and body image: quantitative and qualitative analysis. *Int J Eat Disord*. 1998;23(2):153-159.
21. Loh LC. Racial differences influence health-related quality-of-life measurements. *Chest*. 2012;141(2):570-571, author reply 571.
22. Calvert M, Duffy H, Freemantle N, Davis R, Lip GY, Gill P. Population health status of South Asian and African-Caribbean communities in the United Kingdom. *BMC Health Serv Res*. 2012;12(1):101. doi:10.1186/1472-6963-12-101.
23. United Nations Educational, Scientific, and Cultural Organization. *EFA Flagship Initiatives: Multi-partner Collaborative Mechanisms in Support of EFA Goals*. Paris, France: UNESCO; 2004:51. <http://unesdoc.unesco.org/images/0013/001356/135639e.pdf>. Accessed September 1, 2012.
24. Dyble T, Ashton J. Use of emollients in the treatment of dry skin conditions. *Br J Community Nurs*. 2011;16(5):214, 216, 218 passim.
25. Xhaufilaire-Uhoda E, Haubrechts C, Piérard-Franchimont C, Piérard GE. Quality of life, emollients and hydrating agents. *Rev Med Liege*. 2006;61(4):233-236.

## シンポジウム「臨床薬理のアウトカム」

## よい雑誌とよい論文：臨床試験登録と CONSORT 声明

Good journal and good article: clinical trial registry and CONSORT statement

津谷喜一郎

Kiichiro Tsutani

東京大学大学院薬学系研究科医薬政策学

Department of Drug Policy and Management, Graduate School of Pharmaceutical Sciences, The University of Tokyo

## はじめに

先ほど、中野重行先生が「阿蘇九重カンファレンス30回の歩み：臨床薬理学の過去、現在、未来」と題してこの研究会の30年の歴史とともに、その前史とも言うべき臨床薬理学そのものについても述べられました\*1。私が東京医科歯科大学にいた時の恩師の佐久間昭先生のことについても触れられました。それを聴いて私も臨床薬理学に関連する研究デザイン、その中でもランダム化比較試験 (randomized controlled trial: RCT) の歴史について若干お話ししたいと思います。

通常、教科書的には世界で最初の人を対象としたRCTは、英国の streptomycin trial とされます。最初の患者のリクルートは1946年に始まり、1948年10月30日の *British Medical Journal* に報告されました。私は、世界で最初には日本の731部隊が旧満州で、細菌兵器開発のために行ったものではないかと思っています。実験は1941年に行われたようです。彼らはすでに Fisher RA の *Statistical methods for research workers, 3rd ed. (1930)* や、*The design of experiment (1935)* を読んでいたと思われる。2005年に日本医史学会学術総会で報告しました\*2。

## 1. The Global Health Library

さて、WHO は2004年から The Global Health Library (GHL) というプログラムを始めました。“Universal access to reliable health sciences information” がキーとなるコンセプトです。「誰でもアクセスでき信頼できる健康科学情報」。ここで「信頼できる」は「エビデンスのグレードが高い」というとやや難しいです。Fig. 1は、私がWHOで勤務していた頃にラオスで会った「伝統医学書」を持つ僧医の写真です。ラオスはインド系の医学ですが、パーリ語で書いてあります。こ

Fig. 1 パーリ語の「伝統医学書」を持つラオスの僧医



\*1 中野重行. 阿蘇九重カンファレンス30回の歩み：臨床薬理学の過去、現在、未来. 臨床評価. 2012; 39(3): 442-50.

\*2 津谷喜一郎. 世界初の人を用いたランダム化比較試験は七三一部隊によるか?. 日本医史学雑誌. 2005; 51(2): 278-9.

れが“reliable health sciences information”なのかは議論があるところでしょう。

このGHLはvirtual libraryです。知識(knowledge)はあるけれども必要とする人がそれにアクセスできない、健康と医療体制の改善は貧困に対するantidote(解毒剤)であるということで、WHOは様々な試みに着手しました。玉石混淆のインターネットホームページで「玉」を見つけてWHOロゴをつける、は誰でも考えることです。しかしこれはうまくいきませんでした。手間が大変でコストもかかる。GHLで実際に動いたのは2つあります。HINARIとGIMです。

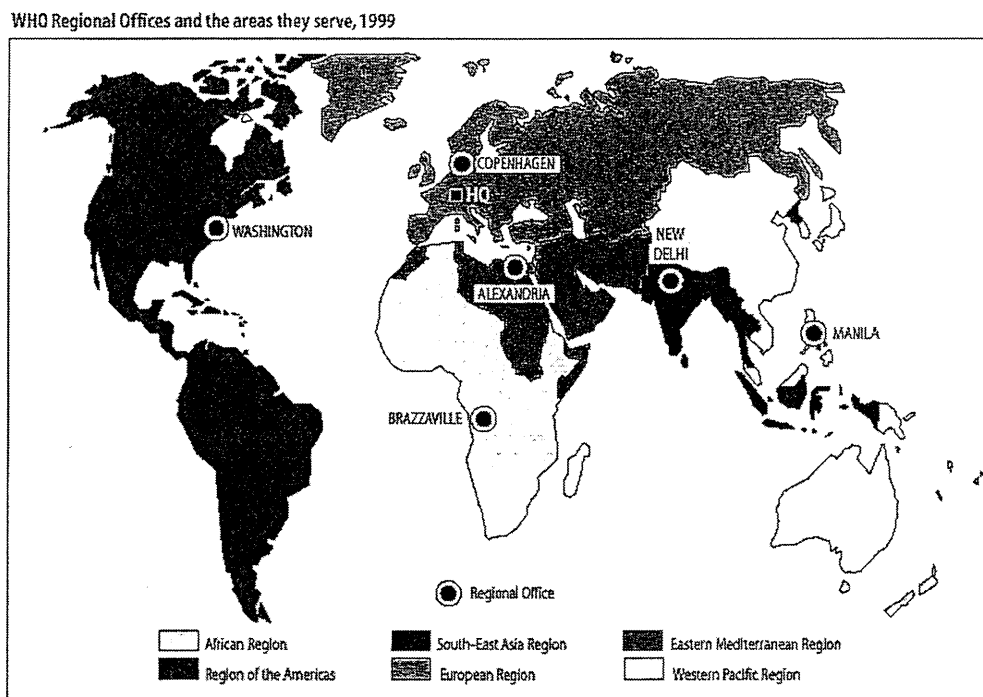
HINARI (Health InterNetwork Access to Research Initiative)<sup>\*3</sup>は、最貧国の研究所の医学関係者や研究者らが世界的な医学雑誌に無料アクセスできるシステムで、Elsevier, Wiley, Springerなど世界的な出版社の連盟やその他の雑誌発行機関とWHOが協議してつくったシステムです。日本に近いところではベトナム、ラオス、カンボジアなどが参加しています。朝鮮民主主義人民共和国

(北朝鮮)は参加していないようです。最初だけ施設としての登録が要るようですが、皆さんもこれらの国で仕事をする場合にはアクセスできます。

つぎのGIM (Global Index Medicus)で言うIndex Medicus (IM)は、昔は医学図書館に行くと電話帳のような冊子で本棚にあったものです。現在のMedline, PubMedの元祖のようなものです。

このGIMはWHOのジュネーブ本部でつくっているのではありません。Fig. 2に示すようにWHOは地域主義で、南北アメリカ、アフリカ、東地中海、ヨーロッパ、南東アジア、西太平洋、と6つの地域からなり、それぞれ地域事務局があり、そこで別々につくっています。この図を見てお分かりのように、北朝鮮は韓国と一緒に座るのが嫌で、ニューデリーに事務局のある南東アジア地域に属しています。昔はモンゴルもそうでしたが、ソビエト連邦崩壊後マニラに事務局のある西太平洋地域に入りました。インドの隣のパキスタンは、カイロに事務局がありイスラム諸国からなる東地中海地域 (Eastern Mediterranean Region)

Fig. 2 WHO world wide



\*3 <http://www.who.int/hinari/en/>



に属しています。

このうち地域版のIMは、東地中海事務局と南北アメリカでわりと早くできました。ついでアフリカ、南東アジア、西太平洋地域でもつくろうということになりました。西太平洋地域でつくったものはWestern Pacific RegionのIMで、WPRIM (Western Pacific Region Index Medicus) と称されます。なおヨーロッパ地域事務局は財政難で図書館すら十分に機能していない状態でこの方面の活動はありません。

## 2. WPRIM

WPRIM開発の活動は2005年のクアラルンプールの会議から始まり、当初日本からは図書館関係の方が参加されました。西太平洋地域事務局 (Regional Office for the Western Pacific: WPRO) でこのプロジェクトの中心となるのは“knowledge management officer”でフランス人のCharles Rabbyです。彼は、図書館、翻訳、出版の3部門を担当しています。私は2006年3月30-31日マニラで開催された第2回の会議から関わりました。私はWPROで1984-1990年に初代の伝統医学担当医

官として勤務していました。その後、2代目は中国、3代目も中国、4代目は韓国からのスタッフです。この4代目の韓国の韓医師Choi Seung-Hoonの活動に韓国政府から資金が出ていました。Knowledge management関連だけではWPRIM開発の予算が不足しており、一方、伝統医学の情報発信が必要だとChoiも考え伝統医学の予算の一部をWPRIMに回すこととなり、その関係で私にも参加するよう話が回って来たのです。

これに参加しすぐに気がついたのは、このプロジェクトに対応する明確な母体が日本に存在しないことです。日本には国立医学図書館やそれに類する機関がありません。そこで、国立情報学研究所や医学中央雑誌刊行会 (医中誌) などを訪ね議論を重ねました。WPRIM国内委員会 (WPRIM Japan: WPRIMJ) は、Table 1の方々に参加していただき、2006年11月21日に設立しました。

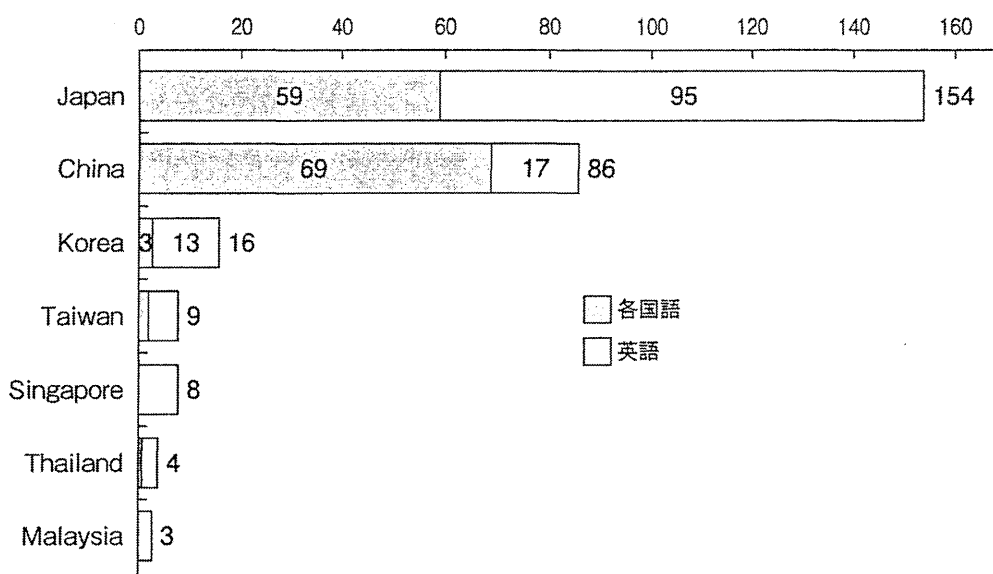
2006年当時、Medlineに入っている日本からの雑誌は154でした (Fig. 3)。これは漢字文化圏諸国からだけを集めたものですが、日本はアジアでは最大です。現在は180を超えています。また日本の医学雑誌数は約2,500ですが、これもアジアでは一番多いです。

Table 1 WPRIMJ 構成

2008.5.20 現在

＜委員＞abc順	
北川 正路	東京慈恵会医科大学 学術情報センター
北村 聖	東京大学医学教育国際協力研究センター教授 「臨床病理」編集長
根岸 正光 (副委員長)	国立情報学研究所教授
津谷 喜一郎 (委員長)	東京大学大学院薬学系研究科医薬政策学特任教授
＜顧問＞	
高久 史麿	日本医学会会長
＜オブザーバー＞	
木内 貴弘	東京大学医学部附属病院 大学病院医療情報ネットワーク研究センター教授
宮川 謹至	科学技術振興機構
城井 康信	(株)メテオ
＜サポートスタッフ (国立情報学研究所 学術基盤推進部)＞	
阿蘊品 治夫, 尾城 孝一, 上村 順一	

Fig. 3 Serials indexed in Medline from Han-character-using country/area (2006)



一方、米国国立図書館 (National Library of Medicine : NLM) の Medline には約 5,500 の雑誌が入っているとされますが、約半分がアメリカの雑誌です。キューバの雑誌は入っていません。敵国だからです。NLM の前身は US Army's Medical Library でした。冒頭に 731 部隊のことを話しましたが、そこと関係のあった日本の陸軍軍医学校のアメリカ版が発展したものです。医学図書館は国益が関係します。ベトナムは現在はアメリカとは国交がありますが、まだ Medline には入っていません。

WHO はそうした政治的なものを排除し、よりグローバルで、かつ local knowledge を伝統医学の雑誌も含めて GIM に入れることを考えたわけです。

日本医学会には 2006 年に 102 の分科会があり英文誌も入れて 158 誌が刊行されていました。それらは 4 つのタイプに分けられました。Table 2 に示します。WPRIM に日本から入るべき雑誌としては、すでに Medline に載って PubMed で無料で見ることができるものは不要と考え、また海外からの原論文のアクセス可能なものと考えました。海外から「あの論文を送ってくれ」と言われてもそれに応ずる人的資源がないためです。こう

して Type C の医中誌 (JAMAS) に掲載され full-text が web で見られるものから始めることになりました。ここでは、full-text が CiNii (Citation Information by National Institute of Informatics) からのもので分類しました。現在では J-STAGE (科学技術情報発信・流通総合システム) に掲載されているものも対象にしています。

同年 12 月 11-12 日には北京で大きな会議がありました。この時にはジュネーブの本部から図書館長代行でフランス系カナダ人の Yvonne Grandbois が参加しました。ブラジル人の Abel Packer は、アメリカ地域事務所 (American Regional Office

Table 2 Four types of 158 journals published by 102 societies under JAMS (日本医学会) in 2006

Type A	Communication journals	6
Type B	Indexed by JAMAS (医中誌) (some journals are available at J-STAGE or Medical Online Co.)	49
Type C	Indexed by JAMAS and full-text is available at CiNii	16
Type D	Indexed by Medline	87
Total		158



for Americas : AMRO) の図書館長で大変優秀な人です。このAMROは歴史的理由からPan American Health Organization (PAHO) とも呼ばれています。この会議が開催された場所は中国中医科学院 (China Academy of Chinese Medical Sciences : CACMS) でした。言わば中国の漢方医学である中医学の研究院です。この科学院 (academy) というのはソ連スタイルのシステムでスタッフは教育の義務はなく研究に専念するというものです。建物もソ連の援助で作られたものです。会議の中国側のホストは、この科学院の中にある医学情報研究所 (Institute of Medical Information) でした。そこの所長を長く務めた、中国の中医学関係の医学情報学の母と呼ばれるXue Qinglu (薛清录) もこの会議に参加していました。

西洋医学の中国医学科学院 (Chinese Academy of Medical Sciences : CAMS) の医学情報研究所からも3人の参加者がいました。彼女らはRichard Freyのことを知っていました。Frey (中国名: 傅莱) はオーストリア・ウィーン生まれ、欧米からの外国人として中国共産党に入り、毛沢東の中国革命に参加した人で、晩年この医学情報研究所の名誉所長、また北京協和医科大学図書館の名誉館長をしており、ときどきWPROでの会議に参加していました。

私はWHO勤務時この双方の科学院に予算を付けて若手スタッフをWHO fellowとして米国のNLMに送り、MedlineやMeSHなどを学んでもらったことがあり、懐かしい顔ぶれです。

この会議での中国の発表スライドは興味深いものでした。中国は雑誌の数が9,488誌と細かく分かるのです。雑誌は国に登録をしないといけないからでしょう。そのうち1,268誌 (13%) が医学領域です。そのうち94誌がMedlineに入っています。またこの9,488誌には台湾と香港の281誌がカウントされています。勝手にカウントされて台湾の人は困るのではないかと思います…。中医学は160誌 (12%) です。

WPRIMの会議は、1年に1回か2回、WPROのあるマニラか加盟国で持ち回りで開催しており、翌年2007年5月17-18日には東京で開催しました。中国などに比べると日本ではマンパワーが不足しています。このような時にはwebsiteを作ったほうがよいと考え、会議の前に公開しました\*4。

### 3. WPRIM収載誌の選択基準

東京会議ではWPRIMの選択基準が決まりました。1) 医学分野、2) 査読システムを持つ、3) 英文抄録を持つ、4) 年2回以上発行している、5) 投稿規程がしっかりしている。日本だと当然ですが、途上国では年に2回発行できない雑誌もあります。日本の学会雑誌はほとんどこの基準を満たします。しかし、Medlineに入り損ねた雑誌がここに来ると二流の雑誌の集まりになってしまうのでは、という懸念があります。

各国では追加のクライテリアを決められることになりました。そこで最近の世界的な動向を含めた日本の追加のクライテリアとして、1) RCT論文はCONSORT声明を使うこと、2) 臨床試験は登録をする、3) 全文がインターネットで見られることと、4) また当初は、日本医学会分科会の雑誌から収載することとしました。1)と2)は投稿規程に入っていることが必要です。将来は、薬学、歯学、看護などの領域や、商業誌へと増やしていこうと思っています。

また東京会議の後「信頼できる医学情報の共有—医療専門家から消費者まで—」(Sharing Reliable Health Information: From Health Professional to Consumer) と題する一般向けのシンポジウムを5月19日に東京で開催しました。ここで日本にも医学雑誌編集者の組織が必要ではないかと議論されました。

臨床試験登録制度ですが、日本では3つの登録サイトがあります。2010年2月までそれぞれに登録されている臨床試験数は、UMIN (University

\*4 <http://wprimj.umin.jp/>

Hospital Medical Information Network : 大学病院医療情報ネットワーク)に約3,000, JAPIC (Japan Pharmaceutical Information Center : 日本医薬情報センター)と日本医師会によるサイトを合わせると約4,200ほどです。おそらく現在は5,000を超えているでしょう。

臨床試験登録の必要性は、第1に出版バイアスを避けることです。例えば同じリサーチクエスチョンに対し世界で15の臨床試験があり、うち10本は効いて安全であった、5本は効かないあるいは危険であった、とすると、真の姿は15戦10勝5敗なのに、効いて安全な10本だけを発表すると、10戦10勝0敗に見えてしまいます。

第2は倫理的な問題です。臨床試験は将来の患者のために行われるものなので、結果はポジティブであろうがネガティブであろうがそれは社会に還元されるべきものです。特に危険な結果、副作用情報を公表しないと、似たような臨床試験が行われる可能性があります。

第3は臨床試験へのアクセスです。日本は医療法の関係で実施医療機関名を掲示できません。米国はClinicalTrials.govというシステムをNLMが運用していますが、エイズなどの生命を脅かす病気の患者が臨床試験に参加するには、実施場所の情報公開が必要ということで、FDAからNLMに話が行ってできたのが発端です。

日本の3つの臨床試験登録サイトを串刺しにしたものを国立保健医療科学院がつくっています\*5。これがJapan Primary Registries Network (JPRN)として、WHO International Clinical Trials Registry Platform (ICTRP)に組み込まれています\*6。当初、JAPICからのものはWHOの基準に満たしていませんでしたが改善されこのJPRNに参加しました。ただこの経緯の情報公開がやや不十分と思われる。

#### 4. CONSORT 声明

CONSORT 声明は2010年の第3版が最新で、全部で25項目のチェックリストです\*7。例えば、「タイトル」にはランダム化比較試験であると記載する。原稿の第1ページなので、チェックリストの表の第1項の右端のページ欄に1と書く。「参加者の適格基準」が原稿の4ページなら4と書く。このように25項目を記載して、それを原稿のトップにつけて編集部へ送れば、編集部がざっと見て、かつそれを査読者に送れば、RCT論文の質を保証するポイントをおさえられるというものです。

2010年の第3版で増えたのはつぎの3つです。第1に、臨床試験の登録番号と試験の登録名。第2に、可能ならば、完全なプロトコルの入手方法。これは臨床試験のデータを解析したところ意味のある差が出なくて他のエンドポイントを探してそこで差を出そうとすることがあるためです。論文にはあたかもそれが最初からのプライマリエンドポイントであるかのごとく書くということがあります。このため当初のプロトコルが入手できるようにしてバイアスを避けようというものです。

第3は、資金提供者と他の支援者(薬剤の供給者など)、試験提供者の役割、です。企業資金のものは、より効くほうに結論がなっているという研究結果に基づき、この項が追加されました。

CONSORTはこの表と一緒にフローチャートがあります\*7。組入れ(enrolment)、割振り(allocation)、追跡(follow-up)、解析(analysis)の順番で、対象とした症例数、またドロップアウトした理由を明記する。これもバイアスを避けるための方法です。

CONSORTは1996年に第1版ができて、2001年に第2版、2010年に第3版になりましたが、こ

\*5 <http://rctportal.niph.go.jp/>

\*6 <http://www.who.int/ictpr/>

\*7 CONSORT 声明については本誌649-57頁投稿規定を参照。

れによって実際に論文の質が上がり、臨床試験の質も上がったということで、他の研究デザインについても対応する報告様式のガイドライン、声明が出てきました。診断はSTARD、システムティック・レビューとメタアナリシスはQUOROM、これは昨年(2010年)PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)に変わりましたが、観察研究はSTROBE、そのメタアナリシスはMOOSE、などです。ICMJE (The International Committee of Medical Journal Editors: 国際医学雑誌編集者会議)による統一投稿規程がUniform requirements for manuscripts (URM)と呼ばれるものですが、その2006年版にはこれらがリストされていましたが、最新の2010年版にはそのリストはなくなってCONSORTだけが書いてあり、雑誌が研究デザインに従って適切なステートメントを使うようになっていきます。2008年に京都大学の中山健夫先生と一緒にこれらを翻訳して本にまとめました\*8。

日本の医学雑誌のfull-textを収載しているプラットフォームで主なものは3つあります。先に述べたCiNiiとJ-STAGEは無料で、もう1つのメディカルオンラインは有料です。他に学会が自分のプラットフォームで公開しているものもありますし、商業出版社が持っているものもあります。WPRIMJとしてはCiNiiとJ-STAGEに収載された分からはじめていますが、将来は他のプラットフォームについても取り込む計画です。課金のシステムを作る必要がありますが、WHOの他の地域局のシステムでも課金のシステムを持っていることもあります。各地域事務局のIMが統合されたGIMのwebsiteも立ち上がっています。

## 5. APAMEの設立

WPRIMの関係者の作業は雑誌の選択とデータベースシステムの運用です。この議論の中から雑

誌選択も大事だけれども、雑誌の質そのものが上がらないといけないという考えが出てきました。そのためには各国の関係者を集めたネットワークが必要ということでいろいろ議論し、Asia Pacific Association of Medical Journal Editors (APAME)が、2008年5月5日にソウルで設立されました。

キーとなる人物を何人か紹介します。APAMEの初代会長はChang-Kok Hahnで、ソウルの漢陽(ハンヤン)大学の放射線科の教授です。研究倫理や出版倫理に詳しいです。韓国には、Korean Association of Medical Journal Editors (KAMJE)という組織が1996年からあります。日本はこの組織をまったく知らなかったのですが、WPRIMの会議に参加するようになり分かりました。HahnはKAMJEに長くかかわってきた人です。

副会長は、マレーシアのMelaka Manipal Medical Collegeの学部長のJohn Arokiasamyです。その他理事として10ヶ国から入っており、主だったところでは、中華医学会雑誌社社長のWang De (王徳)、中国では医学雑誌の冒頭に「中華」が付くのが一番上のクラスということで約80誌あります。それらはすべて彼の出版社が発行しています。フィリピン大学医学部病理の教授のJoselito Mario C Avilaは、フィリピンはMedlineに1つも雑誌が収載されていませんので、学部長から何とかフィリピンの雑誌を入れろとしばしば言われているとのことです。以前は看護の雑誌が入っていたことはあるそうです。モンゴルの前厚生大臣のPagbajabyn Nymadawa、シンガポール大学のWilfred CG Peh、ソウル大学医学部学部長Jung-Gi Im、日本からは東京大学医学教育国際協力研究センター主任教授の北村聖先生が理事として参加しています。

APAMEの事務局長として中心的な活動をしているのが、ソウル大学医学部病理学教授のJeong-Wook Seoです。ソウル大学の医学図書館長もされています。穏やかな人柄ですが組織力と実行力に富む人です。

\*8 中山健夫, 津谷喜一郎, 編著. 臨床研究と疫学研究のための国際ルール集. 東京: ライフサイエンス出版; 2008.

## 6. JAMJE (日本医学雑誌編集者会議)

日本はAPAMEの設立メンバーでもあります。KAMJEのことを知ったこともあり、日本でも日本医学雑誌編集者の会をつくろうという話になりました。いろいろ議論し、やはり日本医学分科会が発行している雑誌が出発点だということで、日本医学会にその組織を置くことを日本医学会会長高久史麿先生にお願いし、2008年8月1日に日本医学雑誌編集者会議 (Japanese Association of Medical Journal Editors: JAMJE) の設立総会が開催されました。北村先生が、JAMJE組織委員会の委員長になりました。設立総会には日本医学会副会長の矢崎義雄先生に参加していただき、その後第1回目のシンポジウムを開催しました。私は先ほどのGHLやWPRIMのことを話しました。ここでのスライドなどは第2回以後を含めてすべてJAMJEのwebsite<sup>\*9</sup>で見ることができます。

第2回シンポジウムは2009年7月17日に、「出版倫理—Publish or Perish」というテーマで開催されました。山崎茂明先生がこのタイトルの本を書かれておられ<sup>\*10</sup>、その内容を話していただきました。多重出版 (duplicate publication) について私が話しました。北村先生は出版倫理の教育について話されましたが、日本における最近のミスコンダクトを列記したスライドを提示されました。2004年理化学研究所でデータ改ざんがあり2人辞職、2005年大阪大学医学部では学生による捏造があり2人停職、2005年京都大学農学部では助手の貢献で教授停職、2005年東京大学工学部ではデータ捏造疑惑、教授停職、2006年神戸大学工学部で特許申請捏造で特許取り下げ、2006年山形大学医学部ではデータ捏造で教授辞職、2006年大阪大学生命機能でデータ改ざん、教授解雇、告発者が自殺しています。2007年鹿児島

大学医学部ではデータ改ざん、助教が自殺しています。この件と、2008年の東北大学農学部、東京大学医科学研究所のデータ捏造はまだ処分がはっきりしません。若い何人かが自殺をしていますので、やはりこういった人を守らなければいけない。Publication ethicsの教育は人助けの意味もあることになります。ミスコンダクトを予防するためのロールプレイの医学教育の紹介をされました。

第3回シンポジウムは2010年7月15日に開催されました。JAMJEは5人の組織委員から成り立ち私も入っており、シンポジウムのテーマなどを決めています。そろそろCOI (conflict of interest: 利益相反) をとということになり、「臨床研究成果の発表に係る利益相反管理のあり方」をテーマとして日本医学会のCOI委員会と共催で開催されました。日本医学会は現在108の分科会があります。煩惱の数だけあると言われていますが、分科会横断型の活動としては医学用語委員会が戦前から存在します。108の分科会は大きく、基礎、社会、臨床の3つの部会に分かれています。この臨床部会にCOI委員会があり、COIのガイドライン案を作成していましたので、このシンポジウムで、COI委員長で、徳島大学臨床試験センター長も務められた曾根三郎先生が発表されました。また旭川医科大学の高後裕先生が国内学会のCOI管理の現状調査の結果を発表されました。

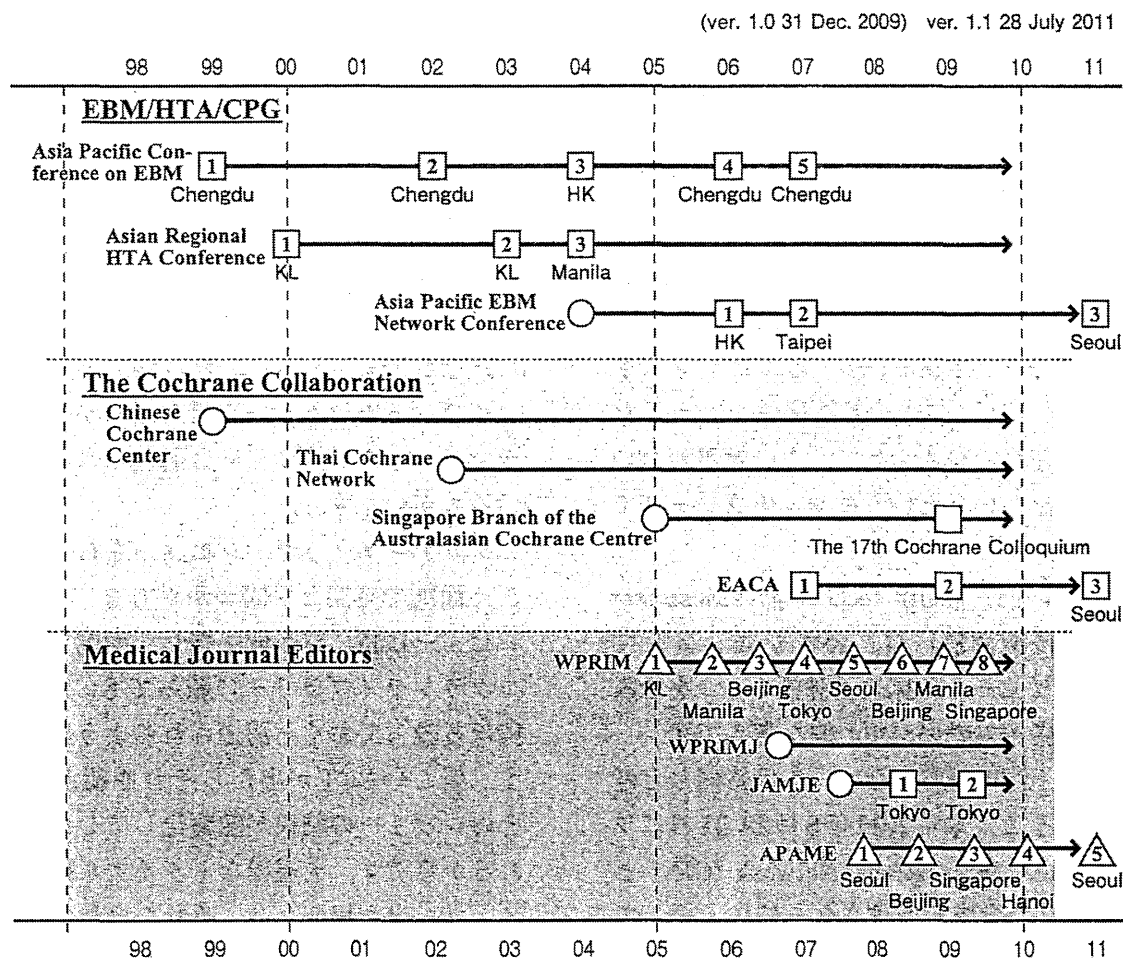
## 7. アジアのEBM関連の動き

Fig. 4はアジアのEBM (evidence-based medicine) 関連の動きを示したものです。一番上のAsia Pacific Conference on EBMは1999年に中国の成都で始まり2-3年に一度開催されてきているものです。2004年に1回だけ香港で開催されました。第2の流れのAsian Regional HTA (Health Technology Assessment) Conferenceは、2000年

\*9 <http://jams.med.or.jp/jamje/>

\*10 山崎茂明. パブリッシュ・オア・ペリッシュ 科学者の発表倫理. 東京: みすず書房; 2007.

Fig. 4 APAME/WPRIM and EBM/CPG related activities in Asia-Pacific



©Tsutani K, Kitagawa M

と2003年にマレーシアのクアラルンプールで開催され、第3回がマニラで開催されました。これが先の第3回の香港でのAsia Pacific Conference on EBMと同じ週となっていました。私を含めて何人かは双方の大会に参加しました。

私はWHOに勤務したこともありこの種の調整には慣れていましたので、香港側の主催者の香港中文大学のTang Jinglinと相談し、その会に参加しているアジア各国からの人を集めて、Asian-Pacific EBM Network (APEBMN) を設立しました。Tangは北京大学EBMセンター長も併任していますが、彼は2006年には、Asia Pacific EBM Network Conferenceを香港で開催しました。翌2007年は台北で開催されました。

これらの会議では、論文の批判的吟味、システマティック・レビュー、診療ガイドラインなどが報告・議論されていました。

第2の流れは、コクラン共同計画(The Cochrane Collaboration) 関連です。コクラン・レビューはあるリサーチクエスチョンについて、システマティック・レビューにより世界中からRCTを集めてプールした結論を出すというもので、エビデンスを「つたえ」るものです。1999年に中国コクランセンターが成都に開設され、またThai Cochrane Networkや、オーストラリア・メルボルンに事務所のあるAustralasian Cochrane CentreのSingapore Branchができています。毎年開催されるコクランコロキウム(The Cochrane Colloquium)はこの

Singapore Branchが主となり2009年にアジアで最初にシンガポールで開催されました。

また2007年にAsia Pacific EBM Network Conferenceが台北で開催された際には、East Asian Cochrane Alliance (EACA) が設立されました。名称案はいくつかありましたが、日本の寿司の「イカ」「タコ」のイカ (EACA) だと冗談で言ったら、それが覚えやすいと、EACAに決まりました。

3つ目に、先にお話しましたWPRIMとAPAMEの流れがあります。EBMで言うところのエビデンスを「つたえ」るメディアの中の雑誌関係ということになります。雑誌の質も上げて「よい雑誌がない」とよい論文が出ない」というキャッチフレーズで動いています。

最初の流れの内、health technology assessment (HTA, 医療技術評価) は、大きくはシステマティック・レビューと臨床経済評価の2つを含みます。後者は費用対効果が中心的な考えです。

すなわち、上記の最終的なゴールは皆同じと考えられます。そこで本年2011年8月28-31日にソウルでGuidelines International Network (GIN) という4日間ほどの会議が開催される時に、そこでこの3つの関係者が集まり議論する予定です。

#### <質疑応答>

座長 (野元) 文献を集めてレビューするグ

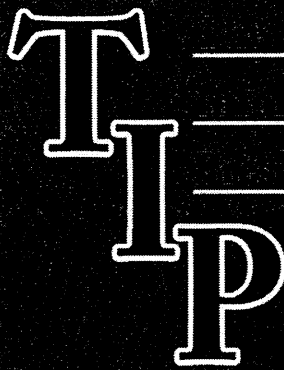
ループに対する評価はありますか。

津谷 コクラン共同計画は数万人が関わっているとされますが、世界的に高い評価があります。1993年に設立されました。しかしシステマティック・レビューはメタアナリシスとほぼ同義と違ってよいのですが、当初、日本では「そのようなものは研究ではない」と言われました。現在、コクラン・レビューのインパクトファクターは5ぐらいで通常の医学雑誌よりもはるかに高くなりました。それでも今でも、あれは他人のスタディをまとめているだけで、一次研究だけが研究だ、システマティック・レビューは研究ではないと思っている人がいます。

システマティック・レビューを行ってみるとすごく面倒で時間と手間がかかります。コクラン・レビューは特に質管理が厳しく面倒で、完成してコクランライブラリーに載っても、現在では3年間ほど放っておくと消去されるようになっています。新しいRCTが出たらそれを組み込まないとそのレビューの結果が陳腐化してしまうとみなされるからです。大規模なトライアルの結果がネガティブと出た場合など、それを含めてレビューすると全体で「効かない」とそれまでのレビューの結果が逆転する可能性もあります。

というわけで、コクラン・レビューではなく、自分であるいはグループで別にシステマティック・レビューを行い発表する人もいます。

\* \* \*




医薬品・治療研究会

# 正しい治療と薬の情報

*Critical Choices in Drugs and Therapeutic Alternatives*
**May, 2012**  
**Vol.27 No.5**


## シリーズ「医療の近接領域および社会科学における EBP の動向」 連載開始に当たって

 原田隆之<sup>1)</sup> 津谷喜一郎<sup>2)</sup>

- 1) 目白大学人間学部心理カウンセリング学科
- 2) 東京大学大学院薬学系研究科医薬政策学

医療やヘルスケアの分野においては、「エビデンスに基づいた医療」(evidence-based medicine: EBM)への動きが日本では、1990年代中頃から始まり、今世紀になって主流になりつつある。だが、それ以外の領域ではどうであろうか。例えば、近接領域である心理学においては、「エビデンスに基づいた実践」(evidence-based practice: EBP)、あるいは「エビデンスに基づいた意思決定」(evidence-based decision-making)は、ほとんど浸透していない。それどころか、あからさまな「エビデンス嫌い」や、EBPに対する初歩的な誤解もめずらしいことではない。

例えば、心理学領域で最大の会員数を誇る心理臨床学会の学会誌の巻頭言では、「臨床の神髄は昨今流行りの evidence based なり数量化したり統計に叶ったりという低いレベルのものでは、到底測りがたいものが多い」<sup>1)</sup>と述べ、このような「風潮」は「断じてゆるしがたい」と非常に強い調子で EBP を批判している。

また、犯罪心理学会の EBP に関するシンポジウムでは、あるシンポ

ジストが「エビデンス論は、犯罪臨床全体のパワーアップを目的とし(筆者注:「目的とすべきであり」の意か)、『有効・無用』の議論が、特定のアプローチを批判するために用いられることは、厳に慎むべきであろう」<sup>2)</sup>との批判を展開している。

このような例は枚挙に暇がないのが実情である。前者について言えば、「人間の心は複雑で微妙なものであるから、数字や統計では到底わからない」という的外れな批判をよく聞くが、これと軌を一にするものであろう。エビデンスを重視するからと言って、何も数字だけで人間の心がわかると思っているわけではないのは当然である。さらに言えば、エビデンスやデータに拠らないで、何に基づいて「高いレベル」の臨床をするというのだろうか。

おそらくは、自らの臨床経験、解釈という名の主観的判断、直観、好みなどでアセスメントをしたり、意思決定をしたりするのであろう(これが心理臨床の「スタンダード」だからだ)。例えば、子供に絵を描かせて虐待の有無を判断したり、インクの染みを見せて性格診断をしたりするなどの実践は、日常的に行われて

いるが、残念ながらこれらを支持するエビデンスはない。しかし、それらが正しいとア priori に信じ込んでおり、「エビデンスがない」という批判を受けようものなら、感情的な反発が噴出する。例に挙げたシンポジストの発言はまさにその典型である。しかし、アカデミズムにおいて、「エビデンスによって批判をするな」という態度はきわめて危険であることは言うまでもない。

さらに、EBP に違和感を持つ者たちは、EBP へのアンチテーゼとして narrative-based practice (NBP) を持ち出すのであるが、それは EBP と対立するものではない。なぜなら、EBP とは何もエビデンス至上主義ではなく、EBM の世界的な教科書の中で、Sacket et al.<sup>3)</sup> は、エビデンスを専門家の臨床技能やその場の状況、そして患者の好みや価値観と統合して用いることの重要性を説いており、これはとりもなおさず患者の narrative を重視すべきとの態度にほかならないからである。

このように、EBM や EBP に対する基本的な知識すら欠いた反論があまりに多いことに対して、筆者らは危惧を抱いている。まず、何よりもエビデンスを欠いた実践によって、クライアントに害をなしてしまう、あるいは効果のない対処をしてしまうという点である。例えば、震災後「こころのケア」と称して大勢の臨床心理士が被災地を訪れているが、その際、不用意に子供に被災体験の絵を描かせ、かえって心情不安定にさ



せてしまったということが大きく新聞でも報じられた(朝日新聞, 2011年6月10日付朝刊)<sup>3</sup>。トラウマ・ケアにおいて, まだ体験後間もないうちに, その体験を掘り起こすようなことをするのは禁忌であるが, それを知らずに自らの日常的臨床経験を元にして, 絵を描かせるのがよいという意味決定をしたのであろう。言うまでもなく, いくら善意からのこととは言っても, 専門家を標榜する者としてその責任は大きい。

さらに, もう1つ危惧することは, エビデンスを無視または軽視した実践が, 心理学というディシプリンの発展を阻害し, この学問領域への信頼を大きく損ねてしまうということである。健全な批判精神なくして学問の発展は望めない(先述のように, 「エビデンスによって批判するな」という意見が危険である所以である)。どれだけ権威者の説であっても, どれだけ長年の定説であっても, あるいは長年馴染んできた実践であっても, 実証的な支持が得られなければ, その事実を認めて受け入れるだけの寛容さが必須である。学問は自らの満足や誇りのためにあるのではない。

長々と心理学領域におけるEBPについて述べたが, このほか, 社会学, 教育学, 経済学, 政治学など

の社会科学, あるいは代替医療などの領域でも, 分野による差はあれ, EBPがまだ十分に浸透していなかったり, EBPへの反発が見られたりすることが現状であろう。その一方で, これらの分野にEBPを浸透させていこうという動きも見られる。欧米では, コクラン・レビューの社会科学版であるキャンベル・レビューが, 数々のシステマティック・レビューを発表し続けている。その日本語のサイトも既に公開されており, 良質のエビデンスに無料でアクセスできる(<http://fuji.u-shizuoka-ken.ac.jp/~campbell/index.html>)。

また, 筆者らのグループは, 様々な分野でのランダム化比較試験の活用を刺激すべく, 一昨年2010年に『ランダム化比較試験(RCT)の設計: ヒューマンサービス, 社会科学領域における活用のために』(日本評論社)を刊行した。また『教育研究とエビデンス: 国際的動向と日本の現状と課題』(国立教育政策研究所(編), 大槻達也, 惣脇宏, 豊浩子, トム・シュラー, 初井圭子, 津谷喜一郎, 秋山薊二, 岩崎久美子(著), 明石書店)が今年(2012年5月)に刊行され, さらに『いろいろな臨床試験のケースレポート: 温泉のRCTから看護のSRまで』(津谷喜一郎

(編), ライフサイエンス出版)も近刊の予定である。我が国の研究の動向や課題を紹介する予定である。こうした地道なアドボカシーや研究と実践を継続することが, 現在の我が国の状況を鑑みるに最も必要なことと言えるであろう。

さて, 今月から5回にわたって, 本誌においても連載企画として, 医療の近接領域や社会科学領域におけるEBPについて, その現状や課題を紹介したい。これらの領域において, どのようなエビデンスが産出され, どのように活用されているのか, あるいはどのような問題があるのかなど, その最前線の現状を各分野の専門家が解説し, その上で, 今後のEBPの発展に対して何らかの示唆を得ようとするのが目的である。貴重なご意見を頂戴することができればこの上ない喜びである。

#### 参考文献

- 1) 山中康裕. 心理臨床学研究, 28(1): 1-3, 2010
- 2) 生島浩. 犯罪心理学研究, 48(特別): 263, 2010
- 3) Sackett et al., Evidence-Based Medicine: How to Practice & Teach EBM, 2<sup>nd</sup> ed. London: Churchill Livingstone; 2000
- 4) <http://www.asahi.com/edu/news/TKY201106090698.html>

## 医療の近接領域および社会科学における EBP の動向(第1回)

### 違法薬物乱用対策のエビデンス

原田 隆之<sup>1)</sup>

1) 目白大学人間学部心理カウンセリング学科

#### 1. 違法薬物乱用の現状

覚せい剤や大麻など違法薬物の乱用は, 我が国のみならず, 世界的に大きな社会問題である。今や世界の違法薬物市場は, 3,200億ドル(約30兆円)に達しているが<sup>1)</sup>, この金額がいかほどの規模かと言えば, 世界の国々のうち, GDPがこの額を上回っている国は, わずか12%しかない<sup>2)</sup>。

我が国においては, 戦後一貫して覚せい剤が最大の違法薬物であり, ここ何十年も検挙人員は1万人を下回ったことがなく(図1), 刑務所人口の約20%を占め続けている。しかも, その再犯率は70%にも及ぶ<sup>3)</sup>。

こうした現状に対し, 我々の社会はどのような対策を取ってきたのだろうか。まず思い浮かぶのは, 「覚

せい剤やめませか, それとも人間やめませか」「ダメ, 絶対」などのキャッチ・コピーによる反薬物キャンペーンである。あるいは, 「ゼロ・トランス(不寛容)」政策と呼ばれる薬物犯罪に対する厳重な取締りである。

海外に目を向けてみても, 米国では1980年代, 増え続ける薬物犯罪に対し, 政府は「War on Drug(薬物戦争)」を宣言し, 厳罰化で臨んできた。このような対策は, 結局失敗に終わり, 薬物問題は一向に解決されなかった。オバマ大統領は, 薬物戦争政策の失敗を認め, より科学的なデータに基づいた薬物政策の必要

〈記録〉第32回 日本臨床薬理学会年会 2011年12月1~3日 浜松  
シンポジウム4: ヘルシンキ宣言の過去と現在

### 3. ヘルシンキ宣言と臨床試験登録

津 谷 喜一郎\*

#### 1. はじめに

1990年代から徐々に高まった臨床試験登録への関心は、どのような背景を持つのだろうか? ヘルシンキ宣言にはそれがどう表れてきたのであろうか? 現在の臨床試験登録の現状とヘルシンキ宣言とはどういう関係になるのであろうか?

#### 2. ヘルシンキ宣言中の臨床試験登録と出版倫理

ヘルシンキ宣言での臨床試験登録に関連する事項は、2000年のエディンバラ改訂版からである。第16項に「すべての研究計画は一般に公開されて (publicly available) いなければならない」の一文が入った (日本医師会の訳による、以下同様)。

2008年のソウル改訂版で、第19項に「すべての (every) 臨床試験は、最初の被験者を募集する前に、一般的にアクセス可能なデータベースに登録されなければならない (must)」が入った。

また第30項には「著者、編集者および発行者はすべて、研究結果の公刊 (publication) に倫理的責務を負っている。著者は人間を対象とする研究の結果を一般的に公表する (publicly available) 義務を有し、報告書の完全性と正確性に説明責任を負う。彼らは、倫理的報告に関する容認されたガイドラインを遵守すべきである。消極的 (negative) 結果および結論に達しない (inconclusive) 結果も積極的 (positive) 結果と同様に、公刊または他の方法で一般に公表されるべきである。刊行物 (publication) の中には、資金源、組織との関わりおよび利益相反が明示される必要がある。この宣言の原則に反する研究報告は、公刊のために受理されるべきではない」と出版倫理全般のうちに結果の公刊 (publication) が入った。

#### 3. なぜ臨床試験登録が必要か?

臨床試験登録制度には、出版バイアスの防止、倫理、

患者の臨床試験へのアクセス向上、の3つのニーズがあり、日本ではこの順番で記されることが多い。それぞれの歴史的背景がある。

##### 1) 出版バイアスの防止

エビデンスを「つたえる」メタアナリシスの方法は、1976年の教育心理学者の Gene V. Glass の講演が最初である。当時は学級規模と成績の関係が話題となっていた。当時から出版バイアスの可能性が議論されていたとされる。

臨床試験登録により出版バイアスを防止しようという動きは1990年代からの「エビデンスに基づく医療」(evidence-based medicine: EBM) で大きくなる。1993年には「コクラン共同計画」(The Cochrane Collaboration) が設立され、システマティック・レビューが一定の質管理とともに数多く実施されるようになり、出版バイアスに対する関心が高まった。2000年にはロンドンで臨床試験登録の会議が開催された<sup>1)</sup>。同年にエディンバラ改訂がなされた。

バイアスの入ったエビデンスは「つかう」べきではない。そのためには臨床試験登録が必要である。それによって、システマティック・レビューやメタアナリシスを行う者は、その臨床試験を実施した者に問い合わせることができる。だが実際の登録は低調であった。

世界的な関心が高まったのは、2004年6月3日に GlaxoSmithKline (GSK) による小児に対する paroxetine (Paxil®) の臨床試験で自殺企図が報告されなかったというスキャンダル記事が New York Times の一面に掲載されたのが契機である<sup>2)</sup>。同年9月には医学雑誌編集者国際委員会 (ICMJE) が臨床試験は登録していなければその論文は受理しないという声明を出し、10月のカナダでのコクラン・コロキウムでは「臨床試験登録に関するオタワ宣言」<sup>3)</sup>が発表されるなどした。

翌2005年4月にWHO本部でこのテーマの会議が開催された<sup>4,5)</sup>。6月には日本で UMIN-CTR が日本で初の臨床試験登録システムを開発し登録を開始した。

\* 東京大学大学院薬学系研究科医薬政策学  
〒113-0033 東京都文京区本郷7-3-1

## 2) 倫理

臨床試験 (clinical trial) と診療 (clinical practice) との倫理の違いは、1970 年代後半から米国で議論されるようになった。患者は公衆 (public) を構成する一員として将来の患者のために臨床試験に参加するのであるから、そこで「つくら」れたエビデンスはポジティブ、ネガティブを問わず公衆に還元されるべきであり、それは公共財 (public goods) である、との利他主義 (altruism) に基づく考えは先のオタワ声明で明確に表れている。

この間、どの臨床試験を登録すべきかについて、製薬企業と医療消費者との間で激しい論争がなされた。企業は知財を守るために早期の臨床試験登録には否定的であった。2008 年 10 月のヘルシンキ宣言ソウル改訂版は「すべて (every) 登録されなければならない」となった。

一方、国際製薬団体連合会 (International Federation of Pharmaceutical Manufacturers & Associations : IFPMA) は、米国の PhRMA、欧州の EFPIA、日本製薬工業協会 (JPMA) の参加のもと、2005 年 1 月 6 日の共同指針で、登録対象試験を「探索的試験を除く全ての試験」とした。その後、2008 年 1 月の第一次改訂で「全ての検証的試験及び全ての有効性を見る探索的試験」、2009 年 11 月の第 2 次改訂で「患者を対象とした全ての臨床試験」へと拡大してきた<sup>6)</sup>。だがこの指針は 2008 年 10 月のヘルシンキ宣言には違反した状態となっている。

また、米国 FDA は、従来、承認申請に使われる海外での臨床研究はヘルシンキ宣言に従うこととしてきたが、2008 年 10 月 27 日にこの方針をとりやめ、ICH-GCP に変えることとした<sup>7)</sup>。

## 3) 臨床試験へのアクセス向上

臨床試験登録機関としては米国の clinicaltrials.gov が世界最大の登録数をもつ。その契機は、1980 年代後半からの臨床試験中の抗エイズ薬へのアクセスの権利の議論であった。1997 年 11 月の FDA 近代化法により、生命を脅かす疾患 (life threatening disease) の臨床試験への患者のアクセスを向上するため、米国国立図書館のインターネット上で臨床試験登録公開をすることとなった。

2000 年 2 月に開始された当初のシステムは、Clinical Trials Data Bank と称され、NIH-sponsored trial が主であった。その後、登録される疾患領域や臨床試験の実施国が拡大された。また、臨床試験へのアクセス向上の動きは、その後コンパッションレート・ユース (compassionate use : CU) 制度にも発展した。

他方、日本では臨床試験を実施している病院名などの開示は医療法違反となるため、この臨床試験へのアクセス確保は大きな意味は持たない。

## 4. おわりに

臨床試験登録はエビデンスを「つたえる」局面の出版倫理 (publication ethics) の一環として捉えることができる。ヘルシンキ宣言はエビデンスを「つくる」局面の研究倫理 (research ethics) として始まったものである。EBM はインターネットなどの情報システムの発達をインフラにもつが、そこでは社会の多様な人々がエビデンスを「つかう」状況が生じた。この EBM の興隆がヘルシンキ宣言のスコープをエビデンスを「つかう」局面から「つたえる」局面へと押し広げたとも言えよう<sup>8)</sup>。ただし、ヘルシンキ宣言は、従来の権威を失っており、各セクター、各国で異なる倫理指針が用いられており、バルカン半島化<sup>7)</sup>していると言える。

## 文 献

- 金子善博, 津谷喜一郎. 臨床試験登録・公開の現況と展望—ランダム化比較試験の登録に関するロンドン会議報告. *臨床評価*. 2000; 27(3) : 491-501.
- 津谷喜一郎. 臨床試験の登録と結果の公開(ポジティブ, ネガティブを含めて)—医薬経済学の立場から. *臨床医薬*. 2005; 21(1) : 47-57.
- 津谷喜一郎, 菊田健太郎. 「臨床試験登録に関するオタワ声明」の背景・内容・今後. *薬理と治療*. 2005; 33(6) : 553-5.
- 津谷喜一郎. WHO 技術諮問会議・臨床試験の登録基準. *薬理と治療*. 2005; 33(6) : 556-9.
- 松葉尚子, 津谷喜一郎. 「WHO 技術諮問会議・臨床試験の登録基準」参加報告. *薬理と治療*. 2005; 33(6) : 560-6.
- 高見和夫. 企業にとっての臨床試験と情報公開. *臨床薬理*. 2011; 42(4) : 265-6.
- Kimmelman J, Weijer C, Meslin EM. Helsinki discords : FDA, ethics, and international drug trials. *Lancet*. 2009; 373(9657) : 13-4.
- 津谷喜一郎. よい雑誌とよい論文—臨床試験登録と CONSORT 声明. *臨床評価*. 2012; 39(3) : 475-84.

# 米国における Postmarketing Requirements と 日本における市販後の研究に関する比較と提言 —米国と日本の市販後研究の比較と日本の安全性監視計画への 提言に関するタスクフォースからの最終報告—

古閑 晃<sup>\*1</sup> 甲斐靖彦<sup>\*2</sup> 景山 茂<sup>\*3</sup> 久保田潔<sup>\*4</sup> 津谷喜一郎<sup>\*5</sup>  
西 利道<sup>\*6</sup> 前田 玲<sup>\*7</sup> 政田幹夫<sup>\*8</sup> 宮川 功<sup>\*9</sup>

**目的：**日米での市販後安全性研究を比較することにより，日本で最良の安全性監視の方法を提言する。

**方法：**2010年に日本で新たに承認された医薬品のうち，すでに米国では市販されている12品目を対象とした。まずは米国で承認された時点での薬剤に対する市販後の安全性の懸念と，それに対応して企業に研究実施が求められる Postmarketing Requirement を抽出し，日本については同じ薬剤で審査の段階で重点調査項目として議論された安全性の課題と，それに対応する市販後の安全性の研究について検討した。

**結果：**両国とも安全性の課題については共通することが多かったものの，対応する研究については，米国では安全性の懸念に応じた個別の研究であるのに対して，日本では通常の製造販売後調査や全例調査といった定型的なパターンであった。日本での理想の市販後安全性研究の提案については，比較群を必要とする研究は，学会，第三者研究機関に委ねるべきものや，全国的ながん登録の構築が望まれるものなど，企業独自の努力ではその達成に限界があるものがほとんどであった。

**結論：**日本でも2013年から医薬品リスク管理計画が実装されることとなるが，その際には国際的に通用する科学的な安全性監視計画の実施が患者の安全性確保につながる。これまでの市販後の調査を基盤として，日本でも実施可能な方策を産官学で検討していくことが喫緊の課題である。  
(薬剤疫学 2012 ; 17 (1) : 55-66.)

**Key words：**安全性監視計画，市販後調査，postmarketing requirement，治療リスク管理計画

## 1. 目的

日本では2005年に審査管理課長，安全対策課長の連名でICH E2E「医薬品安全性監視の計画」<sup>1)</sup>が通知として発出され，Step 5となった。しか

し，その「3.2 医薬品安全性監視の方法」に記載されている「特定の状況における安全性監視に取り組む最良の方法は，医薬品，適応疾患，治療対象の集団及び取り組むべき課題によって異なる。」

\*1 日本イーライリリー株式会社

\*2 参天製薬株式会社

\*3 東京慈恵会医科大学

\*4 東京大学大学院医学系研究科

\*5 東京大学大学院薬学系研究科

\*6 大原薬品工業

\*7 米国製薬研究工業協会

\*8 福井大学医学部附属病院薬剤部

\*9 武田薬品工業株式会社

別刷請求先：〒651-0086 兵庫県神戸市中央区磯上通7-1-5 日本イーライリリー株式会社安全性情報部 古閑晃

の趣旨はほとんど反映されることなく、製造販売後使用成績調査や全例調査などの定型的な調査が、いまだ大多数を占めているのが現実である。2003年以降、海外では rofecoxib や rosiglitazone の心血管系リスクが重大な安全性の懸念事項として問題化し、また、pioglitazone の長期投与後の膀胱がんのリスクなどが問題となった。心血管リスクは背景発現率が高く、薬による膀胱がんのリスクについても、背景発現率との比較は不可欠である。したがって、これらの問題の検証には、異なる投与量使用者グループ間や背景発現率との比較が必要であるが、これまでの日本で実施されてきた調査では十分な対応は困難である。

米国では2007年のFDA再生法成立に基づき、市販後に企業に対して必要があれば安全性の課題の検討を義務付けるための postmarketing requirements (PMR) が一部の製品に課せられ、さらに不順守の場合には罰則規定まで盛り込むなど、市販後の安全性監視が強化された<sup>2)</sup>。日本でも「薬害肝炎の検証及び再発防止検討会」からの最終報告における「薬剤疫学的手法の安全監視への導入：医薬品の副作用や有用性の検証のため、市販後の新しい状況下で情報の収集と評価を科学的に適切に実施するための計画策定についてのガイドライン (ICH E2E) に示された薬剤疫学的な調査手法を具体的な市販後調査に活用すべきであり、上記の医薬品のリスク管理計画においても、これらを取り入れて対策の実施状況のフィードバックがなされる仕組みが望まれる。」との提言<sup>3)</sup>に基づき、2012年の4月に医薬品リスク管理計画書指針が公表された。

日本には世界に類をみない再審査制度があり、市販後の安全対策は世界に誇るものであるとの認識が大勢を占め、市販後安全性評価に対する危機感希薄である。30年前に発足した再審査制度は、当初は十分世界に通用する内容であったと評価できるかもしれない。しかし、最近の欧米の安全性の課題に対する評価力の進歩を勘案するなら、再審査制度、特に調査、研究の内容の世界水準に照らしての再検討が必要である。

そこで日本薬剤疫学会では「米国と日本の市販

後研究の比較と、日本の安全性監視計画への提言「タスクフォース」を2011年4月に立ち上げ、計5回の会合で、直近の日本で承認を受けた新医薬品と同じ製品のうち、すでに米国で承認を受けている医薬品を複数取り上げ、市販後の追加の安全性監視計画の内容を比較することにより、日米の差を明らかにし、日本でも可能な最良の市販後の安全性監視の在り方について提言することを目的として活動を行った。

なお、本成果は2011年11月6日の第17回日本薬剤疫学会学術総会にて発表を行った。

## 2. 方 法

2010年に日本で承認された部会審議品目71品目のうち、米国にて日本と同一効能を含む薬剤で、FDAホームページで公開されているPMRが課されている12品目を抽出した。

日本における市販後の安全性監視の現状については、該当する審査報告書を、以下の医薬品医療機器総合機構ホームページのリンク先から検索して検討した。

<http://www.info.pmda.go.jp/approvalSrch/PharmacySrchInit?>

公開されている範囲の審査報告書から、市販後の安全性の課題事項として取り上げられた「重点調査項目」を抜き出すとともに、対応する市販後に予定される研究を抜き出した。

米国のPMRの内容については、2011年6月時点で公開されている以下のFDAホームページからのリンク先から検索した。

<http://www.accessdata.fda.gov/scripts/cder/pmc/index.cfm>

ただし、PMRは2007年のFDA再生法成立以降に新たに定められたものであり、それまでは postmarketing commitment として、安全性、品質、有効性を含む幅広い課題事項に対して、FDAと企業の間での市販後の研究実施の強制力を伴わない約束として実施されていたものがほとんどである。この点を考慮し、2007年以前に承認されている薬剤については、上述のリンク先においてPMRに検索を制限するためのFDAAA Section