4, although MeCP2 mutation analysis was not performed in 2 cases. As controls we used 4 brains without neuropathologic changes from individuals aged 12 to 31 years. Frozen samples of 3 RTT patients were obtained from the Harvard Brain Tissue Resource Center.

With informed consent, we performed immunohistochemistry and immunoblotting of human RTT and non-RTT brains. These 4-μm-thick paraffin-embedded specimens were immunostained with the antibodies for MeCP2 and IGFBP3. For immunoblot analysis, the primary antibodies of MeCP2 and IGFBP3 were used for all frontal cortices of RTT patients and controls.

Statistical Analysis

We used StatView software (SAS Institute Inc., Cary, NC) for to analyze significance between wild-type male and *mecp2*$^{-/y}$ male, and the Student *t*-test was performed on the data. We considered $p < 0.05$ as a significant difference.

RESULTS

MeCP2 Binding to IGFBP3 Promoter

We obtained promoter sequences of mouse and human *IGFBP3* (Fig. 1A) from the GenBank of the National Center for Biotechnology Information (accession numbers AL607124 for mouse *IGFBP3* and M35878 for human *IGFBP3*). We found many CG-rich regions within 500 bp from exon 1. A chromatin immunoprecipitation assay showed that MeCP2 bound directly to the promoter of *IGFBP3* (Fig. 1B).

IGFBP3 Expression Analyses of Mouse Brains

Reverse transcriptase-PCR showed mRNA of mouse *IGFBP3* expressed in the cerebral cortices (Fig. 1C). *IGFBP3* mRNA expression was higher in the prenatal period than that in the postnatal period. From the standpoint of

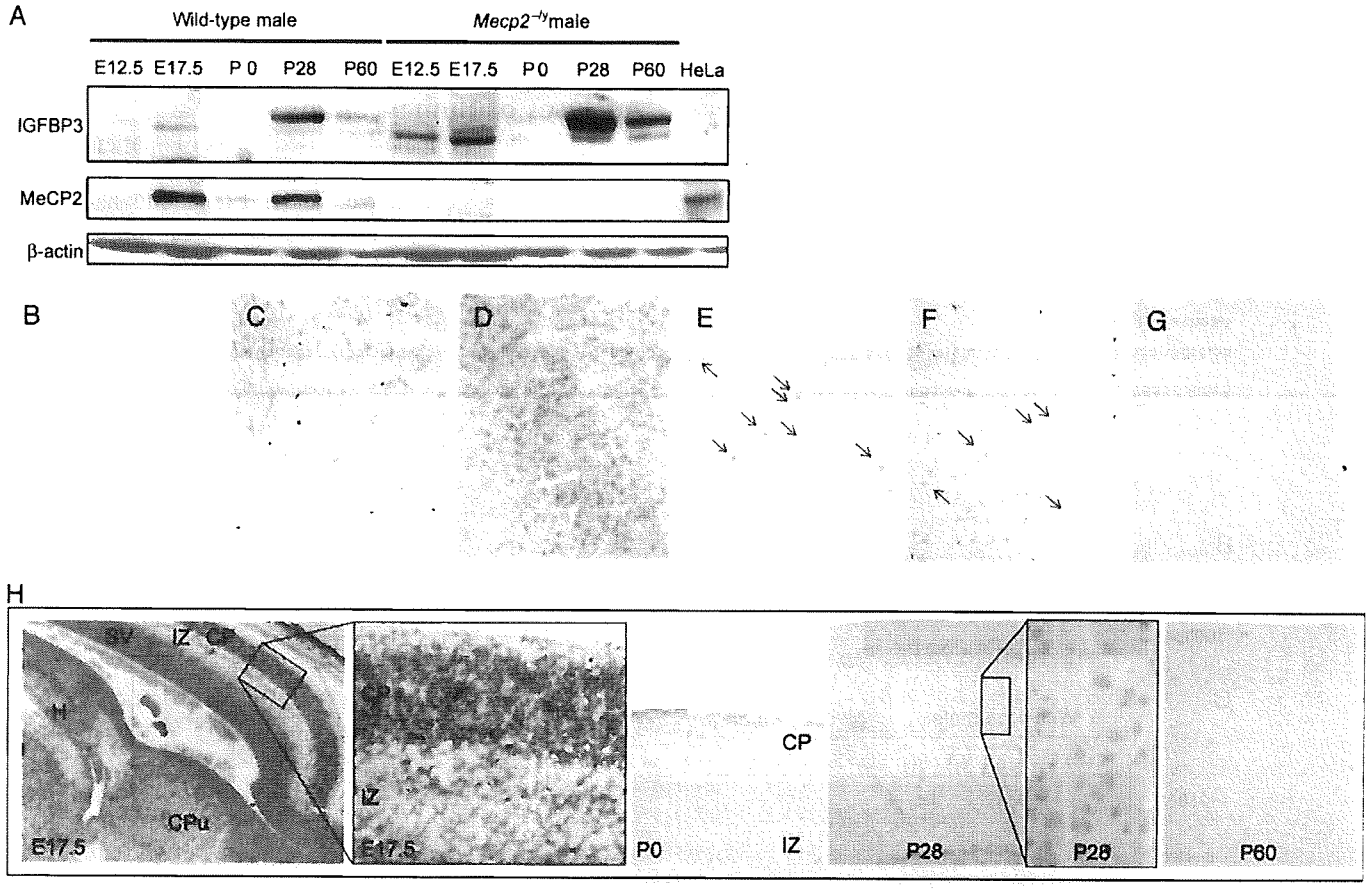

FIGURE 2. Insulin-like growth factor binding protein 3 (IGFBP3) expression analyses of mouse brains. **(A)** IGFBP3 protein expression is higher in *mecp2*$^{-/y}$ than in male wild-type mice. IGFBP3-positive cells scatter and show low signal in the frontal cortex of wild-type **(B)**. Small numbers of IGFBP3-positive cells and fibers are observed in the frontal cortex of *mecp2*$^{+/-}$ **(C)**, but numerous IGFBP3-positive cells and fibers are in the frontal cortex of *mecp2*$^{-/y}$ **(D)**. Wild-type **(E)** and *mecp2*$^{+/-}$ **(F)** mice show some MeCP2-positive cells (arrows) in the frontal cortex **(E)**; however, *mecp2*$^{-/y}$ has no MeCP2-positive cells **(G)**. **(B–G)** Age postnatal day (P) 28; Original magnification: 100×. **(H)** IGFBP3 is widely expressed in the cortical plate (CP), subventricular zone (SV), caudoputamen (CPu), and hippocampus (H) but not in the intermediate zone (IZ). At P0, a small number of IGFBP3-positive cells are restricted in the frontal cortical plate. At P28, IGFBP3-positive cells are scattered in the frontal cortex. At P60, no IGFBP3-positive cells were observed. Original magnification: postconceptual day 17.5, 4× and 200×; P0, P28, and P60, 4×; enlargement of P28, 200×.

genotype, *IGFBP3* expression increased more in *mecp2*$^{-/y}$ male mice than in the wild-type male mice of the same age.

Brain cDNA from four mice at each genotype (4 pools each comprising 3 individual brain cDNA preparations) was compared with real-time PCR analysis, using *MAP2* mRNA as an internal control and *β-actin* mRNA as a reference. As a result, both the *IGFBP3* mRNA expression level and the *BDNF* mRNA expression level of the cortices of the *mecp2*$^{-/y}$ mice were significantly higher (approximately twice as high) than those of wild-type mice and *mecp2*$^{+/-}$ mice, whereas there was no significant difference between wild-type mice and *mecp2*$^{+/-}$ mice (Fig. 1D). Both *IGFBP3* and *BDNF* mRNA were upregulated in P28 *mecp2*$^{-/y}$ mice, at the presymptomatic stage, compared with wild-type controls and *mecp2*$^{+/-}$ mice. The *MAP2* expression ratio was 0.85 ± 0.12 (average ± SD) of wild-type male, 0.95 ± 0.15 of *mecp2*$^{+/-}$ female mice, and 0.99 ± 0.17 of *mecp2*$^{-/y}$ male mice. The *IGFBP3* expression ratio was 0.85 ± 0.21 of wild-type male mice, 1.30 ± 0.27 of *mecp2*$^{+/-}$ female mice, and 1.75 ± 0.12 of *mecp2*$^{-/y}$ male mice and showed a significant difference ($p < 0.001$, Student *t*-test). The *BDNF* expression ratio was 0.78 ± 0.21 of wild-type male mice, 1.25 ± 0.51 of *mecp2*$^{+/-}$ female mice, and 1.88 ± 0.13 of *mecp2*$^{-/y}$ male mice and also showed a significant difference ($p < 0.001$). This *IGFBP3* increase due to *MeCP2* deficiency was thought to be at the same level as *BDNF*, increased expression of which due to lack of *MeCP2* was consistent with the previously reported findings (10,11). Interestingly, *IGFBP3* expression in *mecp2*$^{+/-}$ showed an intermediate amount between *mecp2*$^{-/y}$ and wild-type mice, and there was no significant difference. *BDNF* expression also had the same pattern as *IGFBP3* expression.

Immunoblot analysis confirmed the higher expression of IGFBP3 as protein level in the *mecp2*$^{-/y}$ brain, especially in the period after P28 (Fig. 2A). IGFBP3 age-dependent expression may be revealed, because of faint bands at P0 and P60 (Fig. 2A). Expression of IGFBP3 protein in *mecp2*$^{-/y}$ mice increased more than that in wild-type mice. A light molecule of IGFBP3 was detected only in fetuses of both mice, whereas a heavy band was predominantly detected and the light band was faintly seen at postnatal ages. The multiple bands observed in postnatal mice were consistent with those already reported, in which the heavy band is thought to be a glycosylated and mature form (14, 17). Compared with wild-type males, *mecp2*$^{-/y}$ mice showed significant increases of IGFBP3, especially after P28.

IGFBP3 immunohistochemical analysis revealed an increase in the number of positive cells and fibers in the mutant brains. In wild-type mouse brain, IGFBP3-positive cells distributed in the cortex, and these cells were very faint (Fig. 2B). However, some IGFBP3 immunoreactivity was observed in the cytoplasm and fibers in the cortex of *mecp2*$^{+/-}$ mice (Fig. 2C). *Mecp2*$^{-/y}$ mice showed more IGFBP3-positive cells and fibers than wild-type male and *mecp2*$^{+/-}$ mice (Fig. 2B–D). MeCP2 was recognized in many nuclei in the cortex of wild-type male and *mecp2*$^{+/-}$ mice (Fig. 2E, F); however, *mecp2*$^{-/y}$ mice had no MeCP2 immunoreactivity (Fig. 2G). In the brain of E17.5 mice, IGFBP3 positivity was observed mainly in the cortical plate, hippocampus, and caudoputamen (Fig. 2H). In the cortical plate, IGFBP3 immunostaining located in the cytoplasm and/or intercellular space. In the P0 brain, IGFBP3-positive cells and fibers were limited to the cortex. The IGFBP3 immunostaining decreased with age. The mature brain of P60 mice evidenced no IGFBP3 immunoreactivity, except for striatum and fibers in the white matter (data not shown). In the white matter, IGFBP3 immunoreactivity was shown in some fibers and a small number of glial cells (data not shown).

In the brains of P28 mice, whose IGFBP3 protein concentration was the highest of the three ages examined, IGFBP3 immunohistochemical features in *mecp2*$^{-/y}$ were recognized to increase those reactivities, as determined by Western blot analysis (Fig. 2A).

Immunoexpression of IGFBP3 in Human Brains With RTT and Controls

We found similar immunostaining patterns in human brains. Immunoblot data of 5 RTT patients showed a higher level of IGFBP3 expression than that of controls, indicating that MeCP2-positive neurons express less IGFBP3 (Fig. 3A). Cortical neurons of the controls revealed no IGFBP3 immunoreactivity (Fig. 3B). However, a few IGFBP3-positive neurons and fibers were observed in a heterozygous female RTT patient (Fig. 3C). These results showed a pattern similar to that of mice brains. MeCP2-positive cells sometimes expressed low or undetectable levels of IGFBP3 (Fig. 3D, E). However, these data suggest that the presence of MeCP2 in a neuronal cell is negatively compatible with IGFBP3 expression. It is possible that the role of MeCP2 at the promoter is not only to silence the genes but also to modulate levels of expression.

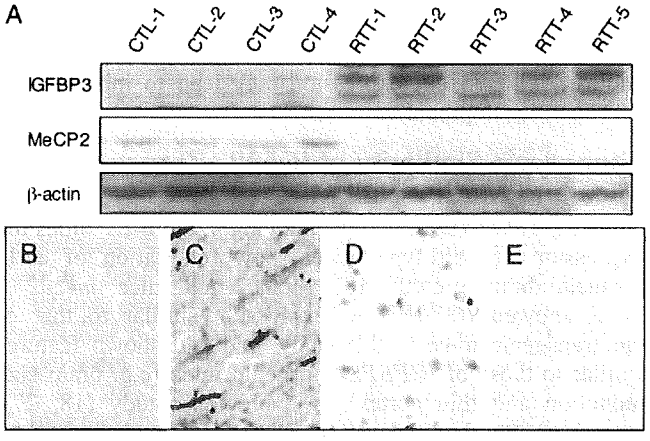

FIGURE 3. Insulin-like growth factor binding protein 3 (IGFBP3) expression analyses of human brains. **(A)** IGFBP3 immunoblot in human brain shows the same pattern as in mice; these patients are presented in the Table. Control (CTL) brain (15 years old) shows very few IGFBP3-positive cells **(B)**, but RTT brain (14 years old) demonstrates many IGFBP3-positive cells and fibers **(C)**. MeCP2-positive cells are observed in control **(D)**, but few in RTT **(E)**. All tissues are the frontal cortices. Original magnification of all histologic figures: 200×.

DISCUSSION

Genetic Relationship Between IGFBP3 and MeCP2

We found that the promoter of *IGFBP3* had a MeCP2 binding site and that *IGFBP3* expression in mouse brains was regulated by MeCP2. Structural analysis of the human *IGFBP3* promoter revealed a dense cluster of CpG islands spanning the region from −220 to −20 bp (14). The mouse *IGFBP3* promoter showed CpG islands from −300 to −120 bp. These CG-rich sites of the *IGFBP3* promoters are thought to regulate its expression. They were earlier reported to be hypermethylation in some cancer cells (15). Chromatin immunoprecipitation assays showed that the *IGFBP3* promoter bound to MeCP2 in the brain tissue. MeCP2 is strongly suggested to contribute to the transcriptional expression of *IGFBP3*. These facts suggest that MeCP2 may be a regulator of *IGFBP* expression and that MeCP2 mutations may lead to *IGFBP3* overexpression, as previously discussed in *BDNF* (10, 11).

Functions of IGFBP3 and Relationship to Phenotypes

It is known that IGFBP3 binds to IGF-1 and IGF-2, major factors for cell growth and elongation of neuronal dendrites and axons and inhibits these functions (17, 23). IGF-1 and IGF-2 are present in most biologic fluids as a complex with IGFBPs. Of the 6 IGFBPs, IGFBP3 is the most abundant in cerebrospinal fluid (23). A role for IGFBP3 in the transport and modulation of the biologic actions of the IGF has been proposed (17). Moreover, IGFBP3 appears to inhibit cell proliferation and differentiation in some cell lines (24, 25).

Overexpression of IGFBP3 shows an elevation of IGF-1 concentration and is also associated with intrauterine and postnatal growth retardation (26). These effects on growth are most likely explicable on the basis of IGF action inhibition. Although its exact physiologic function remains unclear, IGFBP3 overexpression may lead to inhibition of neuronal development.

Phenotypic Similarity of IGFBP3-Transgenic Mice and MeCP2-Null Mice

IGFBP3-transgenic mice, having 4.9 and 7.7 times the expression of wild-type mice, show retardation of early postnatal brain growth (18). MeCP2-null mice, *mecp2*$^{-/y}$ mice, showed *IGFBP3* overexpression similar to that of the transgenic mice in this study, and had shown features similar to those of *IGFBP3*-transgenic mice (e.g. brain size reduction and thin cortex) in a previous study (18, 27). *Mecp2*$^{-/y}$ mice had delayed synaptogenesis and diminished brains (27, 28). These findings allow us to speculate that overexpression of IGFBP3 due to lack of MeCP2 leads to delayed brain maturation and plays a role in these conditions in humans as well. Interestingly, MeCP2 and IGFBP3 in mouse brain showed the highest expression at P28 before the occurrence of the initial symptoms in *mecp2*$^{-/y}$ mice, such as spastic paraplegia and hypoactivity, so these features may be associated with IGFBP3 overexpression in mice brains.

IGFBP3 Expression of MeCP2 Heterozygous Mutated Mice and RTT Patients

Genetically, *mecp2*$^{+/-}$ mice are the same genotype as human RTT. Both have heterozygous mutation of *MeCP2*. *IGFBP3* and *BDNF* expression in *mecp2*$^{+/-}$ showed an intermediate amount between *mecp2*$^{-/y}$ and wild-type mice from our real-time quantitative PCR. The 1.5-fold increase of *IGFBP3* may also act as one of the pathogenetic factors, whereas there were no significant differences, or this pathologic discrepancy in heterozygous mutation may be due to the difference in species. From the standpoint of *IGFBP3* expression, *mecp2*$^{-/y}$ mice phenotypes may be compatible to those of human RTT. *IGFBP3*, which was identified as a novel MeCP2-downstream gene in the present study, exhibited an approximately 2-fold elevation of its mRNA expression in the *mecp2*-null brain, the same level as that of *BDNF*.

MeCP2 as Transcriptional Repressor

It is noteworthy that the absence of *mecp2* leads to elevated levels of *IGFBP3* transcripts before the onset of overt symptoms in these mice. The effect of overexpression of *IGFBP3* is an early phenotype that manifests during postnatal brain development, possibly including dendritic changes in both rodents and humans. Related changes occur in brains of RTT patients (28, 29). Furthermore, seizures and heightened anxiety of mice with mutations in the *mecp2* gene (30) and the glucocorticoid-related gene products, which were growth-regulated proteins, overexpressed in RTT model mice (31).

Because MeCP2 is a transcriptional repressor, the predominant current hypothesis to explain RTT is that critical genes are aberrantly expressed in its absence. Misregulation of several genes is conceivable. Aberrant expression of the MeCP2-target genes, *BDNF*, *Dlx5*, *Ube3a*, *Gabrb3*, *Igfbp3*, or others, might be a contributor to the phenotype. MeCP2 misregulation of the expression of these genes expression may have neurologic consequences that could give rise to disease.

It is important to clarify the gene expression changes in molecular pathways of MeCP2-downstream systems in mice and human brains with the phenotypes. Such genes, which are linked with neuronal function at critical developmental stages and play important roles in specific neuronal development, are thought to contribute to the characteristic and pathognomonic neuronal involvement in RTT.

ACKNOWLEDGMENTS

We thank Dr. K. Suijo, Tokyo Gakugei University, for his generous assistance; Mr. S. Kumagai, NCNP, for technical assistance; and Dr. K. Endoh, University of Yamanashi, for helpful advice on genetic information. Thanks are also due to Dr. K. Ishi, a pathologist of Juntendo University Urayasu Hospital, for performing the autopsy of a Japanese case; and the Harvard Brain Tissue Resource Center for providing RTT samples.

REFERENCES

1. Nan X, Campoy FJ, Bird A. MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. Cell 1997;88:471–81
2. Jones PL, Veenstra GJ, Wade PA, et al. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nat Genet 1998;19: 187–91
3. Nan X, Ng HH, Johnson CA, et al. Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature 1998;393:386–89
4. Amir RE, Van den Veyver IB, Wan M, et al. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. Nat Genet 1999;23:185–88
5. Milani D, Pantaleoni C, D'Arrigo S, et al. Another patient with MECP2 mutation without classic Rett syndrome phenotype. Pediatr Neurol 2005;32:355–57
6. Watson P, Black G, Ramsden S, et al. Angelman syndrome phenotype associated with mutations in MECP2, a gene encoding a methyl CpG binding protein. J Med Genet 2001;38:224–28
7. Itoh M, Takashima S. Neuropathology and immunohistochemistry of brains with Rett syndrome [in Japanese]. No To Hattatsu 2002;34: 211–16
8. Shahbazian MD, Antalffy B, Armstrong DL, et al. Insight into Rett syndrome: MeCP2 levels display tissue- and cell-specific differences and correlate with neuronal maturation. Hum Mol Genet 2002;11:115–24
9. Kishi N, Macklis JD. MECP2 is progressively expressed in post-migratory neurons and is involved in neuronal maturation rather than cell fate decisions. Mol Cell Neurosci 2004;27:306–21
10. Martinowich K, Hattori D, Wu H, et al. DNA methylation-related chromatin remodeling in activity-dependent BDNF gene regulation. Science 2003;302:890–93
11. Chen WG, Chang Q, Lin Y, et al. Derepression of BDNF transcription involves calcium-dependent phosphorylation of MeCP2. Science 2003; 302:885–89
12. Horike S, Cai S, Miyano M, et al. Loss of silent-chromatin looping and impaired imprinting of DLX5 in Rett syndrome. Nat Genet 2005;37: 31–40
13. Samaco RC, Hogart A, LaSalle JM. Epigenetic overlap in autism-spectrum neurodevelopmental disorders: MECP2 deficiency causes reduced expression of UBE3A and GABRB3. Hum Mol Genet 2005; 14:483–92
14. Chang YS, Wang L, Suh YA, et al. Mechanisms underlying lack of insulin-like growth factor-binding protein-3 expression in non-small-cell lung cancer. Oncogene 2004;23:6569–80
15. Hanafusa T, Yumoto Y, Nouso K, et al. Reduced expression of insulin-like growth factor binding protein-3 and its promoter hyper-methylation in human hepatocellular carcinoma. Cancer Lett 2002;176: 149–58
16. Darwanto A, Kitazawa R, Maeda S, et al. MeCP2 and promoter methylation cooperatively regulate E-cadherin gene expression in colorectal carcinoma. Cancer Sci 2003;94:442–47
17. Payet LD, Wang XH, Baxter RC, et al. Amino- and carboxyl-terminal fragments of insulin-like growth factor (IGF) binding protein-3 cooperate to bind IGFs with high affinity and inhibit IGF receptor interactions. Endocrinology 2003;144:2797–806
18. Modric T, Silha JV, Shi Z, et al. Phenotypic manifestations of insulin-like growth factor-binding protein-3 overexpression in transgenic mice. Endocrinology 2001;142:1958–67
19. Guy J, Hendrich B, Holmes M, et al. A mouse Mecp2-null mutation causes neurological symptoms that mimic Rett syndrome. Nat Genet 2001;27:322–26
20. El-Osta A, Wolffe AP. Analysis of chromatin-immunopurified MeCP2-associated fragments. Biochem Biophys Res Commun 2001;289:733–37
21. Kudo S. Methyl-CpG-binding protein MeCP2 represses Sp1-activated transcription of the human leukosialin gene when the promoter is methylated. Mol Cell Biol 1998;18:5492–99
22. Itoh M, Suzuki Y, Akaboshi S, et al. Developmental and pathological expression of peroxisomal enzymes: Their relationship of D-bifunctional protein deficiency and Zellweger syndrome. Brain Res 2000;858: 40–47
23. Cubbage ML, Suwanichkul A, Powell DR. Insulin-like growth factor binding protein-3: Organization of the human chromosomal gene and demonstration of promoter activity. J Biol Chem 1990;265:12642–49
24. Chang YS, Wang L, Liu D, et al. Correlation between insulin-like growth factor-binding protein-3 promotor methylation and prognosis of patients with stage I non-small cell lung cancer. Clin Cancer Res 2002; 8:3669–75
25. Niblock MM, Brunso-Bechtold JK, Riddle DR. Insulin-like growth factor I stimulates dendritic growth in primary somatosensory cortex. J Neurosci 2000;20:4165–76
26. Itoh M, Fukuda T. Comparative neuropathology and immunohistochemistry of brains with human Rett syndrome and its model mice. Brain Dev 2004;26:415–16
27. O'Kusky JR, Ye P, D'Ercole AJ. Insulin-like growth factor-I promotes neurogenesis and synaptogenesis in the hippocampal dentate gyrus during postnatal development. J Neurosci 2000;20:8435–42
28. Fukuda T, Itoh M, Ichikawa T, et al. Delayed maturation of neuronal architecture and synaptogenesis in cerebral cortex of *mecp2*-deficient mice. J Neuropathol Exp Neurol 2005;64:537–44
29. Armstrong DD. Review of Rett syndrome. J Neuropathol Exp Neurol 1997;56:843–49
30. Shahbazian M, Young J, Yuva-Paylor L, et al. Mice with truncated MeCP2 recapitulate many Rett syndrome features and display hyper-acetylation of histone H3. Neuron 2002;35:243–54
31. Nuber UA, Kriaucionis S, Roloff TC, et al. Up-regulation of glucocorticoid-regulated genes in a mouse model of Rett syndrome. Hum Mol Genet 2005;14:2247–56

Journal of Exposure Science and Environmental Epidemiology (2007) 17, 84–87

www.nature.com/jes

Is the midnight-to-midnight average concentration of pollutants an appropriate exposure index for a daily mortality study?

MASAJI ONO[a], TAKASHI OMORI[b] AND HIROSHI NITTA[c]

[a]*Environmental Health Sciences Division, National Institute for Environmental Studies, Onogawa 16-2, Tsukuba, Japan*
[b]*Department of Biostatistics, Kyoto University School of Public Health, Onogawa 16-2, Tsukuba, Japan*
[c]*PM2.5 & DEP Research Project Team, National Institute for Environmental Studies, Onogawa 16-2, Tsukuba, Japan*

Midnight-to-midnight average pollutant concentration has been widely used as the exposure index in many epidemiological studies. This index, however, does not take account of the variability of time of death, because it simply takes 24-h average concentration from midnight to midnight, and this is equivalent to assuming that all deaths happened at the end of the day. This assumption is clearly inappropriate in the real situation, and might be a significant information bias in the exposure assessment. In order to take account of the variability of time of death, a new exposure index, the number-of-death-weighted 24-h average, is introduced in this study. This new index is calculated for each subject by taking 24-h average concentration preceding the time of death. In this study, this new index is applied to the suspended particulate matter (SPM) exposure assessment in Japan, and compared with the result obtained from the conventional index, the midnight-to-midnight average. There are quite large differences in the SPM average concentration between these two indices. Moreover, risk ratio calculated by the new index at time lag 0, where the mortality data and the average concentration data on the same day are used, is a lot larger than that obtained by the conventional one. In addition to the difference in its magnitude, time lag structure (from 1–5 days) also shows the different patterns between these two indices. This study suggests that the conventional exposure index, the midnight-to-midnight average, may evaluate the risk of SPM incorrectly, because of its inadequacy of capturing the temporal variability.
Journal of Exposure Science and Environmental Epidemiology (2007) **17,** 84–87. doi:10.1038/sj.jes.7500527; published online 27 September 2006

Keywords: *daily mortality study, daily average concentration, midnight-to-midnight average, time of death.*

Introduction

Because of increasing public concern about air pollution, many researches on an association between air pollutant and its adverse health effect have been conducted. Particulate matter (PM) is one of the airborne pollutants that are widely studied all over the world, and many epidemiological studies have demonstrated an association between its concentration and daily mortality (Dockery et al., 1992; Schwartz and Dockery, 1992; Katsouyanni et al., 1997; Peters et al., 2000; Samet et al., 2000a). Our group also reported a relationship between suspended particulate matter (SPM: particles whose diameter is less than 10 μm) and daily mortality in 13 major cities in Japan (Omori et al., 2003).

In most studies, simple midnight-to-midnight average pollutant concentration has been used as the exposure index. This exposure index, however, may fail to detect the effect of short-term temporal variation in pollutant concentration on mortality, as the index simply takes 24-h average from midnight to midnight, and this is equivalent to assuming that all deaths happened at the end of the day. This assumption is clearly inappropriate in the real situation and may cause significant information bias in the exposure assessment. However, the effect of temporal variation on the exposure index has not been discussed so far. Therefore in this study, we focus on this issue and introduce the new exposure index, the number-of-death-weighted 24-h average.

The number-of-death-weighted 24-h average is obtained by taking 24-h average concentration preceding the time of death for each case. This new index enables us to take into account the variability of time of death and to examine the effect of short-term variation on mortality. In this study, this index is applied to the SPM exposure assessment in Japan. Due to the effort of the Ministry of Health, Labour and Welfare, and National Institute for Environmental Studies in Japan, hourly mortality data and air pollution data are available at all cities in Japan. We used these data sets in order to calculate the new exposure index and compared the result with that obtained from the conventional index, the midnight-to-midnight average.

1. Address all correspondence to: Dr. M. Ono, National Institute for Environmental Studies, Onogawa 16-2, Tsukuba, 305-8506 Japan.
Tel.: +81 29 850 2421. Fax: +81 29 850 2588.
E-mail: onomasaj@nies.go.jp
Received 19 March 2005; accepted 25 August 2006; published online 27 September 2006

Methods

Data

In this study, we selected the same data that was used in our previous study (Omori et al., 2003) to demonstrate the difference between these two indices. Mortality data were obtained from the Ministry of Health, Labour and Welfare, air pollutant data from the National Institute for Environmental Studies, and meteorological data from the Japan Meteorological Agency. We used the mortality data, which is collected from 1990 to 1994 in 13 major cities in Japan (Tokyo, Osaka, Yokohama, Nagoya, Kyoto, Kobe, Sapporo, Kitakyushu, Fukuoka Hiroshima, Kawasaki, Sendai, and Chiba) for residents whose age is over 65 years.

In this study, we examined following three causes of death: all causes other than accidents, respiratory diseases, and cardiovascular diseases. Residential area of subjects was assigned based on their home address rather than where they died. The concentration of five air pollutants, suspended particulate matter (SPM), sulfur dioxide (SO_2), nitrogen dioxide (NO_2), carbon monoxide (CO), and oxidant (Ox), and two meteorological factors, temperature and relative humidity, were used to calculate the risk ratio.

In order to obtain the accurate average concentration, we only used the cases in which there were more than 20 hourly concentration records within 24-h-period preceding the death.

Exposure Indices

The new exposure index, the number-of-death-weighted 24-h average, is given by

$$x^{(1)}k,i=\sum_{j=0}^{23}\left\{y_{ij}\times\left(\frac{1}{24}\sum_{l=j-23}^{j}x_{k,il}\right)\right\}/y_i \qquad (1)$$

where $x_{k,ij}$ is the concentration of air pollutant k, and y_{ij} is the number of deaths at hour j on day i. y_i is daily mortality given by the following formula.

$$y_i=\sum_{j=0}^{23}y_{ij} \qquad (2)$$

This index $x^{(1)}k,i$ provides the exact 24-h average preceding the death for each case. On the other hand, the conventional method, the midnight-to-midnight average, is given by the following formula.

$$x^{(2)}k,i=\frac{1}{24}\sum_{j=0}^{23}x_{k,ij} \qquad (3)$$

This average provides the 24-h average from 0:00 to 24:00 at day i, and was widely used as exposure index in many daily mortality studies.

In Figure 1, we illustrate images of the number-of-death-weighted 24-h average and the midnight-to-midnight average. Obviously, the number-of-death-weighted 24-h average produces more accurate 24-h exposure than the midnight-to-midnight average.

Statistical Analysis

In order to calculate the risk ratios for each city, the daily mortality, y_i, is modeled by using the generalized additive model (GAM) with log link function. In the model, a logarithm of the mortality is expressed as the function of the concentration of SPM and the other cofactors, such as, the concentrations of the other air pollutants, temperature, and relative humidity. Following formula shows the GAM model used in this study.

$$\begin{aligned}\log y_i=&\ A_1*SPM_i+A_2*SO_{2i}\\&+A_3*NO_{2i}+A_4*CO_i+A_5*Ox_i\\&+f(Temp_i)+A_6*Hum_i\end{aligned} \qquad (4)$$

where A_1–A_6 are estimated parameters, and SPM_i, SO_{2i}, NO_{2i}, CO_i, Ox_i, $Temp_i$ and $Humi_i$ are variables that correspondent to $x^{(1)}k,i$ or $x^{(2)}k,i$.

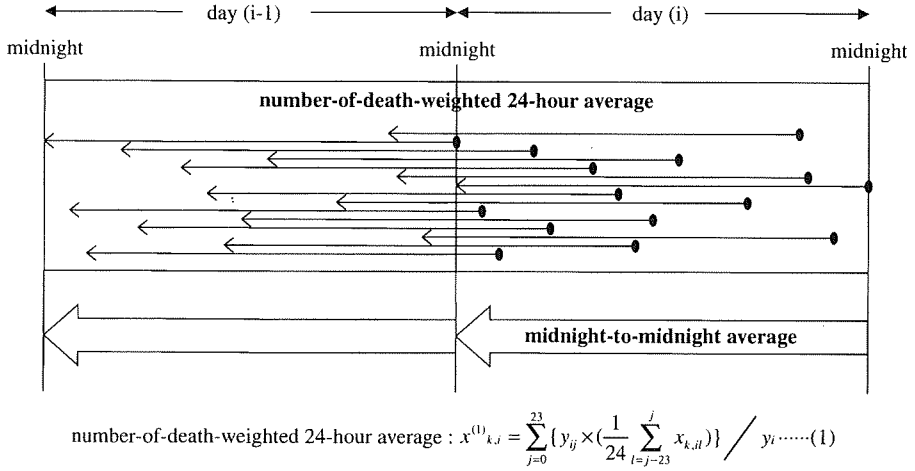

Figure 1. Illustration of exposure indices. In this figure we assume that the pollution concentration over 24 h before death affect death.

Therein, concentrations for air pollutants and relative humidity ($Humi_i$) are treated as linear term, whereas the temperature ($Temp_i$) as non-linear smooth functions. When using the model, risk ratios of SPM for each city are obtained as $\exp(A_1)$. Then, the summarized risk ratio is calculated using the meta-analysis technique with random effect model proposed by DerSimonian and Laird (1986).

In order to analyze the effect of the time lag, we calculated the risk ratio using the five different time lags (from 1–5 days).

The details of these calculation methods were explained in our previous study (Omori et al., 2003). For the calculation, the SAS statistical software package version 8.2 (SAS Institute Inc., NC, USA) was used.

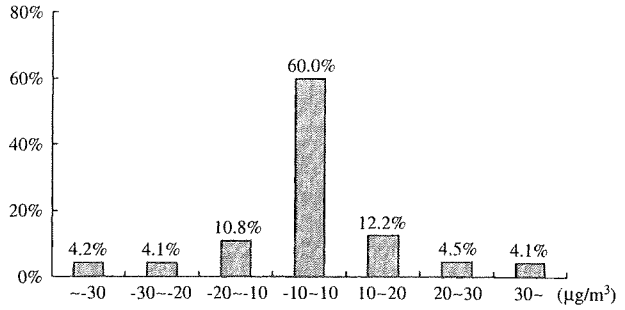

Figure 2. Distribution of the difference in daily average SPM concentration between the number-of-death weighted 24-h average and the midnight-to-midnight average.

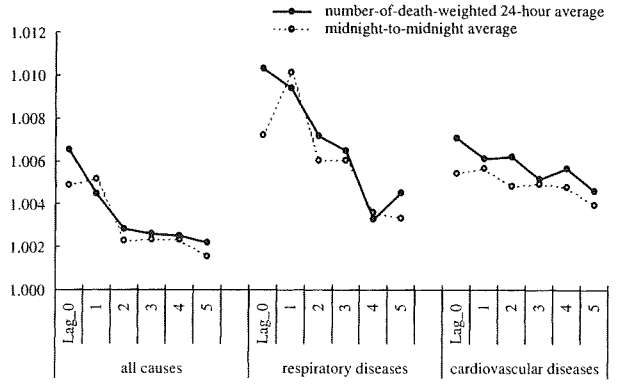

Figure 3. Comparison of risk ratios and time lag in two methods. The risk ratios are shown as an increase of mortality for an increase in SPM of 10 $\mu g/m^3$. Closed circle shows the number-of-death-weighted 24-h average and open circle shows the midnight-to-midnight average.

Result

Difference in SPM Concentration Between Two Exposure Indices

In Figure 2, we show the distribution of the difference in 24-h average SPM concentration between the number-of-death-weighted 24-h average and the midnight-to-midnight average. This figure indicates that about 40% of data show the gap of more than 10 $\mu g/m^3$. This figure is based on the mortality data classified as all causes other than accident in Tokyo, but the distributions of the difference between these two indices are quite similar in the other locations and in other causes of death (not shown in figure or table).

Risk Ratios

In Figure 3, we show the risk ratio and time lags from lag 0 to lag 5 for the mortality data classified as all cases other than accident, respiratory disease, and cardiovascular disease. At lag 0 where the mortality data and the 24-h concentration on the same day are used, risk ratios calculated by the number-of-death-weighted 24-h average concentration are larger than those obtained by the midnight-to-midnight average concentration for all three types of mortality data. The risk ratios obtained by the number-of-death-weighted 24-h average concentration are 1.0065, 1.0103, and 1.0071 for the mortality data classified as all causes other than accident, respiratory disease, and cardiovascular disease, respectively. On the other hand, those obtained by the midnight-to-midnight average concentration are 1.0049, 1.0072, and 1.0054.

In addition to the magnitude of the risk ratio, the lag structure also shows the different pattern between these two indices. For all three types of mortality data, the risk ratios obtained by the number-of-death-weighted 24-h average have their peaks at lag 0. On the other hand, the risk ratios obtained by the midnight-to-midnight average reach their peaks at lag 1.

Discussion

Although many studies on exposure assessment of the air pollutants have focused on the validity of the spatial representation such as locations where the subjects exposed, that of the temporal representation, the effect of temporal misclassification of the exposure index, has not been discussed.

As we have shown in Figure 2, the difference between two exposure indices is quite large, about 40% of data show the gap of more than 10 $\mu g/m^3$ and the largest difference reaches 106.7 $\mu g/m^3$. This indicates that the new exposure index, the number-of-death-weighted 24-h average, leads the significantly different results. The effects of the choice of the exposure index are exemplified in the results shown in Figure 3.

In Figure 3, we show the risk ratio and time lag for three different types of mortality data, all causes other than accidents, respiratory diseases, and cardiovascular diseases. This result shows the difference of magnitude of the risk at lag 0 and the completely different time lag structure between

two exposure indices. At lag 0, risk ratios calculated by the number-of-death-weighted 24-h average concentration are larger than those obtained by the midnight-to-midnight average concentration. This result indicates that our previous study underestimated the magnitude of the risk of SPM exposure at lag 0 for all three categories of death.

Figure 3 also shows the different lag structure between these two indices. Risk ratios obtained by the number-of-death-weighted 24-h average concentration have their peaks at lag 0. On the other hand, in our previous study, in which the midnight-to-midnight average was used, risk ratios reached their peaks at lag 1. Many studies (Peters et al., 2000; Samet et al., 2000b; Dominici et al., 2003) also reported the same lag pattern as our previous study. Dominici et al. (2003) showed that in the lag structure, the effect of PM_{10} on mortality was greatest at lag 1, and that this effect remained for total mortality, cardiovascular diseases, and respiratory diseases in the updated analysis of the 90 US cities previously studied. This disagreement of the lag structure suggests that risk ratio obtained by the midnight-to-midnight average concentration possibly leads the biased estimate of the temporal variability of the risk of SPM.

As we demonstrated above, estimation of the risk ratio is sensitive to the choice of the exposure index. Two different exposure indices, the number-of-death-weighted 24-h average concentration and the midnight-to-midnight average concentration, lead completely different result not only of the magnitude of the risk ratio at lag 0 but also of its temporal variability. These results suggest that using the midnight-to-midnight average concentration as exposure index might be the serious information bias, and we have to pay more attention to the choice of exposure index, especially when fine data are available.

References

DerSimonian R., and Laird N. Meta-analysis in clinical trials. *Controlled Clin Trials* 1986: 7: 177–188.

Dockery D.W., Schwartz J., and Spengler J.D. Air pollution and daily mortality — associations with particulates and acid aerosols. *Environ Res* 1992: 59: 362–373.

Dominici F., McDermott A., Daniels M., Zeger S.L., and Samet J.M. Mortality among residents of 90 cities. Revised analyses of time-series studies of air pollution and health. In: *Health Effects Institute*. Boston, MA, 2003.

Katsouyanni K., Touloumi G., Spix C., Schwartz J., Balducci F., and Medina S., et al. Short term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: results from time series data from the APHEA project. *Br Med J* 1997: 314: 1658–1663.

Omori T., Fujimoto G., Yoshimura I., Nitta H., and Ono M. Effects of particulate matter on daily mortality in 13 Japanese cities. *J Epidemiol* 2003: 13: 314–322.

Peters A., Skorkovsky J., Kotesovec F., Brynda J., Spix C., and Wichmann H.E., et al. Associations between mortality and air pollution in central Europe. *Environ Health Perspect* 2000: 108: 283–287.

Samet J.M., Dominici F., Curreiero F.C., Coursac I., and Zeger S.L. Fine particulate air pollution and mortality in 20 U.S. cities, 1987 to 1994. *N Engl J Med* 2000a: 343: 1742–1749.

Samet J.M., Zeger S.L., Dominici F., Curriero F., Coursac I., Dockery D.W., Schwarz J., and Zanobetti A. The National Morbidity, Mortality, and Air Pollution Study (NMMAPS): Part 2. Morbidity, Mortality and Air Pollution in the United States. *Health Effects Institute*. Health Effects Institute, Cambridge, Mass, 2000b.

Schwartz J., and Dockery D.W. Increased mortality in Philadelphia associated with daily air pollution concentrations. *Am Rev Resp Dis* 1992: 145: 600–604.

日本統計学会誌
第36巻，第2号，2007年3月
195頁～204頁

重篤な有害事象を速く検出するためのシグナル検出法の検討

大森　崇*

A study of speedier signal detection for serious adverse events

Takashi Omori

近年，市場での医薬品の有害事象報告のデータベースから副作用の疑われる医薬品を，客観的な指標に基づき見つけだすシグナル検出という方法が注目され，安全性監視のための一つのツールとして使用されている．一般に有害事象が未知で重篤である場合には，シグナルの検出は速いほど望ましいと思われる．しかしこのような観点からシグナル検出方の検討は行われてこなかった．そこで，有害事象が重篤な場合にはより速くシグナルが検出できるように，ベイズ流のアプローチを用いた検出の指標を構成し，シミュレーション研究によって既存の方法と比較した．検討した方法は重篤ではない場合には既存の方法と同程度の能力であり，重篤な場合にはより早く検出できる方法であることを確認した．

When a marketed drug has some suspicions for unknown adverse events (AE), we want to detect it as a signal for its evaluation. In practice, if the AE is serious, speedier detection might be more desirable than the detection by evidence with accumulation of the number of events. Though recently, some statistical methods have been proposed to detect the signal from a database of spontaneous reports, seriousness has not been incorporated in these. We developed a simple Bayesian method for signal detection taking account of seriousness, and compared to an existing method (BCPNN; the Bayesian confidence proportional neural network). When the true reporting ratio (TRR) was set to 1, results for both methods were similar. On the other hand, when the TRR was set to 2 and 3, the proposed method, when applied to a serious AE, was speedier than that when applied to a non-serious AE, and the BCPNN.
Key Words and Phrases: signal detection, serious adverse event, Bayesian method

1. はじめに

医薬品が承認されて市場に出る前には，非臨床試験，臨床試験などによりその有効性や安全性が確認される．しかし，承認前に実施されるそれらの試験では規模や対象が限られているため，完全な評価を行うことは不可能である．このため，製造販売された後の医薬品の適正な使用方法を確立することを目的に，わが国では再審査制度および安全性定期報告，再評価制度ならびに副作用・感染症報告制度から構成される Post-Marketing Surveillance（PMS）の制度が導入されている．再審査制度および安全性定期報告では，新医薬品等について製造販売承認後，使用成績などに関する調査を行い，有効性および安全性について再確認する．再評価制度は，すでに承認された医薬品について品質，有効性，安全性を見直す制度である．副作用・感染症制度は，企業報告制度，WHO 国際モニタリング制度などを含んでいる．企業報告制度では，

* 京都大学大学院医学研究科医療統計学分野，〒606-8501 京都市左京区吉田近衛町

企業は医薬品による有害事象を知ってから早期に報告をしなくてはならない.

再審査制度および安全性定期報告と再評価制度はある計画された時点で該当する医薬品の評価を行なうので，時間計画的に医薬品を監視する方法といえる．一方，副作用・感染症報告制度は，安全性が懸念されるような事態が生じた時点で問題が把握できる制度であるので，市場の医薬品の状態を監視するという側面がある.

近年，医療機関・薬局から報告される有害事象の報告を集めてデータベースを作成し，副作用の疑いのある医薬品をみつける方法が各国の規制当局などで用いられるようになってきた．この方法はシグナル検出と呼ばれている．このような方法が注目されている理由は，報告を受ける側にとって膨大な数の報告を処理する必要が生じているためであること，シグナル検出の指標の計算により客観的な評価が行えることがあげられる.

現在までに，シグナル検出のためのさまざまな手法が提案されている．その基本的な原理は，医薬品と有害事象から構成される2×2分割表の各セルの頻度を用いて関連の指標を計算し，この値が大きいことでシグナルと判断するというものである（例えば，van Puijenbroek ら (2003)，Bata ら (2002)，Szarfman ら (2002)）．シグナル検出は医薬品の状態を監視することを目的に用いられるであろうから，検出の速さは個々の方法の性能を評価する上で重要な要素だと思われるが，多くの研究はあまりその点が重視されていないようである.

Szarfman ら (2002) は，米国の Food and Drug Administration (FDA) の自発報告のデータベースを用いてシグナル検出の方法と過去に行われていた方法との比較を行った．この結果，過去の方法の判断に比べシグナル検出の方法では速く検出できることが多かった．しかしながら，一方ではシグナル検出の方が過去の方法に比べて遅かった有害事象もあり，それらの多くは重篤な有害事象であった．これはおそらくシグナル検出の方法が精度よく安定した推定値を得るためにそれなりのデータの量（報告数）が必要となることに起因していると思われる．しかし，安定した検出を得るという理由から多くの報告数を求めることが，重篤な有害事象のように報告数が少なくても早期に行動をとらなくてはならないような場合に適しているであろうか.

本稿では，報告された有害事象が重篤である場合により速くシグナルを検出する方法を構築することを目的とする．主観確率を用いたベイズ統計学の利用が一つの方法であると思われる．そこで，ここではベイズ統計学のアプローチによりシグナル検出の指標を構成し，その性能を検討することにする.

第2節では，シグナル検出の原理を解説するとともに，シグナル検出によって重篤な有害事象を検出するための方法を構築する．そして，その方法の性能を評価するためのシミュレーション研究の方法を記述する．第3節ではシミュレーションの結果を示す．第4節では，検討する方法とシグナル検出のあり方について議論を行う.

2. 方　　法

2.1 シグナル検出の原理

はじめに想定するデータを説明する．得られるデータは個々の医薬品における何らかの有害事象の報告であり，各報告は医薬品名と有害事象名が記載されている．データベースにはこのような報告が蓄積され，I 個の医薬品と J 個の有害事象名からなる $I\times J$ の分割表の形をしているとする.

今，特定の医薬品が特定の有害事象を多く引き起こす傾向があるのではないかということに興味があるとしよう．このとき，先の $I\times J$ の分割表を，医薬品が対象としている医薬品であるか否か，また有害事象が対象としている有害事象であるか否かの2×2の分割表に要約する

ことができる（表1）．表1における $x_{11}, x_{21}, x_{21}, x_{22}$ はシグナル検出の指標を計算する時点で報告されている頻度を表している．シグナル検出の方法は基本的にはこの分割表に基づいて行われる．

シグナルを検出する指標として，いくつかの方法が提案されている．Proportional Reporting Ratio（PRR）は，対象としている有害事象の報告割合の比として定義される．つまり，

$$PRR=\frac{x_{11}/(x_{11}+x_{12})}{x_{21}/(x_{21}+x_{22})} \tag{1}$$

である（Evans (2001)）．通常，データベースに含まれる医薬品の数はかなり多いので，対象とする医薬品の報告度数は全体に比べるととても小さい．このため，PRR は以下のように計算される場合もある．

$$PRR=\frac{x_{11}/(x_{11}+x_{12})}{(x_{11}+x_{21})/N}=\frac{x_{11}/N}{(x_{11}+x_{21})/N\times(x_{11}+x_{12})/N} \tag{2}$$

Beta (1998) は，x_{11}, (x_{11}, x_{21}), (x_{11}, x_{12}) がそれぞれ二項分布に従うとして，それぞれの二項分布の割合のパラメータ $p_{x_{11}}, p_{(x_{11}+x_{21})}, p_{(x_{11}+x_{12})}$ で（2）式の PRR を表したものに対数をとった指標（IC）を提案した．

$$IC=\log_2\frac{p_{x_{11}}}{p_{(x_{11}+x_{21})}\times p_{(x_{11}+x_{12})}} \tag{3}$$

この方法は Bayesian Confidence Propagation Neural Network（BCPNN）と呼ばれており，現在 World Health Organization（WHO）の Uppsala Monitoring Center で用いられている．BCPNN ではパラメータがベイズ流に推定される．その事前分布には無情報事前分布を持つように構成されている（Gould, 2003）．

指標としては PRR の他に，報告オッズ比（ROR）も指標として提案されている（例えば van Puijenbroek ら（2003）参照）．

$$ROR=\frac{x_{11}\times x_{22}}{x_{12}\times x_{21}} \tag{4}$$

この他にもいくつかの方法が提案されているが（van Puijenbroek ら（2002）），主に用いられているのは PRR や ROR と BCPNN のようなそれらの変法である．PRR や ROR の検出の基準としては，指標の信頼区間の下限（BCPNN ではベイズ区間の下限）が特定の値を超えた場合であったり，指標が特定の値となりかつ報告数が一定数以上ある場合であったりする．

Rothman (2004) は，PRR と ROR について議論をしている．彼は表1に示されるようなデータはケースコントロール研究のような視点でとらえるべきであるので PRR よりも ROR が望ましいと述べている．本研究でも ROR を用いて指標の構築を考えることにする．

表1　医薬品と有害事象の 2 × 2 分割表

	対象となる有害事象	それ以外の有害事象	合計
対象となる医薬品	x_{11}	x_{12}	$x_{11}+x_{12}$
それ以外の医薬品	x_{21}	x_{22}	$x_{21}+x_{22}$
合　計	$x_{11}+x_{21}$	$x_{12}+x_{22}$	N

2.2 検討する方法

上記のシグナル検出の指標は，いずれも興味のある薬剤と有害事象の有無で構成される2×2分割表からのみ算出されるので，特に重篤か否かということは考慮されていない．しかし，対象とする有害事象が重篤か否かはわかるので，少なくともこの定性的な情報は利用できるはずである．ここでは，先に述べたような状態を監視するという観点から重篤な有害事象は早く検出したいという要求に，重篤か否かという情報を利用することを考える．そこで，ベイズ統計学の枠組みでこの問題を考えることにし，重篤か否かという情報を異なる事前分布として反映させることにする．

表1のデータ $\boldsymbol{x}=(x_{11}, x_{12}, x_{21}, x_{22})$ がそれぞれパラメータ $\boldsymbol{p}=(p_{11}, p_{12}, p_{21}, p_{22})$ の多項分布に従っているとする．つまり

$$\boldsymbol{x}|\boldsymbol{p}\sim Multin(N, \boldsymbol{p}) \tag{5}$$

とする．そして，パラメータ $\boldsymbol{p}$ は，共役事前分布である超パラメータ $\boldsymbol{\alpha}^{(0)}=(\alpha_{11}^{(0)}, \alpha_{12}^{(0)}, \alpha_{21}^{(0)}, \alpha_{22}^{(0)})$ のディレクレ分布に従っているとする．つまり

$$\boldsymbol{p}|\boldsymbol{\alpha}^{(0)}\sim Dirichlet(\boldsymbol{\alpha}^{(0)}) \tag{6}$$

とする．ベイズの定理により，パラメータ $\boldsymbol{p}$ の事後分布は

$$\boldsymbol{p}|\boldsymbol{x}, \boldsymbol{\alpha}^{(0)}\sim Dirichlet(\boldsymbol{\alpha}) \tag{7}$$

ただし，$\boldsymbol{\alpha}=(\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22})=(\alpha_{11}^{(0)}+x_{11}, \alpha_{12}^{(0)}+x_{12}, \alpha_{21}^{(0)}+x_{21}, \alpha_{22}^{(0)}+x_{22})$ となる．

今，シグナル検出の指標としてRORをデータではなくてパラメータで表現することにする．ここではこれを ψ とすると

$$\psi=\frac{p_{11}\times p_{22}}{p_{12}\times p_{21}} \tag{8}$$

である．この指標の点推定値を $\hat{\psi}$ とすると，

$$\hat{\psi}=\frac{\alpha_{11}\times\alpha_{22}}{\alpha_{12}\times\alpha_{21}}=\frac{(x_{11}+\alpha_{11}^{(0)})\times(x_{22}+\alpha_{22}^{(0)})}{(x_{12}+\alpha_{12}^{(0)})\times(x_{21}+\alpha_{21}^{(0)})}$$

として得ることができる．$\hat{\psi}$ の分布は（8）のサンプリングにより求めることができるが，データが多い場合には以下の近似式（Lindley, 1964）を用いることができるであろう．

$$\ln(\psi)\sim N\left(\ln(\alpha_{11})-\ln(\alpha_{12})-\ln(\alpha_{21})-\ln(\alpha_{22}), \frac{1}{\alpha_{11}}+\frac{1}{\alpha_{12}}+\frac{1}{\alpha_{21}}+\frac{1}{\alpha_{22}}\right) \tag{9}$$

この方法を用いるには，超パラメータ $\boldsymbol{\alpha}^{(0)}=(\alpha_{11}^{(0)}, \alpha_{12}^{(0)}, \alpha_{21}^{(0)}, \alpha_{22}^{(0)})$ の値を決める必要がある．ここでは，対象としている医薬品が重篤か否かでこの値を変えることにする．もし，対象としている医薬品が重篤ではない場合には，従来のシグナル検出の方法と同程度の性能を有することにすると，事前分布は無情報事前分布を用いることが望ましいであろう．よく用いられる無情報事前分布として $\boldsymbol{\alpha}_{11}^{(0)}=\alpha_{12}^{(0)}=\alpha_{21}^{(0)}=\alpha_{22}^{(0)}=0$ または1がある（Gelmanら（2004））．もしも0とすれば $\hat{\psi}$ は（4）と一致する．しかし，分割表のセル度数が小さいときに（4）で計算されるオッズ比にはバイアスが生じることが知られている．本研究では検出の速さを扱いたいので，報告数の少ないところで検出できるかどうかが検討の対象となる．このような状況でバイアスを改善する方法として，

$$\frac{(x_{11}+0.5)\times(x_{22}+0.5)}{(x_{12}+0.5)\times(x_{21}+0.5)}$$

がよく用いられている（Jewell, 2004）．これは $\alpha_{11}^{(0)}=\alpha_{12}^{(0)}=\alpha_{21}^{(0)}=\alpha_{22}^{(0)}=0.5$ としたときの $\hat{\phi}$ に一致するので，有害事象が重篤ではない場合にはこれを用いることにする．

次に対象とする報告が重篤の場合を考える．得られる情報は重篤か否かということだけなので，これ以上の情報がなく，4つのパラメータを決めることはできない．ここでは，単純な設定として $\alpha_{12}^{(0)}=\alpha_{21}^{(0)}=\alpha_{22}^{(0)}=0.5$ とし，残りのパラメータである $\alpha_{11}^{(0)}$ を0.5よりも大きな値に変えることにする．この処置はかなり恣意的であるが，相対的に α_{11} が大きくなるので，結果として報告数の少ないところでは $\hat{\phi}$ の値を大きくすることになる．これらの値をいくつに設定すべきかは検討が必要である．

対象とする薬剤の報告が十分集まれば，$\boldsymbol{\alpha}^{(0)}=(\alpha_{11}^{(0)},\alpha_{12}^{(0)},\alpha_{21}^{(0)},\alpha_{22}^{(0)})$ の効果は相対的に小さくなるので，（4）式に近づくことになる．

2.3　シミュレーションによる性能の評価

他の指標に比べて検討する方法の検出の速さがどの程度か，検出は安定性しているのかを調べる目的で，有害事象が重篤である場合とそうではない場合についてシミュレーションを行った．

比較対照の方法と検討する方法の事前分布のパラメータの値

比較する対照とする指標として，計算の容易さと実際に使用されている指標という点からBCPNNを用いることにした．検討する方法の事前分布のパラメータ $\alpha_{11}^{(0)}$ の値は，0.5，1，2，3とした．0.5の場合が重篤ではない場合に相当することになる．

検出基準

検出の基準は，検討する方法では（8）式の95%ベイズ区間を下限が2を越えた場合とし，BCPNNでは 2^{IC} の95%ベイズ区間の下限が2を越えた場合とした．Gould（2003）はBCPNNの正確な区間の計算方法を与えており，ここではその計算法を採用した．検討する方法のベイズ区間は，対象としている医薬品の対象としている有害事象の報告数が10を満たない場合は事後分布の100,000回のサンプリングにより算出したが，それ以降は（9）式の近似式を用いた．

シミュレーション条件

シミュレーション条件は以下のとおりである．はじめの時点で対象となる医薬品の報告は0であるとした．医薬品対象とする医薬品の使用患者数を5,000,000人，その他の医薬品の使用患者数を500,000,000人と想定した．このうち実際に有害事象が生じかつ報告される割合は，対象とする医薬品で対象とする有害事象に関しては0.00002，0.000025，0.00004，0.00006，その他の医薬品の対象とする有害事象については0.00002，対象とする医薬品とその他の医薬品におけるその他の有害事象に関してはそれぞれ0.0002とした．つまり真の報告割合の比がそれぞれ1，1.25，2，3となるように設定した．

その他の医薬品については，すでにデータベースに多くの報告があるであろうから，使用者数と真の報告割合をパラメータとする二項乱数を発生させ，これをはじめにデータベースに含まれる人数とした．次いで，上記の報告割合に基づきランダムに2×2の分割表のいずれかのセルに報告割合に基づく乱数を発生させた．そして，対象となる医薬品の対象となる有害事象の報告が報告されるたびにその時点での分割表が作成されるようにして，その都度シグナル検出の指標を計算した．この試行を1,000回繰り返した．

評価の方法

速さの評価は以下のように行うことにする．上記の条件でシミュレーションを行った場合，対象とする医薬品について，例えば対象とする有害事象の報告割合が0.00004の場合，1回の試行で，報告される回数は，0.00004×5,000,000＝200回程度が期待される．各回について指標を算出して，最初に基準値を超えたのは何報告目であるのかで評価することにした．

シグナルの検出では，一度基準値を超えてしまえばシグナルと判定されることになるので，指標が安定していることも重要であろう．検討する方法では少ない報告のところで検出を試みようとしているので指標が不安定になると思われる．そこでこれがどの程度になるのかを，各回の試行で得られる検出指標の変動係数を求めることにした．

3. 結　　果

図1はシミュレーションのある1回の試行の結果例である（真の報告割合の比が3の場合）．横軸は対象とする医薬品の対象とする有害事象を検出したときまでに集められた15報告までの報告数であり，縦軸は2^{IC}と$\alpha_{11}^{(0)}$の値を0.5，1，2，3と変えたときの$\hat{\phi}$，そしてそれらの95％ベイズ区間である．4本の線は破線がBCPNN，実線が検討する方法を示している．検討する方法では線が太くなるにしたがって，$\alpha_{11}^{(0)}$の値は0.5，1，2，3と大きくなることを意味している．

この図から事前分布の影響として$\alpha_{11}^{(0)}$の値が大きいほど，指標は不安定になることがわかる．また，報告度数が増すに従い指標は安定してくることがわかる．

図2のa）からd）は，真の報告割合別に，対象とする医薬品の対象とする有害事象を検出したときまでに集められた報告数を横軸に，カプラン・マイヤー法で求めた累積分布を縦軸に示している．図中で○印は，発生した乱数の範囲でその試行の最後までに検出できずに打ち切りになっていることを示している．表2はカプラン・マイヤー法により求めた真の報告割合比におけるそれぞれの方法の最終的な検出割合（％）を示している．真の報告割合比が1のとき

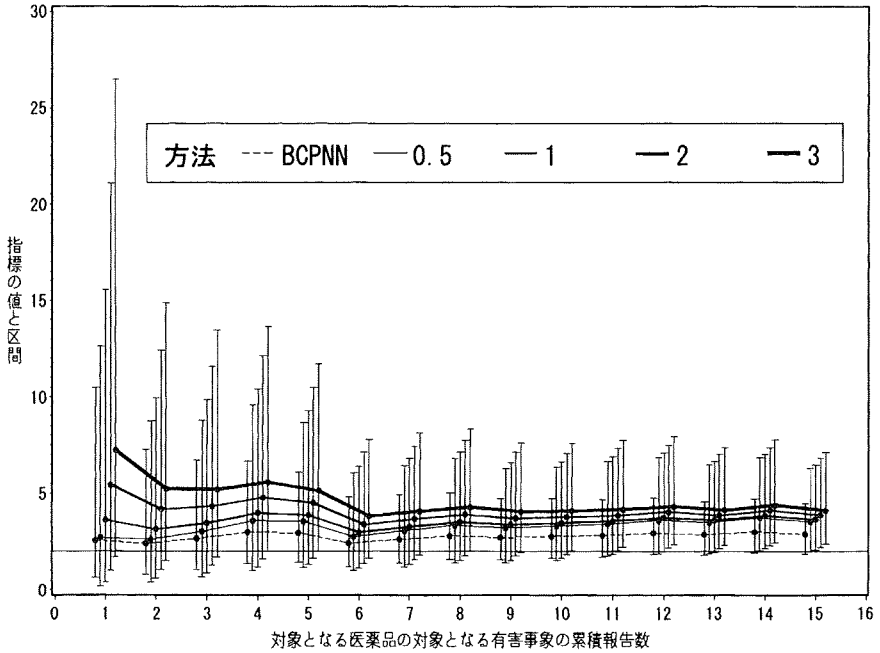

図1　各指標の値と95％ベイズ区間
真の報告割合比が3のときのある1回の試行結果

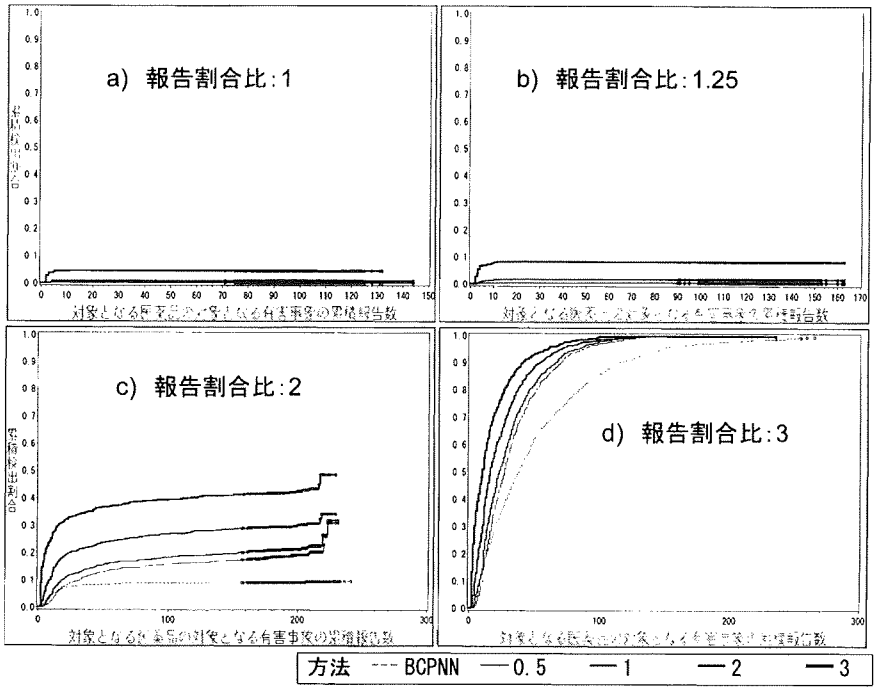

図2　報告数と累積検出割合

表2　真の報告割合比別のシグナルの検出割合（%）
（カプラン・マイヤー法による）

真の報告割合比	BCPNN	提案法の $\alpha_{11}^{(0)}$			
		0.5	1	2	3
1	0.3	0.0	0.2	0.9	4.5
1.25	0.4	0.4	0.5	1.8	8.2
2	9.5	9.5	31.1	34.4	48.7
3	99.5	99.9	99.9	99.9	99.9

のこの検出割合は偽検出を意味することになる．

図2および表2から $\alpha_{11}^{(0)}$ が3としたときに検出は速くなるが，偽検出の割合も増えていることがわかる．このシミュレーション結果では最終的に4.5%の偽検出となっており，それは初期に検出されている．真の報告割合比が1.25のときには，この設定では $\alpha_{11}^{(0)}$ が3以外ではほとんど検出できていない．このことは，対象とする医薬品の有害事象の発生割合がわずかに高い程度では，なかなか検出できないということが言えるであろう．真の報告割合比が2のときには検討する方法が $\alpha_{11}^{(0)}$ の大きさに依存して速く検出できている．しかし，そうであっても150から200程度の報告が集まっても，ここで設定した範囲の報告数ではシグナルとして検出される確率は50%以下となっている．真の報告割合比が3のときには，すべての方法で検出できたが，検討する方法では30報告数で60%以上程度の検出確率を得ることができている．一方BCPNNで60%を得るには40報告程度が必要となり検出は遅れていることが分かる．

表3は各試行において得られた指標の変動係数の平均値を示している．予想されるように $\alpha_{11}^{(0)}$ の値が大きくなるほど変動係数の値は大きくなり，不安定である．$\alpha_{11}^{(0)}=0.5$ とした場合でも，検討する方法の方がBCPNNに比べて若干この値は大きくなっている．真の報告割合比が小さいほどこの値が大きくなる傾向があるが，これは，その場合には対象とする医薬品の対

表3 各試行のシグナル検出指標の変動係数（%）（100回の平均値）

真の報告割合比	BCPNN	提案法の $\alpha_{11}^{(0)}$			
		0.5	1	2	3
1	15.2	17.0	19.9	26.9	34.4
1.25	12.7	14.5	16.2	21.3	27.1
2	9.7	11.8	12.4	15.2	18.9
3	7.5	9.7	9.6	10.6	12.7

象となる有害事象の期待報告度数が小さいことにより，報告数があまり蓄積されず指標が安定していない部分が強く反映されているからであろう．

全体としては，$\alpha_{11}^{(0)}=0.5$ のときには真の報告割合比が1，1.25程度ならばBCPNNと同じくらいの性能であり，真の報告割合比が2，3のときには検討する方法の方が速めの検出ができるようであった．$\alpha_{11}^{(0)}$ の値を1，2，3としたときには，値が増加するに従い速く検出ができることが確認できた．

4. 議　　論

本稿では，有害事象の報告により構成されるデータベースから，シグナルを検出するための方法について論じた．筆者は，副作用報告というは状態監視という観点から医薬品の安全性を監視するという制度であると考えている．この観点からするとシグナルに対する検出の速さが問題となる．そのような視点から本稿で行ったような検討の研究は十分には行われてはこなかった．

本検討では，重篤な有害事象は速く検出するという方針で，ベイズ統計学の方法により指標を構成した．このようなアプローチを採用しなくても，重篤な有害事象の場合には検出の基準値を変えるなどの方法も考えられる．しかし，そのようなことを行うと適当な報告数が集まった後の偽検出も増えてしまうであろう．一方，ベイズ流のアプローチではデータが蓄積されれば事前の分布の影響が少なくなるので，一律に感度を上げることはない．つまり，報告数が少ない場合にはなるべく疑うようにするという方針を表現しているといえる．このことは，Szarfmanら（2002）の検討で，シグナル検出法が重篤な有害事象の検出が既存の方法に比べて遅かった点を改善していることになるであろう．

本稿での考えている状況は，事象を系列で捉え，ある時点で事前に定めた基準値を超えるか否かを判定するということである．これはSequential probability ratio test（SPRT）で扱われる状況と似ている．現に，最近では安全性評価にSPRTを利用するという提案もなされている（Griggら（2003），Spiegelhalterら（2003），Davisら（2005））．しかし，安全性評価にSPRTを用いる場合には，帰無仮説をどのように設定するかという点が実用上の問題となる．Evans（2004）は，SPRTとPRRの関係についてわかりやすい解説をしている．

本稿では，方法の性能を把握するために特定の状況を想定したシミュレーションを行っているが，現実の有害事象の報告のデータはより複雑である．実際の状況の難しさは，渡邉ら（2004）が詳細に解説している．難しさの一つは，自発的な有害事象の報告を集めたデータベースを用いていることによっている．このようなデータベースによる安全性評価の限界は，計算している指標が疫学研究で用いられているリスク比やオッズ比とは異なることである．有害事象が生じても報告がされないことがあるので個々のリスクの正確な分母の数はわからないだけでなく，意図的であろうとなかろうと特定の医薬品についての報告が多くなされるなどのこと

が生じた場合には指標の値は当然偏ることになるからである．このようなデータベースから得られる指標からわかることは効果ではなくシグナルであり，シグナルを検出した後には副作用であるかどうかを評価する段階が必要となる．

本稿で用いた方法は，重篤な有害事象とそうではない有害事象との区別を行っている．しかし，重篤か否かに限らなくても，疑いをもつべきものを早期に検出するということが必要な場合にはこの考え方が適用できる．例えば，添付文書に反映されていない有害事象とそうでない有害事象とすることや，未知の有害事象か既知の有害事象かというような場合にも適用できるであろう．

本研究の限界はいくつかある．一つは事前分布のパラメータを設定するための情報がないことである．検討した方法では，単純にパラメータの1つの値を変えただけであるが，これについての適切な理由はない．また，現実にその値としてどの程度の大きさが望ましいのかはこの検討では十分には明らかにはできなかった．実際にこの方法を用いる場合には何らかの指針が必要になると思われるが，検討したパラメータが早く検出したいという意図の強さを表す以外の意味を示すことはできていない．限界のもう一点は検討の範囲である．シミュレーションの設定が，どの程度現実の実態を反映させているのかが明確ではない．現実には，ある有害事象が生じてそれが報告される割合は医薬品と有害事象の種類に影響されるであろう．それがどの程度かを決めるのは難しく，実態を把握することは困難である．本検討ではわずか1つ結果を報告しているが，探索的に行なった規模の小さな検討（結果は表示していない）では，検出割合は有害事象が生じてかつ報告される割合により大きく変わることがわかった．従って，本検討が示した結果はあくまでも設定に依存することになる．しかしながら，$\alpha_{11}^{(0)}$ の値が増加するにつれて，報告数の少ない初期の時点での検出割合は大きくなることには変わりはない．

本稿で検討した方法は，重篤な有害事象は重篤でない有害事象に比べて早く検出するという方法である．重篤ではない場合には従来の方法と同程度の性能で検出できることになる．このことは，検討した方法が，対象としている有害事象が重篤な場合には早く検出できるように偏った推定を行なっていることにほかならない．しかしながら，そのような方針こそが，製造販売前の承認の段階の限られたデータではチェックできなかった重篤な有害事象についてシグナルを検出するということに求められていることなのではないだろうか．

謝辞

論文の不備のご指摘と貴重なご意見をいただきました査読者に感謝いたします．なお，本研究は科学研究費補助金基盤研究（A）No. 16200022 の助成により実施された．

参考文献

Bate, A., Lindquist, M., Edwards, I. R., Olsson, S., Orre, R., Lansner, A., De Freitas, R. M. (1998). A Bayesian neural network method for adverse drug reaction signal generation, *European Journal of Clinical Pharmacology*, **54**, 315-321.

Bate, A., Lindquist, M., Edwards, I. R. and Orre, R. (2002). A data mining approach for signal detection and analysis, *Drug Safety*, 25, 393-397.

Davis, R. L., Kolczak, M., Lewis, E., Nordin, J., Goodman, M., Shay, D. K., Platt, R., Black, S., Shinefield, H. and Chen, R. T. (2005). Active surveillance of vaccine safety: A system to detect early signs of adverse events, *Epidemiology*, **16**, 336-341.

Evans, S. J. W. (2004). Statistics: Analysis and presentation of safety data, in *Stephens' Detection of New Adverse Drug Reactions* (5th. Ed.), edited by J. Talbot and P. Waller, 301-328.

Evance, S. J. W., Waller, P. C. and Davis, S. (2001). Use of proportional reporting ratios (PRRs) for signal generation from spontaneous adverse drug reaction reports, *Pharmacoepidemiology and Drug Safety*, **10**, 483-486.

Gelman, A., Carin, J. B., Stern, H. S. and Rubin, D. B. (2004). *Bayesian Data Analysis* (2nd. Ed.), Chapman & Hall/CR.

Gould, A. L. (2003). Practical pharmacovigilance analysis strategies, *Pharmacoepidemiology and Drug Safety*, **12**, 559–574.

Grigg, O. A., Farewell, V. T. and Spiegelhalter, D. J. (2003). *Statistical Methods in Medical Research*, **12**, 147–170.

Jewell, N. P. (2004). *Statistics for Epidemiology*, Chapman & Hall/CR.

Lee, P. M. (2004). *Bayesian Statistics, An Introduction* (3rd. Ed.), Oxford University Press Inc.

Lindley, D. V. (1964). The Bayesian analysis of contingency tables, *The Annals of Mathematical Statistics*, **35**, 1622–1643.

Rothman, K. J., Lanes, S. and Sacks, S. T. (2004). The reporting odds ratio and its advantages over the proportional reporting ratio, Pharmacoepidemiology and *Drug Safety*, **13**, 519–523.

Szarfman, A., Machado, S. G. and O'Neill, R. T. (2002). Use of screening algorithms and computer systems to efficiently signal higher-than-expected combinations of drugs and events in the US FDA's spontaneous reports database, *Drug Safety*, **25**, 381–392.

Spigelhalter, D., Grigg, O., Kinsman, R. and Treasure, T. (2003). Risk-adjusted sequential probability ratio tests: applications to Bristrol, Shipman and adult cardiac surgery, *International Journal for Quality in Health Care*, **15**, 7–13.

van Puijenbroek E. P., Bate, A., Leufkens, H. G. M., Lindquist, M., Orre, R. and Egberts A. C. G. (2002). A comparison of measures of disproportionality for signal detection in spontaneous reporting systems for adverse drug reactions, *Pharmacoepidemiology and Drug Safety*, **11**, 3–10.

渡邉裕之，松下泰之，渡辺　篤，前田敏郎，温井一彦，小川嘉正，澤　淳悟，前田　博（2004）．重要な安全性情報を早期に検出する仕組み，計量生物学，**25**, 37–60.

ORIGINAL ARTICLE

Variance of the Stimulation Index for the Local Lymph Node Assay

Takashi Omori[1] and Takashi Sozu[2]

[1]*Department of Biostatistics, Kyoto University School of Public Health, Kyoto, Japan*
[2]*The Center for Advanced Medical Engineering and Informatics, Osaka University, Osaka, Japan*

ORIGINAL ARTICLE

Variance of the Stimulation Index for the Local Lymph Node Assay

Takashi Omori[1] and Takashi Sozu[2]

[1]*Department of Biostatistics, Kyoto University School of Public Health, Kyoto, Japan*
[2]*The Center for Advanced Medical Engineering and Informatics, Osaka University, Osaka, Japan*

Abstract

The local lymph node assay (LLNA) is a well-established method for assessing skin sensitization; it involves the use of mice and provides an alternative to tests requiring guinea pigs. In this assay, the endpoint of interest is the stimulation index (SI) that is defined as the mean disintegrations per minute (DPM)/mouse for the chemical-treated group divided by that for the vehicle control group. Therefore, the SI is a ratio of 2 variables. However, some researchers often ignore the variation in the vehicle control group while assessing chemicals based on the SI. This approach, which we refer to as the ignorance approach, may underestimate the variation in SI. We developed an alternative approach by the delta method, and we evaluated the performance of both approaches through a simulation and a numerical examination. We found that the ignorance approach is not acceptable for estimating the variance of SI because the value obtained is severely biased toward underestimation. Instead, analysis by using the delta method approach is recommended. However, it should be noted that this approach is also slightly biased in the case of large variations in the DPM/mouse values in the vehicle control group.

Key words: *LLNA, stimulation index, statistical analysis, variance, Delta method*

1. Introduction

The local lymph node assay (LLNA) is a well-established method for assessing skin sensitization; it involves the use of mice and provides an alternative to tests requiring guinea pigs. The LLNA has been accepted by the US Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) as a stand-alone alternative to the currently used guinea pig tests. It has also been recognized as a stand-alone test for skin sensitization by the competent authorities of EU member states (Basketter et al., 2005). The test guidelines of the Organisation for Economic Co-operation and Development for the LLNA are available as OECD TG 429 (OECD, 2002).In the LLNA, ^{3}H-methyl thymidine incorporation is measured by β-scintillation counting in terms of disintegrations per minute (DPM). The stimulation index (SI) is a measure for judging whether the response to the tested chemical is positive or negative; it is calculated as the mean DPM/mouse for the chemical-treated group divided by that for the vehicle control group. That is, the SI value is a ratio of 2 variables. In order to assess variations in the SI, some researchers ignore the variation in the vehicle control group and calculate the SI variance by dividing each DPM/mouse value obtained for each chemical-treated group by the mean DPM/mouse value obtained for the vehicle control group. We refer to this approach as the "ignorance approach." The variance of SI obtained by this approach may be underestimated, particularly for chemicals that produce slight stimulation, because it does not take into account variations in the vehicle control group.