Table 1. Proportion of IGRA and TST positivity, stratified by age.

		Non-TB hospital		TB hospital		*p* value
		No. positive/No. tested	(%)	No. positive/No. tested	(%)	
IGRA						
	20–29*	6/26	(23.1)	21/55	(38.2)	
	30–39	4/13	(30.8)	23/48	(47.9)	
	40–49	43/87	(49.4)	20/39	(51.3)	
	≥50	18/24	(75.0)	7/8	(87.5)	
	Combined**					0.13
One-step TST, ≥10 mm						
	20–29*	7/26	(26.9)	28/51	(54.9)	
	30–39	5/13	(38.5)	32/44	(72.7)	
	40–49	47/86	(54.7)	32/37	(86.5)	
	≥50	19/24	(79.2)	6/7	(85.7)	
	Combined**					<0.0001
One-step TST, ≥15 mm						
	20–29*	2/26	(7.7)	5/51	(9.8)	
	30–39	5/13	(38.5)	15/44	(34.1)	
	40–49	27/86	(31.4)	17/37	(45.9)	
	≥50	12/24	(50.0)	2/7	(28.6)	
	Combined**					0.45
Two-step TST, ≥10 mm						
	20–29*	8/26	(30.8)	28/51	(54.9)	
	30–39	5/13	(38.5)	33/44	(75.0)	
	40–49	60/86	(69.8)	32/37	(86.5)	
	≥50	19/24	(79.2)	6/7	(85.7)	
	Combined**					0.0004

TB: Tuberculosis; IGRA: Interferon-gamma release assay; TST: Tuberculin skin test.
*Years old.
**Mantel-Haenszel test for stratified data.
doi:10.1371/journal.pone.0006798.t001

Table 2. Agreement between IGRA and TST using different cut-off values of TST.

		One-step TST						Two-step TST
		≥10 (mm)*			≥11 (mm)*	≥13 (mm)*	≥15 (mm)*	≥10 (mm)*
			BCG (−)**	BCG (+)**				
TST+/IGRA+ (n)		114	44	39	102	86	69	119
TST+/IGRA- (n)		49	14	27	30	21	13	58
TST-/IGRA+ (n)		21	8	8	33	49	66	16
TST-/IGRA- (n)		71	33	23	90	99	107	62
Agreement, %		72.5	77.8	63.9	75.3	72.5	69.0	71.0
Kappa (SE)		0.44 (0.06)	0.55 (0.10)	0.29 (0.09)	0.50 (0.06)	0.46 (0.06)	0.39 (0.06)	0.41 (0.06)
Symmetry test***	Chi-squared value	11.2	1.64	10.3	0.14	11.2	35.6	23.8
	p value	0.0008	0.20	0.0013	0.71	0.0008	<0.0001	<0.0001

IGRA: Interferon-gamma release assay; TST: Tuberculin skin test; SE: Standard error; BCG (−): Without history of BCG vaccination; BCG (+): With history of BCG vaccination;
*N = 255; all subjects with valid data sets.
**N = 99 and 97 for BCG (−) and BCG (+) groups, respectively.
***Equivalent to McNemar test for evaluation of the symmetry of TST+/IGRA- and TST-/IGRA+.
doi:10.1371/journal.pone.0006798.t002

duration were included in the model, BCG was the only parameter showing significant association with this discordance (OR = 2.26 [95%CI, 1.09–4.71]) (Table not shown). No factors analyzed in this study showed association with TST-/IGRA+ discordance.

Factors associated with IGRA positivity

We tried to identify factors associated with IGRA positivity. In univariate analysis, non-occupational factors such as age and the high body mass index (BMI) were significantly associated with having a positive IGRA result (OR = 1.05 [95%CI, 1.02–1.07] per one year and OR = 5.10 [95%CI, 1.45–17.99], respectively) (Table 3), whereas occupational factors including job category, working duration, and mask use did not show significant associations with IGRA positivity (Table 3).

In multivariate analysis, significantly increased odds of IGRA positivity were observed with non-occupational factors such as increase in age (OR = 1.06 [95%CI, 1.00–1.11]), high BMI (OR = 4.18 [95%CI, 1.14–15.36]), education lower or equal to high school level (OR = 4.28 [95%CI, 1.28–14.27]) and pre-university level (OR = 3.54 [95%CI, 1.18–10.59]). Among occupational factors tested, working in TB hospital was the only parameter showing the significant association (OR = 1.94 [95%CI, 1.04–3.64]) (Table 3).

Discussion

Our study demonstrated the high prevalence of latent TB infection estimated by either TST or IGRA positivity among hospital workers and higher risk of infection adjusted for age and other factors in the TB hospital than in a general hospital in Hanoi, Viet Nam. Disagreement between TST and IGRA positivity was largely affected by BCG vaccination history and it was not improved by changing cut-off values of TST. As far as we know, this is the first report on TB infection among HCWs evaluated by IGRA in Southeast Asia.

The overall prevalence of IGRA positivity among HCWs in our study (47.3%) is high and comparable to previous estimates from India, Russia and Georgia (40.0%, 40.8%, and 60.0% respectively) [17–19]. Direct comparison is difficult among the studies, because in the previous studies particularly the Russian one, detailed information about age strata has not been shown, which strongly affects the prevalence of TB infection.

The prevalence of TST positivity in our study population was higher than that of IGRA. High false-positive TST reaction due to BCG vaccination given after infancy has been reported [7,20], especially in individuals less than 40 years old [21]. In fact, the degree of TST/IGRA agreement was low in the group with BCG vaccination in our study, with a significant disproportional increase in TST+/IGRA- over TST-/IGRA+. Furthermore,

Table 3. Logistic regression analysis results for the associations between potential risk factors and IGRA positivity (n = 265).

		Proportion of positive results		Uni-variate		Multi-variate	
		n	(%)	Odds Ratio	(95%CI)	Odds Ratio	(95%CI)
Non-occupational factors:							
Age	/year	NA*	NA*	1.05	(1.02–1.07)	1.06	(1.00–1.11)
Sex	Female	102/197	(51.8)	1.00	(reference)	1.00	(reference)
	Male	40/68	(58.8)	1.33	(0.76–2.32)	1.10	(0.56–2.16)
BMI	18.5≤<25.0	114/223	(51.1)	1.00	(reference)	1.00	(reference)
	<18.5	12/23	(52.2)	1.04	(0.44–2.46)	1.50	(0.57–3.94)
	25.0≤	16/19	(84.2)	5.10	(1.45–17.99)	4.18	(1.14–15.36)
Education	University and higher	47/93	(50.5)	1.00	(reference)	1.00	(reference)
	High school and lower	25/36	(69.4)	2.22	(0.98–5.04)	4.28	(1.28–14.27)
	Pre-university	70/136	(51.5)	1.04	(0.61–1.76)	3.54	(1.18–10.59)
Occupational factors:							
Hospital	Non-TB	71/136	(52.2)	1.00	(reference)	1.00	(reference)
	TB	71/129	(55.0)	1.12	(0.69–1.82)	1.94	(1.04–3.64)
Job	Others	45/74	(60.8)	1.00	(reference)	1.00	(reference)
	Doctor	38/66	(57.6)	0.88	(0.45–1.72)	2.60	(0.82–8.29)
	Nurse	45/98	(45.9)	0.55	(0.30–1.01)	0.78	(0.28–2.19)
	Technician	14/27	(51.9)	0.69	(0.29–1.69)	1.02	(0.31–3.35)
Working years	<2	12/29	(41.4)	1.00	(reference)	1.00	(reference)
	2≤<5	20/47	(42.6)	1.05	(0.41–2.68)	0.94	(0.34–2.58)
	5≤<10	22/44	(50.0)	1.42	(0.55–3.65)	0.85	(0.28–2.56)
	10≤	88/145	(60.7)	2.19	(0.97–4.92)	0.91	(0.27–3.13)
Mask use	Frequently	65/124	(52.4)	1.00	(reference)	1.00	(reference)
	Occasionally	40/82	(48.8)	0.86	(0.50–1.51)	1.02	(0.55–1.88)
	Rarely/never	37/59	(62.7)	1.53	(0.81–2.88)	1.78	(0.81–3.94)

*NA = Not applicable.

doi:10.1371/journal.pone.0006798.t003

among parameters tested in our study, BCG history was the only factor to be associated with TST+/IGRA-discordance in univariate and multivariate analysis. Our finding is consistent with the previous reports [10,11], but different from that of another recent study, where BCG did not account for this discordance [22]. This may be simply due to difference in age of BCG vaccination. Involvement of other unknown factors for the discordance cannot be excluded. Exposure to nontuberculous mycobacterium might be another factor for TST+/IGRA- discordance [11], although nontuberculous mycobacterium is rarely found among smear-positive patients in Viet Nam (unpublished data).

These findings indicate that IGRA is more advantageous than TST with different cut-off values [23]. In Viet Nam, BCG vaccination has been included in the Extended Program of Immunization since 1986 and given within one month after birth. Before this point of time, there were no national guidelines and BCG vaccination was sporadically implemented in several areas and mostly given during childhood. In the heterogeneous background of BCG vaccination, it seems difficult to interpret TST result of the present Vietnamese HCWs even with a higher cut-off value as recommended elsewhere [21,24]. High agreement level between TST and IGRA in a study from India [17] is probably attributed to the fact that most of their participants were vaccinated at birth.

On the other hand, our study did not find significant associations between BCG history and TST-/IGRA+ discordant results and this finding is consistent with the previous reports [10,11,22]. Age was associated with TST-/IGRA+ in one study [10] but this was not confirmed in our study.

The CDC [3] and others [25] recommend performing a two-step TST on all newly employed HCWs to identify HCWs who have had MTB infection. Two-step TST is known to evoke remote infection, weak response by nontuberculous mycobacteria, past BCG or other factors, while IGRA appears to reflect recent rather than remote MTB infection [23]. In our study, the influence of two-step TST on TST/IGRA discordance was not much different from that of one-step TST.

In the TB hospital, the proportion of young HCWs who should lower the overall IGRA positivity was larger than in non-TB hospital. Despite this fact, the IGRA positivity in TB hospital was not low. Occupational factors as well as non-occupational factors have been expected to be associated with latent TB infection. In multivariate analysis using a logistic regression model, working in the TB hospital was significantly associated with twofold increase in odds of TB infection estimated by IGRA. Although previous studies have shown that occupational factors, such as working duration and job category, confer a risk on IGRA positivity [17–19], our results did not support their data. Working duration is closely related to age and it was difficult to assess its independent effect on TB infection in our study. On the other hand, our data imply that many staff members pursuing a variety of job in the TB hospital might have a considerable chance of exposure to infectious droplet nuclei. While non-TB hospital is a large hospital including an eleven-story building and located in a site with a large yard, the TB hospital is smaller and more enclosed, where all TB patients and HCWs share the same ambulatory route from the entrance (Figure 1). Personal protective equipment used is mostly surgical mask, which cannot prevent the transmission effectively. This finding suggests that the overall working environment and currently used administrative measures should be reconsidered. The cores of infection control programs should be understood deeply to avoid health-care associated infection of TB or MDR-TB at the worst, when a number of MDR-TB patients are hospitalized for treatment.

Non-occupational risk factors for IGRA positivity have been shown in several studies [17,19]. Age reflects cumulative exposure to MTB and it was significantly associated with IGRA positivity in our study. Education levels may indicate potential risk of TB infection in non-working environment as well as high risk of nosocomial infection. Although these two risks were not separately assessed in our study design, training may be necessary to increase awareness of prevention of nosocomial infection towards workers with low educational levels. Possible risk of high BMI for TB infection was unexpected, but the effect was highest among all covariates. In fact, high BMI was associated with both TST and IGRA positivity. TB development associated with diabetes accompanied by overweight is known [26], but the relationship between overweight and TB infection itself has not been reported. The results may have been produced by chance. Another independent investigation is necessary to determine whether it can be reproducible.

Among the HCWs with positive results of IGRA, no one took INH for treatment of latent TB infection. INH treatment is a safe and low-cost intervention and recommended by WHO [27] and others [28]. However, in typical health care facilities of TB high burden countries where the risk of TB exposure is high and continuous, HCWs still doubt significance of one-time INH treatment.

Our study has several limitations. Firstly, we were not able to evaluate the risk of infection from non-working environment as mentioned above, although the prevalence of TB infection could be estimated roughly from the annual risk of TB infection based on the TST surveys using a formula recommended elsewhere [29]. In Viet Nam, the infection rate was too different between areas to estimate it (data not shown). Secondly, we did not measure HIV infection in our study population. According to the data from a household survey in Ho Chi Minh city, estimated prevalence of HIV infection there is 0.7% in 2005 [30]. We assume that the prevalence is lower in Hanoi. Thirdly, it was not possible to identify the cause of indeterminate cases of IGRA. We re-performed ELISA for all preserved plasma samples with indeterminate results and obtained completely the same results. In addition, we have paid careful attention to maintain the high quality of this test [16]. All of the indeterminate cases showed low response to both TB antigen and mitogen and the pattern of TST measurements in IGRA indeterminate cases was similar to that of IGRA negative cases. For this reason, while calculating the IGRA positivity we did not include indeterminate cases in the numerator but did include them in the denominator, although our results might have underestimated the true proportion.

In conclusion, there is a potential high risk of TB infection among HCWs, particularly those working in TB health facilities in a TB high burden country. Prompt attention is necessary to prevent TB infection among HCWs, preparing for recent spread of MDR-TB in resource-limited settings. For this purpose, IGRA seems appropriate to estimate latent TB infection accurately, contributing to improve infection control strategy especially for young vulnerable HCWs who have heterogeneous history of BCG vaccination after birth.

Supporting Information

Table S1 Characteristics of population studied. TB: Tuberculosis. *Others mainly consist of administrative staff and pharmacists.

Found at: doi:10.1371/journal.pone.0006798.s001 (0.12 MB DOC)

Acknowledgments

The authors would like to thank Dr. Dang Van Chinh (Thanh Nhan Hospital) and Dr. Hiroshi Ohara (IMCJ-BMH Medical Collaboration

Center) for supporting management. The authors also thank Dr. Nguyen Phuong Hoang, Dr. Pham Tuan Phuong, Dr. Pham Thu Anh, Dr. Le Thi Hong, Dr. Do Bang Tam, Ms. Vu Thi Xuan Thu (Hanoi Tuberculosis and Lung Disease Hospital), Dr. Phan Thi Minh Ngoc, Ms. Nguyen Thi Thuy Hanh (IMCJ-BMH Medical Collaboration Center) for supporting site implementation. The authors thank Kazuko Tanabe D.V.M. and Mr. John Crosskey for their critical reading of this manuscript.

Author Contributions

Conceived and designed the experiments: LTL NTLH NK HY ET SS PHT VCC NH KH LAT NK. Performed the experiments: NTLH VCC IM. Analyzed the data: NTLH AN TM NK. Contributed reagents/materials/analysis tools: LTL NK HY ET SS PHT IM LAT NK. Wrote the paper: LTL NTLH NK.

References

1. Joshi R, Reingold AL, Menzies D, Pai M (2006) Tuberculosis among health-care workers in low- and middle-income countries: a systematic review. PLoS Med 3: e494.
2. Seidler A, Nienhaus A, Diel R (2005) Review of epidemiological studies on the occupational risk of tuberculosis in low-incidence areas. Respiration 72: 431–446.
3. Jensen PA, Lambert LA, Iademarco MF, Ridzon R (2005) CDC. Guidelines for preventing the transmission of *Mycobacterium tuberculosis* in health-care settings, 2005. MMWR Recomm Rep 54: 1–141.
4. da Costa PA, Trajman A, Mello FC, Goudinho S, Silva MA, et al. (2009) Administrative measures for preventing *Mycobacterium tuberculosis* infection among healthcare workers in a teaching hospital in Rio de Janeiro, Brazil. J Hosp Infect 72: 57–64.
5. Yanai H, Limpakarnjanarat K, Uthaivoravit W, Mastro TD, Mori T, et al. (2003) Risk of *Mycobacterium tuberculosis* infection and disease among health care workers, Chiang Rai, Thailand. Int J Tuberc Lung Dis 7: 36–45.
6. Blumberg HM, Watkins DL, Berschling JD, Antle A, Moore P, et al. (1995) Preventing the nosocomial transmission of tuberculosis. Ann Intern Med 122: 658–663.
7. Farhat M, Greenaway C, Pai M, Menzies D (2006) False-positive tuberculin skin tests: what is the absolute effect of BCG and non-tuberculous mycobacteria? Int J Tuberc Lung Dis 10: 1192–1204.
8. Menzies D, Pai M, Comstock G (2007) Meta-analysis: new tests for the diagnosis of latent tuberculosis infection: areas of uncertainty and recommendations for research. Ann Intern Med 146: 340–354.
9. Pai M, Zwerling A, Menzies D (2008) Systematic review: T-cell-based assays for the diagnosis of latent tuberculosis infection: an update. Ann Intern Med 149: 177–184.
10. Nienhaus A, Schablon A, Diel R (2008) Interferon-gamma release assay for the diagnosis of latent TB infection—analysis of discordant results, when compared to the tuberculin skin test. PLoS One 3: e2665.
11. Mazurek GH, LoBue PA, Daley CL, Bernardo J, Lardizabal AA, et al. (2001) Comparison of a whole-blood interferon gamma assay with tuberculin skin testing for detecting latent *Mycobacterium tuberculosis* infection. JAMA 286: 1740–1747.
12. Mazurek GH, Jereb J, Lobue P, Iademarco MF, Metchock B, et al. (2005) Guidelines for using the QuantiFERON-TB Gold test for detecting *Mycobacterium tuberculosis* infection, United States. MMWR Recomm Rep 54: 49–55.
13. World Health Organization (2008) Anti-Tuberculosis drug resistance in the world WHO/HTM/TB/2008.394. Available: http://whqlibdoc.who.int/hq/2008/WHO_HTM_TB_2008.394_eng.pdf. Accessed 2009 Jul 21.
14. Horie T, Lien LT, Tuan LA, Tuan PL, Sakurada S, et al. (2007) A survey of tuberculosis prevalence in Hanoi, Vietnam. Int J Tuberc Lung Dis 11: 562–566.
15. Huong NT, Duong BD, Co NV, Quy HT, Tung LB, et al. (2006) Tuberculosis epidemiology in six provinces of Vietnam after the introduction of the DOTS strategy. Int J Tuberc Lung Dis 10: 963–969.
16. Hang NT, Ishizuka N, Keicho N, Hong LT, Tam DB, et al. (2009) Quality assessment of an interferon-gamma release assay for tuberculosis infection in a resource-limited setting. BMC Infect Dis 9: 66.
17. Pai M, Gokhale K, Joshi R, Dogra S, Kalantri S, et al. (2005) *Mycobacterium tuberculosis* infection in health care workers in rural India: comparison of a whole-blood interferon gamma assay with tuberculin skin testing. JAMA 293: 2746–2755.
18. Drobniewski F, Balabanova Y, Zakamova E, Nikolayevskyy V, Fedorin I (2007) Rates of latent tuberculosis in health care staff in Russia. PLoS Med 4: e55.
19. Mirtskhulava V, Kempker R, Shields KL, Leonard MK, Tsertsvadze T, et al. (2008) Prevalence and risk factors for latent tuberculosis infection among health care workers in Georgia. Int J Tuberc Lung Dis 12: 513–519.
20. Wang L, Turner MO, Elwood RK, Schulzer M, FitzGerald JM (2002) A meta-analysis of the effect of Bacille Calmette Guérin vaccination on tuberculin skin test measurements. Thorax 57: 804–809.
21. Tissot F, Zanetti G, Francioli P, Zellweger JP, Zysset F (2005) Influence of bacille Calmette-Guérin vaccination on size of tuberculin skin test reaction: to what size? Clin Infect Dis 40: 211–217.
22. Machado A Jr, Emodi K, Takenami I, Finkmoore BC, Barbosa T, et al. (2009) Analysis of discordance between the tuberculin skin test and the interferon-gamma release assay. Int J Tuberc Lung Dis 13: 446–453.
23. Kang YA, Lee HW, Yoon HI, Cho B, Han SK, et al. (2005) Discrepancy between the tuberculin skin test and the whole-blood interferon gamma assay for the diagnosis of latent tuberculosis infection in an intermediate tuberculosis-burden country. JAMA 293: 2756–2761.
24. Do AN, Limpakarnjarat K, Uthaivoravit W, Zuber PL, Korattana S, et al. (1999) Increased risk of *Mycobacterium tuberculosis* infection related to the occupational exposures of health care workers in Chiang Rai, Thailand. Int J Tuberc Lung Dis 3: 377–381.
25. Frenzel EC, Thomas GA, Hanna HA (2006) The importance of two-step tuberculin skin testing for newly employed healthcare workers. Infect Control Hosp Epidemiol 27: 512–514.
26. Leung CC, Lam TH, Chan WM, Yew WW, Ho KS, et al. (2008) Diabetic control and risk of tuberculosis: a cohort study. Am J Epidemiol 167: 1486–1494.
27. World Health Organization (2004) Interim policy on collaborative TB/HIV activities WHO/HTM/TB/2004.330, WHO/HTM/HIV/2004.1. Available: http://whqlibdoc.who.int/hq/2004/who_htm_tb_2004.330.pdf. Accessed 2009 Jul 21.
28. Balcells ME, Thomas SL, Godfrey-Faussett P, Grant AD (2006) Isoniazid preventive therapy and risk for resistant tuberculosis. Emerg Infect Dis 12: 744–751.
29. http://www.who.int/tb/surveillanceworkshop/status_analysis/direct_estimation_prevalence_of_infection.htm. Accessed 2009 Jul 21.
30. Tuan NA, Ha NT, Diep VT, Thang PH, Long NT, et al. (2008) Household survey in two provinces in Viet Nam estimates HIV prevalence in an urban and a rural population. AIDS Res Hum Retroviruses 24: 1017–1026.

The Journal of Toxicological Sciences (J. Toxicol. Sci.)
Vol.34, Special Issue II, SP199-SP200, 2009

Overview: "Children's Toxicology", a renovating study field of irreversible "early exposure-delayed effects"

Jun Kanno

Division of Cellular and Molecular Toxicology, Biological Safety Research Center, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan

(Received February 17, 2009)

ABSTRACT — "Children are not small adults". This is a well-known phrase, especially in the clinics for diagnosis, efficacy of treatment, side effect, and prognosis. However, in the field of toxicology, this issue has long been a challenge. The knowledge has been limited to the differences in metabolism and other physiological factors. Currently available test guidelines for fetuses and immature animals are teratogenicity and reproductive toxicity studies. These tests look for straight-forward (essentially macroscopic) outcomes established within a rather short period of exposure to the test substances. However, recent advances in molecular toxicology allow combination of *in vitro* and *in vivo* studies at molecular levels. The target molecules and receptors can be identified in quantitative fashion and at the fine structure levels around and below the resolution of normal light microscopy. Such expansion of the knowledge lead us to consider a rather new category of "receptor mediated toxicity" or "signal toxicity". Such non-organic insults would merely induce transient effects on adults. However, there are growing evidences that such slight insults on the developing and maturating organisms can leave irreversible effects that become overt in adulthood. As an overview, toxicology has entered a new phase where children's toxicology becomes a renovating study field of the irreversible "early exposure-delayed effects".

Key words: Children's toxicology, Receptor-mediated toxicity, Signal toxicity, Early exposure-delayed effect

INTRODUCTION

Toxicology is a study to analyze interaction between living organisms and xenobiotics, and its final goal is to secure the safety of humans and environment in modern life where various products and technologies are used. Up to now, the majority of toxicological tests to evaluate the toxicity of a particular substance are utilizing experimental animals as a surrogate of humans. The results obtained from such animal tests are extrapolated to humans for the settlement of various kinds of regulation on the test substances, i.e. food additives, pesticides, industrial chemicals, medicines, etc. In cases of pharmaceutical products, clinical trials (human tests) are available. However, these are rather exceptional occasions for toxicology as a whole. It would be very difficult for non-pharmaceutical objects to test on humans, and even for pharmaceuticals, human trial for children including fetuses have many difficulties.

Current toxicological testing protocols are based on an assumption that both experimental animals and humans share common basic structure of the body and thus similar biological reaction. Most of those toxicological studies are based on "diagnosis" of the symptoms of experimental animals in a similar fashion to give a diagnosis to human patients. Because the fine structure and function of the bodies are still unknown, both humans and animal bodies are "black boxes" responding to the test substances by showing various symptoms. Usually, the "no observed adverse effect level" (NOAEL) or "no observed effect level" (NOEL) is given by such tests. Since the basic nature of species differences and individual differences are not known, a number called "safety factor" was invented to extrapolate animal NOAEL/NOEL data to humans (Benford, 2000). Normally, a factor of 10 for the species and another 10 for individual differences, thus 100 as a whole, is used to set lower NOAEL/NOEL levels for humans. This approach has been working well for the majority of test substances. Not surprisingly, however, there are some exceptions. Thalidomide is a best-known example (Newman, 1985). Phocomelia, a spectrum of malformation of limbs, was induced in offspring of tha-

Correspondence: Jun Kanno (E-mail: kanno@nihs.go.jp)

lidomide-treated pregnant women, but not observed in offspring of mice and rats. Therefore, more precise toxicity evaluation/prediction is obviously needed for safer assessment. An approach that enables us to point out molecular mechanisms of toxicity would be essential for such needs and to better understand the species-specific responses.

DISCUSSION

To modernize the toxicology and improve the accuracy of safety assessments, we are attempting to describe and understand the organism-xenobiotics interaction at the molecular levels. Different for other exploratory studies, a major prerequisite is that the Toxicology must be prepared for any unexpected or unpredictable responses. Thus, the approach must be comprehensive. Consequently, we adopted a whole-genome cDNA microarray system for a comprehensive monitoring of the transcriptome, and launched the Percellome Toxicogenomics Project, of which the ultimate goal is to illustrate out the whole regulatory pathways induced by xenobiotics in the experimental animals, mainly mice, including embryo (Kanno *et al.*, 2006).

On top of that, there is an important factor of toxicology, that it the "time frame" such as acute, chronic and delayed toxicity. Among them, researches for the assessment of delayed toxicity targeted for children (including fetus and infants) is becoming very important. It is very likely that the children have a chance to be exposed in daily life to a series of substances which can be a cause of delayed toxicity, especially, of the highly evolved systems, that is endocrine, immune and central nervous system. Such chemical substances can affect the developing systems at a dosage lower than the dosage that induces overt cytocidal changes that would link to immediate appearance of symptoms. For example, our recent experience on the perinatal exposure study (Tanemura *et al.*, 2009) which resulted in the emergence of delayed effects on neurobehavioral endpoints can be explained by a metaphor. That is, "No one turns on power when building a computer, but the living brains are built under the "power-on" situation". It is very likely that the developing brain needs proper or normal signals to build up its fine structures and functional networks (Cohen-Cory, 2002). At this stage, if the signals are disrupted by exogenous insults, it may result in malformation of the fine structure of the brain system. In this case, it is not necessary to directly kill the nerve cells during exposure. The malformation of fine structure/functional network will become symptomatic when the animals grow up to adults. On the other hand, most of those insults to adults would end up in reversible and transient changes.

Such delayed toxicity cannot be readily detected by currently available functional observational battery-(FOB-) based neuronal test system. Our new findings fall into the category of "early exposure- delayed effect". As mentioned above, nervous systems of developing organisms are susceptible to signal disruption which could lead to the delayed neurobehavioral anomaly. Toxicology is asked to prepare to respond to such new types of toxicity or "signal toxicity" with a consideration on the mechanisms which could explain the severity and irreversibility specific to children.

In conclusion, the 35th Annual Meeting of the Japanese Society of Toxicology had raised "Children's Toxicology" as one of its main Themes, and organized Special lectures, five Symposiums and two Workshops on Children's Toxicology of various targets and pending problems, which includes central nervous system, immune system, and endocrine system as targets, as well as problems in pharmacology i.e. issues on children's preclinical and clinical trials and on the off-label use of drugs. This special issue of the Journal of Toxicological Sciences gathers the peer-reviewed papers presented by the authors who participated in the lectures/symposiums/workshops on Children's Toxicology at the meeting.

REFERENCES

Benford, D. (2000): The Acceptable Daily Intake: A food for ensuring food safety. ILSI Europe Concise Monograph series. Brussels, Belgium.

Cohen-Cory, S. (2002): The developing synapse: construction and modulation of synaptic structures and circuits. Science, **298**, 770-776.

Kanno, J., Aisaki, K., Igarashi, K., Nakatsu, N., Ono, A., Kodama, Y. and Nagao, T. (2006): "Per cell" normalization method for mRNA measurement by quantitative PCR and microarrays. BMC Genomics, **7**, 64.

Newman, C.G.H. (1985): Teratogen update: clinical aspects of thalidomide embryopathy-a continuing preoccupation. Teratology, **32**, 133-44.

Tanemura, K., Igarashi, K., Matsugami, T.R., Aisaki, K., Kitajima, S. and Kanno, J. (2009): Brain structure impairment and behavioral disturbance induced in male mice offspring by a single intraperitoneal administration of domoic acid (DA) to their dams. J. Toxicol. Sci., **34**, Special Issue II, SP279-SP286.

The Journal of Toxicological Sciences (J. Toxicol. Sci.)
Vol.34, Special Issue II, SP279-SP286, 2009

Intrauterine environment-genome interaction and Children's development (2): Brain structure impairment and behavioral disturbance induced in male mice offspring by a single intraperitoneal administration of domoic acid (DA) to their dams

Kentaro Tanemura, Katsuhide Igarashi, Toshiko-R Matsugami, Ken-ichi Aisaki, Satoshi Kitajima and Jun Kanno

Division of Cellular & Molecular Toxicology, Biological Safety Research Center, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan

(Received February 17, 2009)

ABSTRACT — To demonstrate induction of delayed central nervous toxicity by disturbing neuronal activities in the developing brain, we administered a single intraperitoneal dose of domoic acid (DA; 1 mg/kg), a potent glutamate receptor agonist, to pregnant female mice at the gestational day of 11.5, 14.5 or 17.5. The dams had recovered from acute symptoms within 24 hr, followed by normal delivery, feeding and weaning. All male offspring mice after weaning were apparently normal in response to handlers during cage maintenance, body weight measurement and to mate mice in group housing conditions. At the age of 11 weeks, our neurobehavior testing battery revealed severe impairment of learning and memory with serious deviances of anxiety-related behaviors. The developed brain of prenatally exposed mice showed myelination failure and the overgrowth of neuronal processes of the limbic cortex neurons. This study indicates that the temporal disturbance of neurotransmission of the developing brain induces irreversible structural and functional damage to offspring which becomes monitorable in their adulthood by a proper battery of neurobehavioral tests.

Key words: Domoic acid, Prenatal exposure, Brain structure, Behavior

INTRODUCTION

Adequate neural activities are necessary for the maturation of neural networks during brain development (Rice and Barone, 2000). Historically, the presence of such plasticity-driven mechanisms has been demonstrated by a series of studies of eyelid suture in kittens or monkeys and corresponding findings reported in young human cataracta patients (Wiesel, 1982; Gu *et al.*, 1989; Fonta *et al.*, 2000). These processes require proper stimuli to the brain that trigger the release of neurotransmitters from the neurons and subsequent receptor-mediated signal transduction (Ooi and Wood, 2008; Greer and Greenberg, 2008; Cohen-Cory, 2002). Therefore, it is highly conceivable that disturbance of neural activities by neuroactive xenobiotics leads to malformation of the fine structure of the brain. Even when the exposure was transient, it would result in anomaly of higher brain functions in adulthood without overt signs of brain damage during maturation.

Glutamate receptors begin to express in the late embryonic stages, and their expression increases with the advance of brain development (Luján *et al.*, 2005; Manent *et al.*, 2005). Prenatal exposure of xenobiotic chemicals that interfere with the glutamate receptor function could induce malformation of the fine structure of the brain which should lead to anomaly of higher brain function that is different from acute neurotoxicity known for such chemicals to induce in adults (Bondy and Campbell, 2005). A marine biotoxin domoic acid (DA) which is structurally related to glutamate, and activates ionotropic α-amino-3-hydroxy-5-methyl-4-isoxazolepropionic acid (AMPA) and kinate subtypes of glutamate receptors (Pulido, 2008) is known to cause acute symptoms of diarrhea, seizures and memory loss in adult human by eating contaminated shellfish (Tryphonas and Iverson, 1990), and DA induced acute neurotoxicity in animal

Correspondence: Jun Kanno (E-mail: kanno@nihs.go.jp)

model (Chandrasekaran *et al.*, 2004). Additionally, DA is also known to cross the placenta, and enters prenatal brain tissue in rats (Maucher and Ramsdell, 2007). Therefore, prenatal exposure of DA may disrupt the neural activities by excessive stimulation of glutamate receptors, and should induce fine structural and functional disorganization in the developing brain. Here, we report that a transient transplacental DA exposure *in utero* induced alteration of the neurobehavioral parameters and corresponding fine brain structure of the male C57BL/6 mice in their adulthood.

MATERIALS AND METHODS

Animal treatment

All experiments were carried out under approval of Experimental Animal Use Committee of National Institute of Health Sciences, Japan. Pregnant C57LB/6 female mice obtained from Japan SLC, Inc., were individually housed in plastic breeding cages with free access to water and pellet diet (CRF-1; Oriental Yeast Co., Tokyo, Japan) in a 12 hr light-dark cycle conventional condition. Four groups with five pregnant mice each were prepared. All groups received three intraperitoneal injections on gestational day 11.5 (E11.5) as a late embryonic period, 14.5 (E14.5) and 17.5 (E17.5) as early and late fetal period respectively. Group A (Control) received three i.p. shots of saline on E11.5, E14.5 and E17.5. Group B (DA@ E11.5) received one shot of DA (Calbiochem, San Diego, CA, USA) at a dosage of 1 mg/kg on E 11.5 and two shots of saline on E14.5 and E17.5. Group C (DA@14.5) received a shot of saline on E11.5, a shot of DA on E14.5 and another saline on E17.5. Group D (DA@E17.5) received two shots of saline on E11.5 and E14.5 , and a shot of DA on E17.5. The pups were weaned at 4 weeks of age, and four male mice per liter were randomly selected and housed in one cage with free access to water and CRF-1 pellet until 11 weeks of age.

Immunohistochemical analysis

Brains (n = 4 male mice per group) were fixed with methacarn fixative (methanol: chloroform:acetic acid, 60:30:10 v/v) and paraffin-embedded sections were prepared. Mouse monoclonal anti-microtubule associated protein 2 (MAP2, sc-32791; Santa Cruz, CA, USA), mouse monoclonal anti-neurofilament-m (NF-M, sc-20013; Santa Cruz, CA, USA), rabbit polyclonal anti-myelin associated glycoprotein (MAG, sc-15324; Santa Cruz, CA, USA), and rabbit polyclonal anti MAP2 (sc-20172; Santa Cruz, CA, USA) were used. Deparaffinized sections were pretreated with HistoVT-One (Nacalai Tesque, Kyoto, Japan.) as previously described (Tanemura *et al.*, 2005) and incubated with primary antibodies. Secondary antibodies were Alexa 568-conjugated anti-mouse IgG and Alexa 488-conjugated anti-rabbit IgG (Molecular Probes, Eugene, OR, USA). Fluorescent images were obtained with an FV-300 confocal laser scanning microscope (Olympus, Tokyo, Japan). For semi-quantitative analysis of images, we calculated the ratio of fluorescence intensity compared to control mice (group A), by using the IMAGE J program (http://rsb.info.nih.gov/ij/index.html. National Institute of Health, Bethesda), after adjusting background noise (n = 4 images per mouse).

Neurobehavioral tests

A battery of neurobehavioral tests were conducted on open field test (OF), light/dark transition test (LD), elevated plus maze test (EP) and contextual/cued fear conditioning test (FZ). Experimental apparatuses and image analyzing softwares were obtained from O'Hara & Co., Ltd., Japan. Image analyzing softwares (Image OF4, Image LD2, Image EP2 and Image FZ2) were developed from the public domain IMAGE J program. All experiments were done with 8 mice per group (32 mice total), and were conducted between 13:30 and 16:30. The level of background noise during behavioural testing was about 50 dBA. After each trial, the apparatuses were wiped and cleaned.

Open field test

The locomotor activity was measured for 10 min using an open field apparatus made of white plastic (50 x 50 x 40 (H) cm).

An LED light system was positioned 50 cm above the centre of the field (50 lux at the centre of field). Total distance travelled (cm), time spent in the central area (30% of the field) (sec), and the frequencies of movement were measured (Tanemura *et al.*, 2002).

Light/dark transition test

The apparatus used for the light/dark transition test consisted of a cage (21 x 42 x 25(H) cm) divided into two chambers by a partition with an opening. One chamber is brightly illuminated (250 lux), whereas the other chamber is dark (2 lux). A mouse is placed into the dark area and allowed to move freely between the two chambers through the opening for 5 min. The latency for the first move to the light area, the total number of transitions and the time spent on each side were measured.

Elevated plus maze test

The plus-shaped apparatus consisted of four arms (25

x 5 cm) connected to a central square area (5 x 5 cm). Opposite two arms are enclosed with 20 cm-high transparent walls and other two are left open. The floor of the maze is made of white plastic plate and is elevated 60 cm above the room floor (200 lux at the centre of the apparatus). A mouse is placed to the central square area of the maze, facing one of the open arms, and the behavior was recorded for 10 min: total distance traveled (cm), total time on open arms and central square area (sec) and the total number of entry to any of the arms (Tanemura *et al.*, 2002).

Contextual/cued fear conditioning test

The apparatus consists of a conditioning chamber (or a test chamber) (17 x 10 x 10 (H) cm) made of clear plastic with ceiling and placed in a sound proof box. The chamber floor has stainless steel rods (2-mm diameter) spaced 5 mm apart for giving electric foot shock (0.1 mA, 3 sec duration) to the mouse. The soundproof box consists of white-coloured wood, and is equipped with an audio speaker and light source (35 lux at the centre of the floor). A CCD camera is positioned 20 cm above the ceiling of the chamber. During the conditioning trial (Day 1), mice are placed individually into the conditioning chamber in the sound proof box and, after 90 sec, they are given three tone-shock pairings (30 sec of tone, 75 dB, 10 KHz followed by 3 sec of electric shock at the end of tone, 0.1 mA) separated by 90 sec. Then they are returned to their home cage. Next day (Day 2), as a "contextual fear test", they are returned to the conditioning chamber without tone and shock for a 6-min. On the third day (Day 3), they are brought to a novel chamber of different make without stainless steel rods place in the sound proof box and, after a period of 3 min, only the conditioning tone is presented for 3 min (no shock was presented, 35 lux at the centre of the floor). The freezing response of mice was defined as a consecutive 2 sec period of immobility. Freezing rate (%) was calculated as [time freezing/session time] x 100 (Tatebayashi *et al.*, 2002).

Statistical analysis

Data were indicated as means ± S.D. Statistical analysis was conducted with student`s t-test by using StatView (SAS Institute, Cary, NC, USA). A p-value of < 0.05 compared to the results of control male mice (group A) was considered statistically significant.

RESULTS

Effects on morphology of brain by prenatal exposure to DA

Offspring mice of all groups after weaning up to the age of 11 weeks were apparently normal in response to handlers during cage maintenance, body weight measurement and to mate mice in group housing conditions. Routine histological observation of the brain at 11 weeks old by hematoxylin-eosin staining could not reveal difference among the groups (data not shown). By immunohistochemical study on the same brain sections, reduced immuno-reactivity against the MAG, the marker for myelin, was detected in the cortices of group B (DA@11.5) and C (DA@14.5) compared to control (Figs. 1A-D and I). In contrast, increased immuno-reactivity against MAP2, the marker for neuronal dendrite, was indicated in the same area of group B (DA@11.5), C (DA@14.5) and D (DA@17.5) compared to control (Figs. 1E-H and J). Increased immuno-reactivity against MAP2 was also found in lateral area of CA3 hippocampus of group B (DA@11.5), C (DA@14.5) and D (DA@17.5) compared to control, whereas immuno-reactivity for MAP2 showed no significant difference in medial area of CA3 hippocampus among the groups (Figs. 2A-D and I). Immuno-reactivity against NF-M; the marker for neuronal axon, also showed no significant difference in the same area among the groups (Figs. 2E-H and J).

Effects on behavior by prenatal exposure of DA

In the OF test, the distance traveled was not different among the groups (Fig. 3A), the time spent in center area was significantly prolonged in group D (DA@17.5) mice (Fig. 3B). In the LD test, group C (DA@14.5) mice stayed in light area for longer time (Fig. 4A), and latency for the first move to light area was significantly shorter in group C (DA@14.5) and D (DA@17.5) (Fig. 4B). In the EP test, significantly increased distance traveled and time spent in the open area were detected for group B (DA@11.5), C (DA@14.5) and D (DA@17.5) (Figs. 5A and B). In the FZ test, both Day 1 and Day 2 freezing responses of group C (DA@14.5) and D (DA@17.5) were significantly reduced compared to control (Figs. 6A and B).

DISCUSSION

The expression levels of glutamate receptors starts to elevate at the fetal period, i.e. approximately from E14 (Luján *et al.*, 2005; Manent *et al.*, 2005). Exogenous glutamatergic stimuli at this period would affect the for-

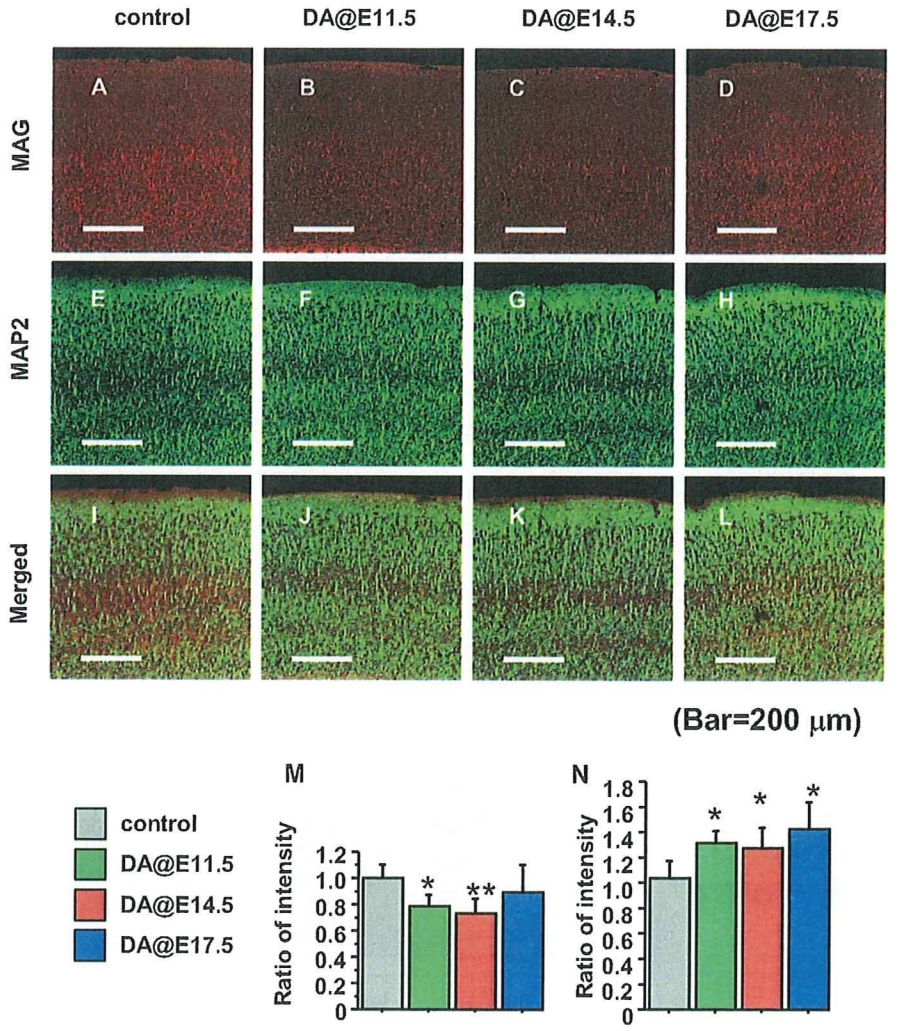

Fig. 1. Delayed effects on cerebral cortex induced by prenatal exposure of DA.
A-D, Immunohistochemical staining against MAG; E-H, immunohistochamical staining against MAP2; I-L, merged images of the cerebral cortex. A, E, I, group A (control), B, F, J, group B (DA@11.5), C, G, K, group C (DA@14.5) and D, H, L, group D (DA@17.5). Scale bar = 200 mm. M, Quantitative analysis in intensity ratio to control of MAG expression, and J, MAP2 expression among the groups (mean ± S.E.M.). Asterisk (**) and (*) indicate significant difference compared to control ($P < 0.01$) and ($P < 0.05$).

mation of the neural circuits. An extreme example to support this hypothesis would be the phenotype of the double knockout mouse of glutamate transportors GLT1 and GLAST (Matsugami *et al.*, 2006). Lack of these transportors is considered to result in abnormally high concentration of glutamate in the brain. In fact, morphological anomaly became apparent in synchronization with the expression of glutamine receptors. In our study, corresponding to the hypothesis, the neurobehavioral symptoms as a whole was severer for those exposed at fetal periods, i.e. E14.5 and E17.5, compared to those at embryonic period, i.e. E11.5 (Fig. 7).

We demonstrated that a prenatal exposure of a relatively low dose of DA induced a spectrum of neurobehavioral anomalies which became monitorable at the adult stage accompanied by alteration in fine brain structures detectable by immunohistochemistry. It is emphasized that this amount of DA did not induce abnormal responses dur-

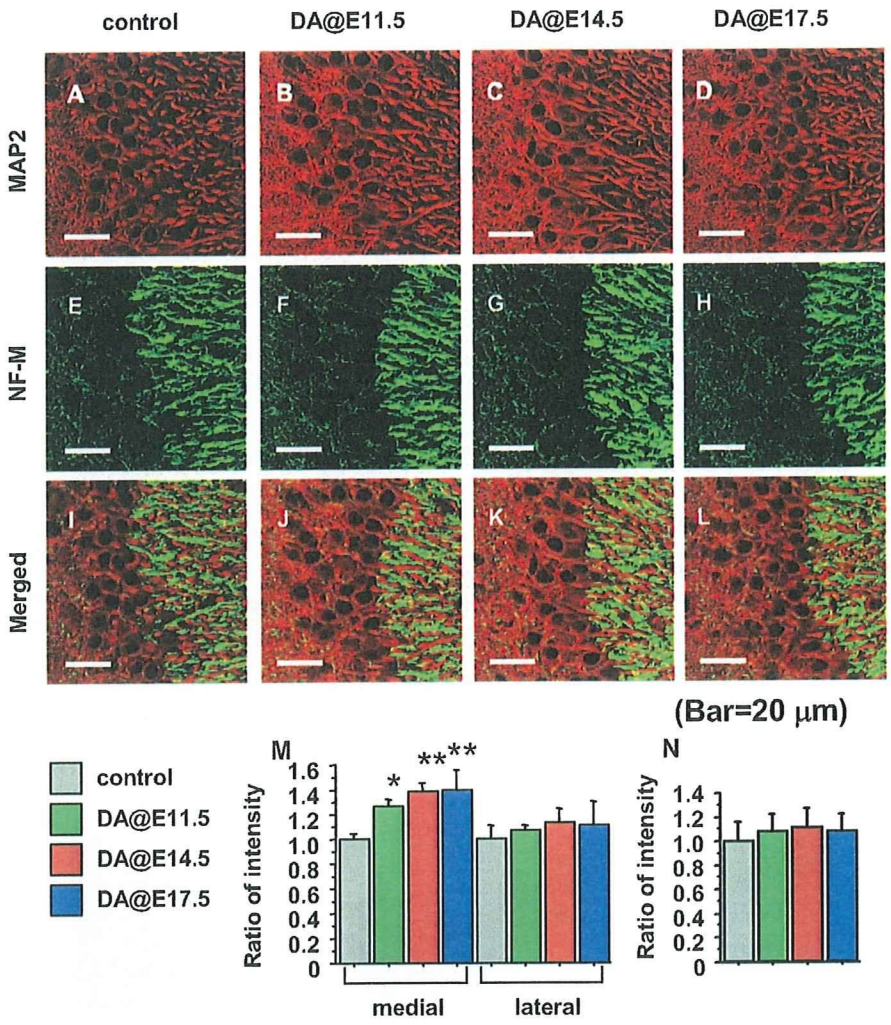

Fig. 2. Delayed effects on hippocampus induced by prenatal exposure of DA.
A-D, Immunohistochemical staining against MAP2; E-H, immunohistochamical staining against NF-M; I-L, merged images, of CA3 hippocampus. A, E, I, group A (control), B, F, J, group B (DA@11.5), C, G, K, group C (DA@14.5) and D, H, L, group D (DA@17.5). Scale bar = 200 mm. M, Quantitative analysis of MAP2 expression, and N, NF-M expression among the groups (mean ± S.E.M.). Asterisk (**) and (*) indicated significant difference compared to control ($P < 0.01$) and ($P < 0.05$).

ing maturation, such as hyperreactivity to handling and to cage mates, and did not present overt malformation of the brain detectable by the routine H&E histology at the age of 2 weeks (data not shown). It is also noted that the spectrum of the neurobehavioral symptoms induced in mice exposed to DA at adulthood was different from those monitored in this study (data not shown).

Although progressive hippocampal neuronal damages were reported to be induced by prenatal administration of DA (0.6 mg/kg intravenous injection to the dam) (Dakshinamurti *et al.*, 1993), we did not find notable neuronal loss or neuronal cell death as the delayed effects in adult mouse brain by prenatal exposure. On the other hand, we found myelination failure (Miller and Mi, 2007) in cortex of group B (DA@11.5) and C (DA@14.5) mice. And we also detected a finding compatible with the overgrowth of neuronal processes in cortex and hippocampus of group B (DA@11.5), C (DA@14.5) and D (DA@14.5)

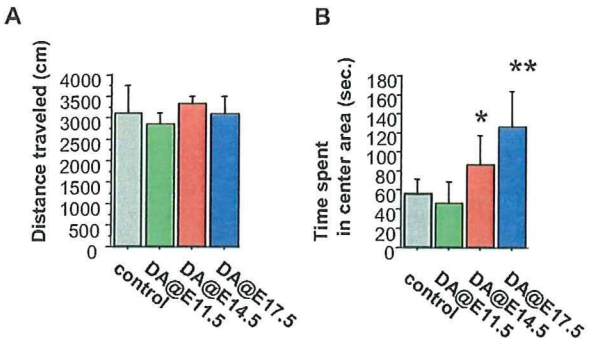

Fig. 3. Delayed effects on locomotor activity (OF test) induced by prenatal exposure of DA.
A, Mean distance travelled (total distances divided by total duration of trial, 10 min) and B, mean time spent in center area (30% of the field) in the open field apparatus (mean ± S.E.M.). Asterisk (**) and (*) indicated significant difference compared to control ($P < 0.01$) and ($P < 0.05$).

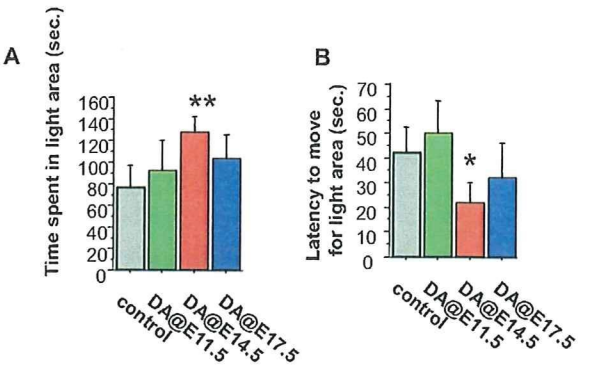

Fig. 4. Delayed effects on anxiety-related behavior (LD test) induced by prenatal exposure of DA.
A, Total time spent in light area, and B, latency time to move to light area in the LD apparatus (mean ± S.E.M.). Asterisk (**) and (*) indicated significant difference compared to control ($P < 0.01$) and ($P < 0.05$).

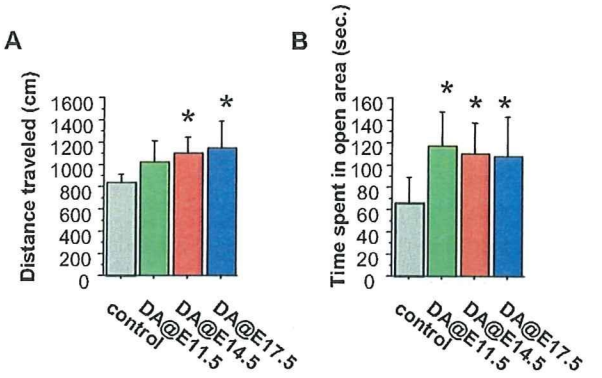

Fig. 5. Delayed effects on anxiety-related behavior (EP test) induced by prenatal exposure of DA.
A, Total distance travelled, and B, total time spent in open area in the elevated plus maze apparatus (mean ± S.E.M.). Asterisk (*) indicated significant difference compared to control ($P < 0.05$).

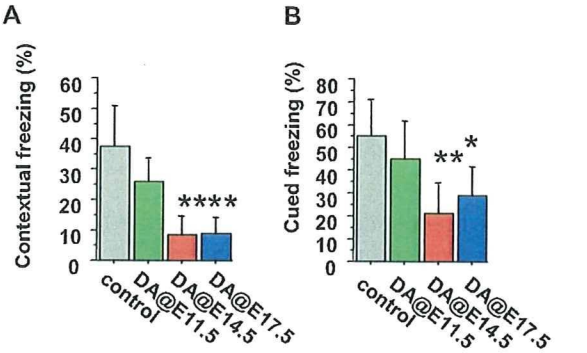

Fig. 6. Delayed effects on learning and memory (FZ test) induced by prenatal exposure of DA.
A, Contextual fear test and B, cued fear test. Memory performance is expressed as a mean percent duration of freezing responses (mean ± S.E.M.). Asterisk (**) and (*) indicated significant difference compared to control ($P < 0.01$) and ($P < 0.05$).

mice by using cytoskeletal marker. These findings indicated that the disorganization of brain was induced by the prenatal exposure of DA, and remained irreversibly up until the maturation period.

Among multiple endpoints of the behavioral test battery we used, serious deviances in anxiety-related behaviors of group C (DA@14.5) and D (DA@17.5) mice were observed. Mice in those groups showed low performances in adaptations for novel circumstances, i.e., strange and broad area in OF test, beamish place in LD test, high and narrow space in EP test. Additionally, we also found severe impairment of learning and memory. Although the low performances of memory task have been reported in rats with prenatal DA exposure (Levin *et al.*, 2005),

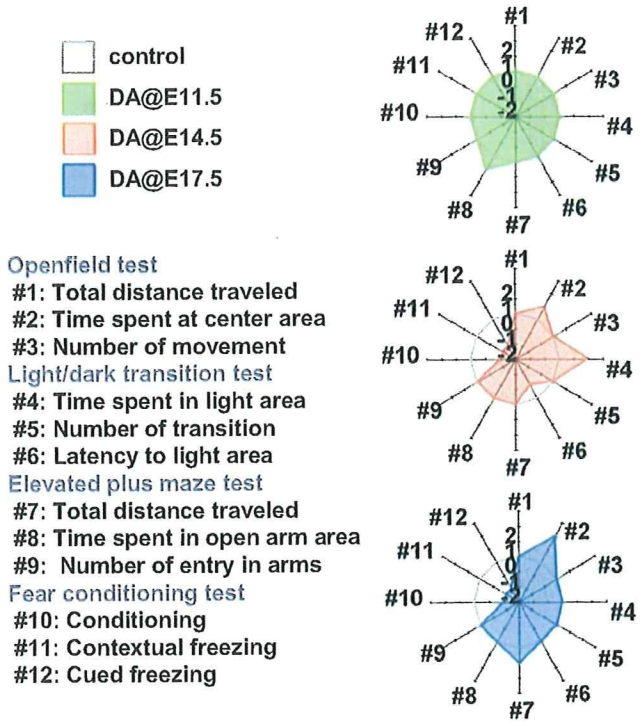

Fig. 7. Summary radar chart of the neurobehavioral battery test results. Radial axis indicates the direction (increase or decrease) of the deviation, and the p value of the endpoints compared to the control (+1 and −1, $0.01 <= p < 0.05$, +2 and −2, $p < 0.01$). Regular dodecagon of radius 0 indicates no deviation from control.

we showed serious deviances about affective (emotional) behaviors additional to severe memory deficit.

In conclusion, we clearly indicated that the disturbance against the adequate neural activity during developmental period when glutamate receptors became active have induced delayed memory defect and unnatural adoptive behaviors that became monitorable at the maturation period in mice. The responsible foci deduced from these behavioral disturbances are the limbic cortex and hippocampus. Our morphological findings are consistent with the interpretation. A combination of neurobehavioral and pathomorphological analysis was shown to be an effective method to assess delayed neurotoxic effects which dose not induce immediate organic brain damage and related symptoms after exposure. Having adopted the hypothesis that exogenous stimuli to neural signaling systems during the development of the brain can be a cause of delayed anomaly of higher brain function, stimuli toward systems other than glutamate receptors should also induce such anomaly of different targets and symptoms in concert with the distribution of the corresponding receptor(s) in the developing brain. Such data on other system would be reported elsewhere.

ACKNOWLEDGMENTS

The authors thank Mr. Yusuke Furukawa and Ms. Maki Otsuka for technical support. This study was supported by Health Sciences Research Grants H17- -Kagaku- 001 from the Ministry of Health, Labour and Welfare, Japan.

This peer-reviewed article is based upon a lecture presented at the 35th Annual Meeting of Japanese Society of Toxicology, June 2008 in Tokyo under the theme of "Children's Toxicology", June 2008 in Tokyo.

REFERENCES

Bondy, S.C. and Campbell, A. (2005): Developmental neurotoxicology. J. Neurosci. Res., **81**, 605-612.

Chandrasekaran, A., Ponnambalam, G. and Kaur, C. (2004): Domoic acid-induced neurotoxicity in the hippocampus of adult rats. Neurotox. Res., **6**, 105-117.

Cohen-Cory, S. (2002): The developing synapse: construction and

modulation of synaptic structures and circuits. Science, **298**, 770-776.

Dakshinamurti, K., Sharma, S.K., Sundaram, M. and Watanabe, T. (1993): Hippocampal changes in developing postnatal mice following intrauterine exposure to domoic acid. J. Neurosci., **13**, 4486-4495.

Fonta, C., Chappert, C. and Imbert, M. (2000): Effect of monocular deprivation on NMDAR1 immunostaining in ocular dominance columns of the marmoset Callithrix jacchus. Vis. Neurosci., **17**, 345-352.

Greer, P.L. and Greenberg, M.E. (2008): From synapse to nucleus: calcium-dependent gene transcription in the control of synapse development and function. Neuron, **59**, 846-860.

Gu, Q.A., Bear, M.F. and Singer, W. (1989): Blockade of NMDA-receptors prevents ocularity changes in kitten visual cortex after reversed monocular deprivation. Brain Res. Dev. Brain Res., **47**, 281-288.

Levin, E.D., Pizarro, K., Pang, W.G., Harrison, J. and Ramsdell, J.S. (2005): Persisting behavioral consequences of prenatal domoic acid exposure in rats. Neurotoxicol. Teratol., **27**, 719-725.

Luján, R., Shigemoto, R. and López-Bendito, G. (2005): Glutamate and GABA receptor signalling in the developing brain. Neuroscience, **130**, 567-580.

Manent, J.B., Demarque, M., Jorquera, I., Pellegrino, C., Ben-Ari, Y., Aniksztejn, L. and Represa, A. (2005): A noncanonical release of GABA and glutamate modulates neuronal migration. J Neurosci., **25**, 4755-4765.

Matsugami, T.R., Tanemura, K., Mieda, M., Nakatomi, R., Yamada, K., Kondo, T., Ogawa, M., Obata, K., Watanabe, M., Hashikawa, T. and Tanaka, K. (2006): Indispensability of the glutamate transporters GLAST and GLT1 to brain development. Proc. Natl. Acad. Sci. USA, **103**, 12161-12166.

Maucher, J.M. and Ramsdell, J.S.(2007): Maternal-fetal transfer of domoic acid in rats at two gestational time points. Environ Health Perspect., **115**, 1743-1746.

Miller, R.H. and Mi, S. (2007): Dissecting demyelination. Nat. Neurosci., **10**, 1351-1354.

Ooi, L. and Wood, I.C. (2008): Regulation of gene expression in the nervous system. Biochem. J., **414**, 327-341.

Pulido, O.M. (2008): Domoic acid toxicologic pathology: a review. Mar. Drugs, **6**, 180-219.

Rice, D. and Barone, S.Jr. (2000): Critical periods of vulnerability for the developing nervous system: evidence from humans and animal models. Environ. Health Perspect., **108**, 511-533.

Tanemura, K., Murayama, M., Akagi, T., Hashikawa, T., Tominaga, T., Ichikawa, M., Yamaguchi, H. and Takashima, A. (2002): Neurodegeneration with tau accumulation in a transgenic mouse expressing V337M human tau. J Neurosci., **22**, 133-141.

Tanemura, K., Ogura, A., Cheong, C., Gotoh, H., Matsumoto, K., Sato, E., Hayashi, Y., Lee, H.W. and Kondo, T. (2005): Dynamic rearrangement of telomeres during spermatogenesis in mice. Dev. Biol., **281**, 196-207.

Tatebayashi, Y., Miyasaka, T., Chui, D.H., Akagi, T., Mishima, K., Iwasaki, K., Fujiwara, M., Tanemura, K., Murayama, M., Ishiguro, K., Planel, E., Sato, S., Hashikawa, T. and Takashima, A. (2002): Tau filament formation and associative memory deficit in aged mice expressing mutant (R406W) human tau. Proc. Natl. Acad. Sci. USA, **99**, 13896-13901.

Tryphonas, L. and Iverson, F. (1990): Neuropathology of excitatory neurotoxins: the domoic acid model. Toxicol Pathol., **18**, 165-169.

Wiesel, T.N. (1982): Postnatal development of the visual cortex and the influence of environment. Nature, **299**, 583-591.

ENVIRONMETRICS
Environmetrics 2009; **20**: 1–13
Published online 8 May 2008 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/env.898

An experimental design for judging synergism on consideration to endocrine disruptor animal experiments

Nobuhito Matsunaga[1*,†], Jun Kanno[2], Chikuma Hamada[3] and Isao Yoshimura[3]

[1]*Kyowa Pharmaceutical, Inc., 212 Carnegie Center, Suite 101, Princeton, NJ 08540, USA*
[2]*National Institute of Health Sciences, Tokyo, Japan*
[3]*Tokyo University of Science, Tokyo, Japan*

SUMMARY

This paper investigates an appropriate statistical design for an animal experiment to evaluate synergism of two test chemicals. It assumes a certain number of animals are divided into groups, each of which is treated with a combination of dose levels of two chemicals. A design is identified by the set of group size for each combination of doses, including the case where the dose of either one chemical is zero. The power of t-test to detect synergism by positive surplus of response on a simultaneous administration group from the additivity plane composed of the responses on single administration groups is adopted as the criterion for the appropriate design. The applicable design is investigated for the application to real cases of endocrine disrupter study conducted at the National Institute of Health Sciences of Japan.

It revealed that the dose level of the simultaneous administration group should be located inside or on the boundary of a triangular region and that the total number of animals should be the same as those for single administration groups. Copyright © 2008 John Wiley & Sons, Ltd.

KEY WORDS: additivity; animal experiment; experimental design; endocrine disruptor; synergism; triangular region

1. INTRODUCTION

In the past, environmental pollutants were regulated according to individual effects. However, recently, there has arisen the problems of combinations of complex pollutants, and regulations that address synergism have become necessary. As a result, experimental researches have been conducted on pollutant synergism. The investigation by Kanno, one of the authors (Kanno *et al.*, 2001) on the synergism of endocrine disruptors, using the rodent uterotrophic assay, is an example of such researches.

In our experiments using multiple test substances, dividing animals such as rats into multiple groups of single administration and simultaneous administration, we estimated the response when there is no synergism based on the response in the single administration group to investigate whether the response in the simultaneous administration group exceeds the estimated response.

*Correspondence to: N. Matsunaga, Kyowa Pharmaceutical, Inc., 212 Carnegie Center, Suite 101, Princeton, NJ 08540, USA.
†E-mail: matsunaga.nobuhito@kyowa-kpi.com

Received 21 February 2006
Accepted 20 October 2007

For the data analysis method used in this type of animal experiment to investigate synergism, Kelly and Rice (1990) proposed a method to evaluate the dose–response curve by smoothing method. Gennings and Carter (1995) and Gennings *et al.* (1997), on the other hand, proposed a method to evaluate synergism by using a model in which the response becomes flattened when there is no synergism. Using a similar plan to that of Gennings *et al.* (1997), Matsunaga *et al.* (2003) proposed a method to evaluate the difference in the responses with simultaneous administration of two substances from those estimated by applying an additive model to the data of single administration of each substance. They applied their proposed method to the actual data analysis. Other data analysis methods are also cited in Laska and Meisner (1989) and Machado and Robinson (1994).

For the experimental design evaluating the synergism, Hasegawa *et al.* (1996) proposed the experimental design of animal experiment for five or ten chemical mixture, and Straetemans *et al.* (2005) investigated a fixed-ratio design on *in vitro* study. Abdelbasit and Plackett (1982) and Tan *et al.* (2003) are also related to this issue. However, the situation of these researches differs from our case study. Our interest is limited to simply checking whether the effects of the chemicals are additive or not. And in the animal experiments, some assumptions and limitations are generated for the applicable information and experimental conditions.

In this kind of research, we thought that the problem of the experimental design is to determine appropriate dose levels for simultaneous administration and to select the appropriate number of animals for allocation to the dose levels. However, in the past research, the research directly related to this problem by the animal experiment was not found. Accordingly, based on the analysis methods proposed by Matsunaga *et al.* (2003), we investigated what type of design would be appropriate.

The paper is organized as follows. Section 2 introduces the conditions in the case study that motivated this paper, while Section 3 formulates the issues. Section 4 derived the appropriate design corresponding to the case studies dealt with this paper. Finally, Section 5 provides Conclusions and Discussions for future issues.

2. MOTIVATING CASE

According to the World Health Organization, an endocrine disruptor is defined as "an exogenous substance or mixture that alters function(s) of the endocrine system and causes adverse health effects in an intact organism, or its progeny or (sub)populations."

The effect of endocrine disruptors is not stimulated directly at the site of the adverse effect, but is mediated by the signal and occurs through nuclear receptor. Furthermore, there is more than one signal transduction system in humans and animals. Because nuclear receptors and transcription factors are redundant, there may be an interaction between different pathways, which leads to possible synergistic endocrine disruption action. Here the definition of "synergistic" is that if two chemicals produce the same endpoint, they bring about a larger response as compared to the anticipated response when these chemicals are purely added. Consequently, it is very important to realize the synergism between two endocrine disruptors such as "Genistein" and "bisphenol A (BPA)" through animal experiments for explaining the mechanism of action.

Both Genistein and BPA bind to estrogen receptors and elicit estrogenic responses to an organism including uterotrophic responses. Genistein is a phytoestrogen found relatively abundant in soybeans and its derivative foods. BPA is the basic monomer of polycarbonate plastic and epoxy resins widely used as a lining for food and beverage cans, in hard plastic baby and water bottles, toys, dental sealants, etc. It has been reported that BPA monomer can leached out to food and drinks especially when the

DOI: 10.1002/env

polymerization process is incomplete and/or the plastic is aged. These two estrogenic compounds can be found very commonly in our food environment. Therefore, it is of great importance from the point of the safety regulations to examine whether the combined effect is additive or synergistic.

For the experimental design of the research based on this background, some assumptions and limitations were generated for the applicable advanced information and experimental conditions. In the interests of simplicity, when explaining synergism research for two substances, the limitations are as follows.

First, before conducting the experiment that investigates synergism, advance knowledge can be obtained to some extent on dose–response curves for the single administration from preliminary experiments using each substance. Therefore, if the dose–response curve is nonlinear, by appropriate transformation of variables for the dose and response, it is possible to assume the dose–response curve to be approximately linear.

Second, the maximum dose of each substance used in the experiment is limited. We, for example, recognize in our experiments that the signal transduction system amplifies the signal at significantly smaller doses as compared to the dose used in normal toxicity studies. Other toxicities appear with higher doses, so that the endocrine disruption effect to be investigated is concealed. Because these maximum doses, $D_{A\max}$ and $D_{B\max}$, generally can be obtained through preliminary experiments, the range of dose levels used in the experiments can be limited.

Third, because various kinds of test substances are to be investigated, the number of animals, n, used in each experiment is relatively small. In our actual experience, the number of animals is approximately 40–50.

Fourth, the fundamental form of the experiment for investigating synergism is roughly decided. In an experiment using two substances A and B, we set single administration groups of G_{00} (at a dose level of 0), G_{A2}, and G_{B2} (at the maximum dose levels of $d_{A2} = D_{A\max}$ and $d_{B2} = D_{B\max}$, respectively), and groups G_{A1} and G_{B1} (at the middle dose levels of d_{A1} and d_{B1}, respectively), and administer the test substances after assigning the same number of animals n_s in every group. Independent of this, we set one or more simultaneous administration groups of G_{AB} with dose levels of the two substances at d_A and d_B. We measure responses by performing the experiment with this type of design, and estimate dose–response curves by forecasting synergism from the single administration group to confirm whether the response obtained in the simultaneous administration group is larger compared to that estimated.

Fifth, the observed response is usually quantitative variable such as the uterine weight of rats, which generally shows normal distribution, because it is difficult to define the additivity/synergism for the response in qualitative values.

Under the above conditions, what should be questioned for the design is what the most appropriate dose for simultaneous administration is, and whether to have more or less animals for the simultaneous administration group as compared to the single administration group. In order to obtain guidelines for these, this paper generalizes and formulates the above mentioned problems to make numerical evaluations under some conditions.

3. FORMULATION OF PROBLEM

3.1. Definition of synergism

In order to simplify the discussion, hereafter, we assume that there are two test substances denoted by A and B.

There have been many discussions in the past for how to define the terms additivity, synergism, and antagonism. Synergism is not defined unconditionally (Hewlett and Plackett, 1959; Berenberm, 1989).

DOI: 10.1002/env

Table 1. Difference between factorial design and triangular design

	Dose of A		
Dose of B	d_{A0}	d_{A1}	d_{A2}
(a) Factorial design			
d_{B0}	(1)	(2)	(3)
d_{B1}	(4)	(6)	(7)
d_{B2}	(5)	(8)	(9)
(b) Triangular design			
d_{B0}	(1)	(2)	(3)
d_{B1}	(4)	(6)	
d_{B2}	(5)		

Single chemical is administered at (1)–(5), whereas combination of two chemicals is administered at (6)–(9).

In fact, in the general remarks of these studies, many ideas are introduced for discussing synergism such as "independent joint action," "potentiation," "simple similar action," "complex action," and "dissimilar action." We will first explain the definition of synergism that is adopted in this paper.

From the standard statistical viewpoint, the dosages set certain dose level for the respective two substances as shown in Table 1(a). If the response at the simultaneous dose level is the sum of the effects generated by single substances, then the effect is considered to be additive. On the other hand, if it is large, there is a positive interaction and the effect is synergistic.

However, with toxic responses like endocrine disruption action, this point of view is not appropriate. Because, for this toxic response, as pointed out by, for example, Hasegawa *et al.* (1996), it is impossible to establish response linearity at doses that exceed the maximum dose for the respective substances, and thus it is impossible to determine whether a positive interaction is attributable to synergism or nonlinearity. Therefore, the following definition that expresses the tenets of Hewlett and Plackett (1959) by isobologram is adopted in this paper.

Label the expected response at dose d_A and d_B (doses for A and B) as $f(d_A, d_B)$. Also, label the single administration dose of A that results in an arbitrary response E as D_A, so that $f(D_A, 0) = E$. Similarly, label the dose of substance B that has expectation E as D_B. In the cases that motivated this study, A and B generate their responses in a similar stimulation process, the expected effects are proportional to the doses of A and B, and the effects of the two substances are additive. If these conditions hold, then $f(d_A, d_B) = E$ whenever (d_A, d_B) satisfies Equation (1).

$$\frac{d_A}{D_A} + \frac{d_B}{D_B} = 1 \tag{1}$$

The reason is that because a combination of dose levels like this represents simultaneous administration of A and B at an arbitrary ratio, using the amount that brings about a response of the same magnitude. If a response of magnitude E is consequently generated as expected, there is no special combined effect between the two chemicals. In this paper, when this relationship holds, the effect of the two substances is additive, or the two substances satisfy additivity.

On the other hand, if the two substances generate a synergistic response in different stimulation processes, it is considered that $f(d_A, d_B) > E$ is established with respect to an arbitrary (d_A, d_B) that satisfies Equation (1). In this paper, when this relationship is established, the two substances are synergistic, or satisfy synergism.

DOI: 10.1002/env

3.2. *Terminology, notation, and assumption*

The two-dimensional plane by plotting d_A (dose of A) on the x-axis and d_B (dose of B) on the y-axis is referred to as the dose plane, and the three-dimensional space by plotting the response on the z-axis above the dose plane is referred to as the response space.

In the experiment, the response is measured for each individual animal. The response that is measured is called the response variable, and is generally expressed by the symbol Z. The response variable Z measured for each individual is set as a random variable that follows a normal distribution independent of other individuals. Since it is assumed that the endpoint is organ weight as a target for application in the case study, such as endocrine disruptor study, it is considered that the assumption of the normal distribution is empirically valid. When the dose of the two substances administered is (d_A, d_B), the expected value is $E\{Z\} = f(d_A, d_B)$.

For single administration, that is, when the dose of one test substance is 0, Equation (2) can be assumed concerning the dose–response curve f.

$$f(d_A, 0) = \beta_0 + d_A\beta_A,\ f(0, d_B) = \beta_0 + d_B\beta_B \tag{2}$$

This assumes the dose–response curve for single administration to be linear. With this assumption, $f(d_A, d_B)$ is expressed by Equation (3). The two substances are additive if the hypothesis H_0 of Equation (4) holds, while synergistic if the hypothesis H_1 holds.

$$f(d_A, d_B) = \beta_0 + d_A\beta_A + d_B\beta_B + \Delta(d_A, d_B) \tag{3}$$

$$H_0 : \Delta(d_A, d_B) = 0,\ H_1 : \Delta(d_A, d_B) > 0 \text{ for all } (d_A, d_B) \tag{4}$$

The value of the dose used in the experiment is called the dose level, the collection of animals allocated for the dose level is called a group, and the number of animals for each group is called the group size, and the dose level of the group on the dose plane is called the group point. With this terminology, it is defined that "design is the set of group point and group size."

For numerical evaluation described in the next section, five groups of group size n_s for single administration and one group of group size n_m for simultaneous administration as described in Table 1(b) are assumed as the design. For the single administration group, the group points are set to be $(0, 0)$, $(d_{A1}, 0)$, $(d_{A2}, 0)$, $(0, d_{B1})$, $(0, d_{B2})$, and the response variables are distributed as normal with variance σ_s^2. For the simultaneous group, the group point is set to be (d_A, d_B), and the response variable is distributed as normal with variance σ_m^2

Let the sample mean of response variable in each group be Z_{00}, Z_{A1}, Z_{A2}, Z_{B1}, Z_{B2}, and Z_{AB}, respectively. It is assumed that $Z_{00}, \ldots, Z_{B2}$ are distributed as normal with the mean of Equation (2) and the variance σ_s^2/n_s, while Z_{AB} is distributed as normal with the mean of Equation (3) and the variance σ_m^2/n_m.

3.3. *Criterion for the appropriate design*

As criterion for the most appropriate design, it is natural to use the power in the hypothesis test of "H_0 versus H_1." Because the model is a linear model and the hypothesis is a linear hypothesis, a one-sided

DOI: 10.1002/env

t-test (or Welch test) with Equation (5) is naturally set as the test statistic.

$$T = \frac{\hat{\Delta}}{\sqrt{\hat{V}(\hat{\Delta})}} \tag{5}$$

Here, $\hat{\Delta}$ is the least square estimator of Δ, and the denominator of the statistics is the square root of the variance estimator.

The critical value of this test statistics is $t(\nu, \alpha)$, the upper 100α percentile of the t-distribution with degree of freedom ν, and the test is a one-sided test. In other words, A and B are judged to be synergistic when $T > t(\nu, \alpha)$.

This test, in short, detects the synergism when there is a statistically significant difference between Z_{AB} and the estimate obtained from the single administration groups assuming the dose–response surface under H_0.

4. EXAMPLES OF RECOMMENDED DESIGN

4.1. Real examples motivated the problem

We conducted many experiments and selected endocrine disruptor study as case study. This study was performed using triangular design such as Table 1(b). This design consisted of seven dose groups which included five group points for single administration and two group points for simultaneous administration. The endpoint was uterine weight gain and the main purpose was to evaluate whether the combined effect was synergistic or not. In order to explain the characteristic of the data, we took up two real examples. The details of these data are as follows.

Example 1. *Chemical A: genistein (mg/kg), chemical B: BPA (mg/kg)*

1. Group points for the single administration: $(d_A, d_B) = (0, 0), (12.5, 0), (25, 0), (0, 35), (0, 70)$.
2. Group points for simultaneous administration: $(d_A, d_B) = (6.25, 17.5), (12.5, 35)$.
3. Group size: $n_s = 6$, $n_m = 6$, $n = 42$.
4. Mean $\pm$ standard deviation of observed values

$$(d_A, d_B) = (0, 0) : 84.0 \pm 7.1$$
$$(d_A, d_B) = (12.5, 0) : 111.2 \pm 9.3, \quad (d_A, d_B) = (25, 0) : 149.5 \pm 33.7$$
$$(d_A, d_B) = (0, 35) : 138.4 \pm 16.2, \quad (d_A, d_B) = (0, 70) : 181.0 \pm 24.4$$
$$(d_A, d_B) = (6.25, 17.5) : 141.1 \pm 11.4, \quad (d_A, d_B) = (12.5, 35) : 180.2 \pm 26.8$$

5. Estimated value of Δ: 15.6.
6. Result of t-test: Significant in one-sided Welch test with significance level 2.5% $T = 2.19$, $\nu = 19$, $p = 0.02$.

Example 2. *Chemical A: diethylstilbestrol (μg/kg), chemical B: genistein (mg/kg)*

1. Group points for the single administration: $(d_A, d_B) = (0, 0), (0.1, 0), (0.2, 0), (0, 12.5), (0, 25)$.
2. Group points for simultaneous administration: $(d_A, d_B) = (0.05, 6.25), (0.1, 12.5)$.

DOI: 10.1002/env