

PCB-Induced Impairments in Older Adults: Critique of Schantz et al.'s Methodology and Conclusions

Shantz (1) provided a valuable scientific service to the field of the developmental neurotoxicity of polychlorinated biphenyls (PCBs) by offering insightful criticisms of the methodologies used by Jacobson et al. (2) and others. Schantz and colleagues proceeded to study PCBs and dichlorodiphenyl dichloroethane (DDE), shifting the focus from effects on infants and children to effects on a cohort of older adults (3–6). The shortcomings of the research design and data analysis used by Schantz et al. are equivalent to the shortcomings of previous studies (1). In their paper published in 2001 (3), Schantz et al. *a*) failed to account adequately for the chance significant findings that occur when many statistical analyses are conducted simultaneously; *b*) used an outdated measure of memory when a much-improved test was available at the time of testing; *c*) failed to consider the implications of the experimental interdependency of two key variables that were significantly related to PCBs; and *d*) controlled IQ (intelligence quotient) only with the Wechsler Adult Intelligence Scale-Revised (WAIS-R) vocabulary subtest.

Schantz et al. (3) conducted 48 multiple regression analyses simultaneously, 24 with DDE and 24 with PCBs, spanning several cognitive domains. They used an alpha level of 0.05, which means that one significant finding is expected to occur by chance alone for every 20 analyses; with 48 analyses, 2–3 significant findings will occur by chance. Schantz et al. identified four significant findings, which is barely above the number expected by chance; however, they focused only on three—the ones that produced the anticipated negative correlation—and they virtually ignored the significant, but opposite, relationship between DDE and delayed recall. Of the three negative associations with PCBs, two were experimentally interdependent—List A, Trial 1, and the semantic cluster ratio—both yielded by the California Verbal Learning Test (CVLT). The semantic cluster ratio is based on performance on Trials 1–5 of List A; consequently, the significant relationship between PCBs and Trial 1 of List A contributed to the relationship between PCBs and the semantic cluster ratio; hence, the two CVLT significant results may be redundant.

Furthermore, the other two significant results—the negative relationship with PCBs and the positive relationship with

DDE—occurred on the delayed recall portion of the logical memory subtest of the Wechsler Memory Scale (WMS). These results occurred on a 1975 revision of a long-outdated 1945 test (7). Why did Schantz et al. use such an old measure when the revision (WMS-R) was readily available when they collected their data? More to the point, the WMS did not even include a delayed recall component; Russell added that component in 1975, using a weak sample and producing a delayed recall measure with poor psychometric properties (8). In contrast, the 1987 WMS-R, which produces reliable and valid measures of immediate and delayed memory, has been given exceptional reviews (9). It is conceivable that the two significant results that occurred on the WMS Logical Memory test are more a function of the weak measurement of delayed recall than of any real relationship to PCBs or DDE.

In view of the multiple simultaneous comparisons and other points raised here, the best explanation for the significant results by Schantz et al. (3) is chance. The investigators discounted the impact of the many analyses because the negative results were confined to PCBs, as opposed to DDE, mercury, and lead, and because of alleged consistency with previous findings with children. In the “Discussion” (but not in the abstract), they urged caution in interpreting their results because of multiple analyses. However, the researchers did not consider the overlap in the two CVLT scores or the weakness of the WMS. The consistency in research findings is also open to considerable debate (10).

Most of all, Schantz et al. (3) understate the problem of multiple analyses. They stated that mercury and lead, as well as DDE and PCBs, were evaluated as exposure variables, but that all of the significant negative relationships occurred for PCBs. Consequently, they apparently conducted 96 analyses, not 48. In addition, Schantz's team analyzed motor functioning variables for the same cohort of older adults, but published the results in a separate paper (4). They conducted a variety of parametric and nonparametric analyses, although the exact number is not easily discernible; they found no relationship between PCB/DDE exposure and either hand steadiness or visual-motor coordination. Surprisingly, they blended DDE and PCB exposure to get a joint measure of contamination. Though the merger of the two might be defensible, it is not intuitive. Did the authors look at an array of statistical analyses of PCBs and DDE separately before deciding to combine the two for the published paper?

Additionally, Schantz et al. published a paper in 1996 while their analyses were still partly in the planning stage (5). In that paper, the emphasis was on two groups, fish eaters and non-fish eaters, matched on age and sex. The groups were statistically compared on a diversity of potential confounders and were generally found not to differ significantly. One of the purposes of the study was to relate consumption of contaminated fish to decline in cognitive and motor function. A second purpose was to relate serum PCB and serum DDE levels to the degree of behavioral dysfunction. Schantz and colleagues have published papers on serum levels, but the only papers that featured the fish eaters versus non-fish eaters compared the groups by potential confounders, such as alcohol consumption and general intelligence (5), and by PCB congener profiles (6). Why did they not relate fish-eating status to neuropsychologic decline? Their published research has addressed serum levels, but not fish-eating status. Do the few significant results published by Schantz et al. (3) represent the only significant findings obtained by this team of researchers despite the large number of analyses they conducted?

Schantz et al. (3) used the WAIS-R vocabulary subtest as the measure of general intelligence to control for this important confounding variable. Vocabulary is reliable, stable, and a good measure of general intelligence, and provides excellent measurement of what Horn (11) called crystallized intelligence (Gc) (7). Gc reflects knowledge and problem solving that is dependent on formal schooling and acculturation, and is referred to by Horn as a “maintained” ability, one that is maintained across the adult life span and is generally resistant to brain damage (11). In contrast, fluid intelligence (Gf), which refers to novel problem solving that is not dependent on education (such as solving abstract analogies), is a “vulnerable” ability that declines rapidly with increasing age and is vulnerable to brain damage (11). The growth curves for Gc and Gf are so different across the adult age range that there is really not a single general intelligence for adults, but two general intelligences, Gc and Gf.

The Verbal IQ yielded by the WAIS-R or the third edition of the WAIS (the WAIS-III) is roughly equivalent to Gc, whereas the Performance IQ is roughly equivalent to Gf (7). The difference in the aging patterns for these two types of general intelligence are dramatic. In an education-adjusted cross-sectional study conducted with the WAIS-III across the 20–89 year age span (12) using a common adult

reference group, Verbal IQ (Gc) averaged about 98 for ages 20–24, peaked at about 104 for ages 45–54, declined gradually to about 98 for ages 80–84, and reached its low point of 96 for ages 85–89. In striking contrast, Performance IQ (Gf) peaked at ages 20–24 (mean = 100), decreasing successively to 92 (ages 45–54), 85 (ages 65–69), 80 (ages 75–79), and 76 (ages 85–89). These data with the WAIS-III are extremely similar to cross-sectional data on the WAIS-R and to longitudinal data with independent samples (7,12).

When controlling for general intelligence for a group of older adults, such as the sample of 49- to 86-year-olds in Schantz et al.'s study (3), it is essential to control for both Gc and Gf in order to rule out the potential confounding of general intelligence. By controlling only for Gc, these investigators did not provide an adequate control for general intelligence. Because the kinds of memory and learning abilities that relate significantly to PCBs are vulnerable abilities whose aging curves more closely resemble the curves for Gf than Gc, it is especially important to rule out the potential confounding of fluid general intelligence. The investigators should have administered a measure of Gf, most notably Raven's Matrices, a fairly pure measure of fluid reasoning that is included in many epidemiologic studies, usually with a vocabulary test, to control for intelligence (13).

The problems of multiple statistical analyses, the experimental interdependence of the two CVLT tasks that related significantly to PCBs, the outdated nature of the WMS and poor qualities of Russell's delayed memory measure, and the inadequate controlling for general intelligence are flaws in the methodology used by Schantz et al. (3) and challenge the conclusions that they reached. Certainly, the problem of multiple comparisons and poor measurement of adults' general intelligence extends well beyond PCB research, affecting the interpretation of the relationship of lead exposure to IQ loss in children (13,14). Although there has been a great deal of research on the relationship of lead to IQ in children, Schantz and colleagues are the only team that has studied PCBs in older adults. Because of the flaws in their methodology, the significant findings reported by Schantz et al. (3) are best attributed to chance.

The author has been remunerated by the General Electric Company for consultation about the quality of PCB studies, but he received no remuneration for preparation of any part of this letter.

Alan S. Kaufman
Yale University School of Medicine
New Haven, Connecticut
E-mail: alanskaufman@hotmail.com

REFERENCES AND NOTES

- Schantz SL. Developmental neurotoxicity of PCBs in humans: what do we know and where do we go from here? *Neurotoxicol Teratol* 18:217–227 (1996).
- Jacobson JL, Jacobson SW, Humphrey HEB. Effects of exposure to PCBs and related compounds on growth and activity in children. *Neurotoxicol Teratol* 12:319–326 (1990).
- Schantz SL, Gasior DM, Polverejan E, McCaffrey RJ, Sweeney AM, Humphrey HEB, Gardiner JC. Impairments of memory and learning in older adults exposed to polychlorinated biphenyls via consumption of Great Lakes Fish. *Environ Health Perspect* 109:605–611 (2001).
- Schantz SL, Gardiner JC, Gasior DM, Sweeney AM, Humphrey HEB, McCaffrey RJ. Motor function in aging Great Lakes fishes. *Environ Res* 80 (2 Pt 2):S46–S56 (1999).
- Schantz SL, Sweeney AM, Gardiner JC, Humphrey HEB, McCaffrey RJ, Gasior DM, Srikanth KR, Budd ML. Neuropsychological assessment of an aging population of Great Lakes fishes. *Toxicol Ind Health* 12:403–417 (1996).
- Humphrey HEB, Gardiner JC, Pandya JR, Sweeney AM, Gasior DM, McCaffrey RJ, Schantz SL. PCB congener profile in the serum of humans consuming Great Lakes fish. *Environ Health Perspect* 108:167–172 (2000).
- Kaufman AS. *Assessing Adolescent and Adult Intelligence*. Boston, MA:Allyn & Bacon, 1990.
- Erickson RC, Scott ML. Clinical memory testing: a review. *Psychol Bull* 84:1130–1149 (1977).
- Powel J. Review of the Wechsler Memory Scale—Revised. *Arch Clin Neuropsychol* 3:397–403 (1988).
- Schell JD Jr, Budinsky RA, Wernke MJ. PCBs and neurodevelopmental effects in Michigan children: an evaluation of exposure and dose characterization. *Regul Toxicol Pharmacol* 33:300–312 (2001).
- Horn JL. Cognitive diversity: a framework of learning. In: *Learning and Individual Differences* (Ackerman PL, Sternberg RJ, Glaser R, eds). New York:Freeman, 1989:61–116.
- Kaufman AS. WAIS-III IQs, Horn's theory, and generational changes from young adulthood to old age. *Intelligence* 29:131–167 (2001).
- Kaufman AS. Do low levels of lead produce IQ loss in children?: a careful examination of the literature. *Arch Clin Neuropsychol* 16:303–341 (2001).
- Phelps L. Low-level lead exposure: implications for research and practice. *Sch Psychol Rev* 28:477–492 (1999).

PCB-Induced Impairments in Older Adults: Schantz et al.'s Response

In his letter critiquing our study of neuropsychologic functioning in older adults exposed to PCBs and DDE (1), Kaufman states that there were some serious flaws in the research design and data analysis. Specifically, he calls into question one of the test instruments we used to assess memory, argues that we did not adequately control for IQ in our analyses, and raises several statistical issues including failure to adequately control for multiple comparisons and failure to consider the implications of the interdependency of two key variables.

With regard to the issue of multiple comparisons, Kaufman charges that we

conducted at least 48 and possibly as many as 96 separate multiple regression analyses without correcting for multiple comparisons. He also implies that we may have conducted additional analyses that were not reported in the published paper. Kaufman's calculations assume that each of the four exposure variables we assessed was considered independently. That is not accurate. The four exposure measures (PCBs, DDE, lead, and mercury) were included in the same regression model, which examined PCBs and DDE as the major independent variables and lead and mercury levels as covariates. We freely acknowledged in the paper, and we reiterate here, that we did perform multiple comparisons to look at a number of different cognitive outcomes. However, as shown in Table 4 of our paper (1), the total number of comparisons was 24, not 48 or 96 or some unspecified number beyond that. We were acutely aware that the use of multiple statistical tests to assess multiple cognitive outcomes raised the possibility that one or more of the associations we encountered could be spurious, and we discussed that issue in detail in our paper. As Kaufman acknowledges in his letter, we urged caution in interpreting our results because of the multiple comparisons.

The issue of correcting for multiple comparisons is not as straightforward as Kaufman apparently believes. In recent years this topic has been the subject of spirited debate, and a number of prominent epidemiologists have argued that adjustments are unnecessary (2–4). Also, in his zeal to make his point about Type I error, Kaufman fails to consider the other side of the coin: Type II error—concluding that there is no effect when one does, in fact, exist. In addressing important public health issues, we believe that Type II error is a serious concern and should not be ignored.

It is important to point out that our study was hypothesis-driven research. We designed our neuropsychologic test battery based on previous research on the neurologic effects of PCBs, and we selected a specific subset of outcome variables from the tests we administered *a priori* on the basis of our hypotheses about which aspects of neuropsychologic function were likely to be affected. These were the only analyses we performed, and they are all reported in our published papers.

Kaufman also charges that because two key outcome variables (List A, Trial 1, and the semantic cluster ratio), both measures from the California Verbal Learning Test (CVLT), are “experimentally interdependent,” two of our significant results may be redundant. The fallacy of this argument

becomes apparent if we consider the correlations between several CVLT measures. The Spearman correlation between List A, Trial 1, and semantic cluster ratio in our sample was a modest 0.376, yet both of these outcomes were significantly associated with log PCB in our models. In contrast, the Spearman correlation between List A, Trial 1, and List A, Trials 1–5, was much greater (0.752), but one of these measures was significantly associated with log PCB and the other was not. This illustrates that we cannot assume that two correlated outcomes will both be significant (or nonsignificant). Finally, although we do not feel it is necessarily the best approach to control for other outcome variables in the regression model, we repeated the regression analysis and found that the association between semantic cluster ratio and log PCB remained significant even when we controlled for List A, Trial 1, in the analysis.

According to Kaufman, another flaw in our research design was the use of an outdated version of the Wechsler Memory Scale (WMS), when a newer, more reliable version, the Wechsler Memory Scale-Revised (WMS-R) was available. He cites one paper (5), which gave the WMS-R “exceptional reviews.” We based our decision not to use the WMS-R on a careful review of all of the literature that was available at the time. Although favorably reviewed by Powel (5) and Holden (6), the WMS-R has been sharply criticized by a number of others. Loring (7) pointed out that in the development of the WMS-R, advances in cognitive, experimental, and clinical psychology over the decades since the introduction of the original WMS (8) were largely ignored. In addition, the WMS-R has been faulted on basic psychometric properties including small sample size and poor subtest reliability (9) as well as interpolation of scores used as norms for one-third of the population (10,11). Finally, two separate reviews in the *Mental Measurements Yearbook* (12,13) were both highly critical of the WMS-R. Although we do not feel that Kaufman’s concerns about the WMS are valid, we would like to point out the fallacy of his argument. He implicitly assumes that use of an “outdated” measure with “poor psychometric properties” would be more likely to lead to a false positive than a false negative association, but he fails to provide a rationale for this assumption. We argue that an unreliable test would be more likely to attenuate correlations than to result in spurious associations.

Kaufman is also critical of our choice of the WAIS-R vocabulary subtest as a measure to control for general intelligence. He acknowledges that WAIS-R vocabulary is a

“reliable, stable and good” measure of general intelligence—precisely the reasons we selected it—but he goes on to argue that vocabulary taps primarily “crystallized intelligence” and does not adequately control for “fluid intelligence,” which he considers to be particularly important in older adults. Kaufman raises an interesting point, but in reality this issue is not as simple as he makes it seem. As he himself acknowledges, fluid intelligence is vulnerable to brain damage, so using it as a control for general intelligence in the presence of exposures that have the potential to damage the brain is of questionable utility. Furthermore, although it is possible that including the Raven Matrices (or alternatively the entire WAIS-R) would have provided better overall control of general intelligence, this would have added significantly to the time required to administer the test battery. The subjects in our study were aging adults who were evaluated for approximately 3–4 hr in their homes. Pilot testing prior to the study indicated that people in this age group were not receptive to a longer testing battery and that fatigue became a factor in test performance if the visit was extended any longer. The homes of the study participants were located 90–200 miles from the research office; thus we did not have the resources for more than one visit per subject. In the selection of dependent variables, we considered the hypotheses to be tested, the instruments available to us at the time, and the amount of time we realistically had available to do the assessments. No single research study will be definitive in every conceivable respect. As Needleman and Bellinger (4) aptly pointed out when Kaufman leveled similar criticisms regarding their lead studies, “... complete control of all confounders is an unattainable goal in real-world epidemiology” (p. 363).

We thank Kaufman for his thorough critique of our manuscript and the editors of *EHP* for giving us the chance to address the misconceptions concerning the design of our study and the statistical analyses performed on the data. As we stated in the original article (1): “[Our] study suggests ... that PCB exposure during adulthood *may* [emphasis added] be associated with impairments in certain aspects of memory and learning” (p. 610) and “... it would be prudent to interpret the findings with caution until they have been replicated in an independent exposure cohort” (p. 610).

Susan L. Schantz
University of Illinois at
Urbana-Champaign
Urbana, Illinois
E-mail: schantz@uiuc.edu

Donna M. Gasior
Elena Polverejan
Harold E.B. Humphrey
Joseph C. Gardiner
Michigan State University
East Lansing, Michigan
Robert J. McCaffrey
University at Albany
State University of New York
Albany, New York
Anne M. Sweeney
University of Texas Health
Science Center at Houston
Houston, Texas

REFERENCES AND NOTES

- Schantz SL, Gasior DM, Polverejan E, McCaffrey RJ, Sweeney AM, Humphrey HEB, Gardiner JC. Impairments of memory and learning in older adults exposed to polychlorinated biphenyls via consumption of Great Lakes fish. *Environ Health Perspect* 109:605–611 (2001).
- Poole C. Multiple comparisons? No problem! *Epidemiology* 2:241–243 (1991).
- Rothman K. No adjustments are needed for multiple comparisons. *Epidemiology* 1:43–46 (1990).
- Needleman HL, Bellinger D. Studies of lead exposure and the developing nervous system: a reply to Kaufman. *Arch Clin Neuropsych* 16:359–374 (2001).
- Powel J. Review of the Wechsler Memory Scale-Revised. *Arch Clin Neuropsych* 3:397–403 (1988).
- Holden R. Wechsler Memory Scale-Revised. In: *Test Critiques*, Vol. VII (Keyser D, ed). Kansas City, MO:Test Corporation of America, 1988;633–638.
- Loring DW. The Wechsler Memory Scale-Revised, or the Wechsler Memory Scale-Revisited? *Clin Neuropsychol* 3:59–69 (1989).
- Wechsler D. A standardized memory scale for clinical use. *J Psychol* 19:87–95 (1945).
- Elwood R. The Wechsler Memory Scale-Revised: psychometric characteristics and clinical application. *Neuropsychol Rev* 2:179–200 (1991).
- D’Elia L, Satz P, Schretlen D. Wechsler Memory Scale: a critical appraisal of the normative studies. *J Clin Exp Neuropsychol* 11:551–568 (1989).
- Chelune GJ, Borenstein RA, Prifitera A. The Wechsler Memory Scale-Revised: current status and applications. In: *Advances in Psychological Assessment*, Vol. 7 (Rosen J, McReynolds P, Chelune GJ, eds). New York:Plenum Press, 1990;66–95.
- Huebner EC. Review of the Wechsler Memory Scale-Revised. In: *The Eleventh Mental Measurements Yearbook* (Kramer JJ, Conoley JC, eds). Lincoln, NE: University of Nebraska Press, 1992;1023–1024.
- Reinehr RC. Review of the Wechsler Memory Scale-Revised. In: *The Eleventh Mental Measurements Yearbook* (Kramer JJ, Conoley JC, eds). Lincoln, NE: University of Nebraska Press, 1992;1024–1025.

The Melatonin Hypothesis: A Matter of Method

In their paper in *EHP*, Graham et al. (1) addressed the effects of electromagnetic fields (EMF) and light at night (LAN) on hormonal regulation, particularly of melatonin and estradiol. The study provides evidence for the predicted role of exposure to LAN on melatonin levels, which marks the credibility of the setting. However, I am concerned that the relationship between breast cancer and EMF and LAN has been reduced to the mechanistic level of estradiol

concentrations, ignoring the importance of the two major factors associated with breast cancer risk: *a*) prolactin levels as a consequence of hypothalamus–pituitary regulation, which act as stimulants for proliferation of the tissue of origin, increasing chances of mutation; and *b*) depressiveness as a consequence of low melatonin levels, which are related to lack of sleep. In addition, insomnia at night leads to leaving lights turned on for reading or other activities instead of remaining in the dark, when melatonin is best synthesized; this lack of darkness decreases melatonin synthesis and thus increases the incidence and degree of effects of depression and depressiveness. In addition, women in an experimental situation do not react the same way as stressed, exhausted, unhappy, or desperate women in daily life who are not stimulated by the setting of an experiment that increases attention and functions otherwise more likely down-regulated than not.

The influence of electricity on the ubiquitous availability of light as an important stimulus, which aids against depressiveness—with and without insomnia—has been described by Davis et al. (2). For the time being the LAN theory has not lost its value, but it seems not to be open for experimental approaches such as the one published by Graham et al. (1).

Rainer Frentzel-Beyme

Bremen Institute for Prevention Research
and Social Medicine
Bremen, Germany
E-mail: beyme@bips.uni-bremen.de

REFERENCES AND NOTES

1. Graham C, Cook MR, Gerkovich MM, Sastre A. Examination of the melatonin hypothesis in women exposed at night to EMF or bright light. *Environ Health Perspect* 109:501–507 (2001).
2. Davis S, Mirick DK, Stevens RG. Night shift work, light at night, and risk of breast cancer. *J Natl Cancer Inst* 93:1557–1562 (2001).

Examination of the Melatonin Hypothesis: Graham et al.'s Response

We appreciate Frentzel-Beyme's comments about our study (1). Our focus was to determine if melatonin and estradiol concentrations are altered, as suggested by the melatonin hypothesis (2), in women exposed to magnetic fields (EMF) at night or to bright light at night (LAN). Frentzel-Beyme expresses concern that we have reduced the relationship between EMF/LAN and breast cancer to the mechanistic level of estradiol concentrations, and ignored the important roles played by prolactin, depression,

insomnia, and LAN in the etiology of the disease. Further, the relevance of our negative results is questioned because hormonal regulation in healthy women, who are supposedly stimulated by being in a laboratory environment, is different from that which occurs in stressed women in real life. Frentzel-Beyme concludes that, for now, this area of environmental health effects does not seem open to experimental approaches such as the one we published.

We take a different point of view. Richard Stevens developed one of the few mechanistic, testable hypotheses in the area of EMF research: namely, that the increased incidence of breast cancer in industrial societies is related to greater exposure to power-frequency EMF and/or the presence of high levels of LAN (2). EMF and LAN are believed to reduce circulating levels of the hormone melatonin which, in turn, allows estrogen levels to rise and stimulate the turnover of breast epithelial stem cells and increase the risk for malignant transformation. This hypothesis has heuristic value precisely because it does describe a mechanistic relationship between environmental exposure and neoplastic disease, one that is subject to experimental observation and manipulation. We believe that testing specific hypotheses under controlled experimental conditions is the foundation of science, and that this process has led to many important advances relevant to human health.

We also disagree with some of the conclusions drawn by Frentzel-Beyme. We did not ignore the hormone prolactin; it simply was not part of the hypothesized chain of events we set out to test. As noted in the recent review on EMF health effects by the National Institute of Environmental Health Sciences (3), studies of EMF exposure in both healthy individuals and electrically hypersensitive people have failed to observe alterations in prolactin concentrations. More generally, the results of multiple human EMF exposure studies provide little evidence for any reliable effect on hypothalamic, pituitary, thyroid, or adrenal hormonal systems.

We see a number of difficulties with Frentzel-Beyme's reasoning. Melatonin levels tend to be stable within an individual, but vary widely from person to person. For example, women in our study showed a 15-fold difference in the total amount of melatonin they secreted overnight (area under the curve range: 86–1,296 pg/mL), and this is not an unusual observation (4). Depression is simply not a function of having low melatonin levels. In fact, endogenous low melatonin levels in humans do not seem to correlate with much of anything. Furthermore, the melatonin rhythm

is a function of the light/dark cycle, not the sleep/wake cycle; thus, the quality of night sleep (or the lack of it) does not alter the nightly rhythm of this hormone. As we reported, even when extremely bright LAN is used to cause a marked (> 90 %) reduction in nocturnal blood levels of melatonin, the natural rhythm is rapidly reinstated in humans after the light is discontinued.

As indicated in our paper (1), we certainly agree with Frentzel-Beyme on the need for further research on LAN and its impact on health, particularly as it relates to shift work. Although melatonin may not be responsive to the sleep/wake cycle, prolactin certainly is. We also feel that the issues raised by Frentzel-Beyme involving LAN, depression, and the deregulation of hormones implicated in carcinogenesis are quite amenable to experimental approaches such as we described in our paper (1). For example, it would be quite feasible to assess the hormonal consequences of controlled exposure to EMF or LAN in healthy women compared to women stratified on various measures of depression, insomnia, or other factors. One should bear in mind, however, that if the initiation of breast cancer were limited only to those women who are "stressed, depressed, exhausted, unhappy, and desperate" prior to diagnosis, prevention would be a much easier matter than it is now. Anti-depressant medications are not a currently recommended prophylactic for breast cancer, and sadly, many women who are happy and well-adjusted develop this disease.

Charles Graham

Mary R. Cook
Mary M. Gerkovich
Antonio Sastre
Midwest Research Institute
Kansas City, Missouri
E-mail: mcg@planetkc.com

REFERENCES AND NOTES

1. Graham C, Cook MR, Gerkovich MM, Sastre A. Examination of the melatonin hypothesis in women exposed at night to EMF or bright light. *Environ Health Perspect* 109:501–507 (2001).
2. Stevens RG. Electric power use and breast cancer: a hypothesis. *Am J Epidemiol* 125:556–561 (1987).
3. Portier C, Wolfe MS, eds. *Assessment of Health Effects from Exposure to Power-Line Frequency Electric and Magnetic Fields*. NIH Publication no 98-3981. Research Triangle Park, NC: National Institute of Environmental Health Sciences, 1998.
4. Cook MR, Graham C, Kavet R, Stevens RG, Kheifets LI, Davis S. Morning urinary assessment of nocturnal melatonin secretion in older women. *J Pineal Res* 28:41–47 (2000).