Eli Lilly and Company Comments on the NTP-CERHR Expert Panel Report on the Reproductive and Developmental Toxicity of Fluoxetine

The report of the NTP-CERHR Expert Panel on the Reproductive and Developmental Toxicity of Fluoxetine was recently issued for public comment. Early on in their evaluation, the Expert Panel concluded that the potential for environmental exposure to fluoxetine was trivial. The NTP-CERHR Expert Panel went on to evaluate the safety of fluoxetine for patients.

Comments by Eli Lilly and Company on the specific findings in the report of the NTP-CERHR Expert Panel can be summarized as follows:

1. There is not sufficient evidence to conclude that fluoxetine causes developmental toxicity manifested as poor neonatal adaptation, increased incidence of minor anomalies, reduced gestation or birth weight, or reduced size of breastfed infants.

2. Beneficial or unwanted side effects of SSRIs on sexual function appear to be reversible and any risks should be evaluated by patients and their physicians. Since effects on sexual function from the use of SSRIs appear to be confounded by the disease state and to be based on reversible pharmacology, fluoxetine should not be classified as a reproductive toxin in humans.

3. Untreated depression can result in substantial medical risks. Fluoxetine is approved in the United States for the treatment of major depressive disorder. Fluoxetine is only available through prescription, so the benefits and potential risks of use should be evaluated on a case-by-case basis by patients and their physicians.

4. Communications about the safety of patients using pharmaceuticals should remain the responsibility of the US FDA.

1. Developmental Toxicity of Fluoxetine

According to the Expert Panel (Section 3.4 Summary),

> *"…there is sufficient evidence in humans to determine that prenatal exposure to fluoxetine results in poor neonatal adaptation (e.g., jitteriness, tachypnea, hypoglycemia, hypothermia, poor tone, respiratory distress, weak or absent cry, diminished pain reactivity, or desaturation on feeding) at typical therapeutic exposures (20–80 mg/day orally) during the third trimester of pregnancy.";*

> *"Therapeutic fluoxetine exposure during early pregnancy may result in an increased incidence of minor anomalies.";*
>
> *"Shortening of gestation and reduced birthweight are also suspected, although the evidence is not sufficient to exclude the underlying disorder, depression, as a cause or contributor to these effects.";*
>
> and
>
> *"The evidence is suggestive that exposure to fluoxetine through breast milk can result in reduced infant growth; however, these effects may be related to prenatal exposure."*

Eli Lilly and Company does not believe that there is sufficient evidence to support any of the conclusions by the NTP-CERHR Expert Panel regarding developmental toxicity.

Poor Neonatal Adaptation

There are published reports about increases in a variety of clinical observations in newborns from mothers treated with SSRIs during the third trimester of pregnancy. The observations resolve within days after birth. Their cause cannot be effectively separated from the underlying disease condition and other confounding factors. Some of the clinical observations are consistent with observations that might be anticipated from pharmacological activity, or withdrawal from exposure to an SSRI. Many of these observations have also been reported in babies from mothers suffering from untreated depression. Given the minor and transient nature of the observations and their similarity to some of the observations that might be expected from pharmacological activity, Eli Lilly and Company does not believe that there is sufficient evidence to classify fluoxetine as a developmental toxin.

The Expert Panel appears to have primarily based their conclusion on the work of Chambers et al. (1996). The Expert Panel noted some weaknesses in this report, including a small number of subjects and models that have many covariates (due to confounding factors such as smoking, alcohol consumption, use of other psychotherapeutic medications, and differences in maternal age). Other epidemiologists consider the study conclusions to be limited precisely because of these confounding factors (Llewellyn et al., 1997).

The description of "poor neonatal adaptation" itself is not a clinically defined category for a distinct pattern of responses that are routinely associated with newborns from women treated with fluoxetine. Instead it is a catch-all descriptor for individual or multiple observations that, when combined, form a statistically significant grouping. All the observations used to define this catch-all grouping are not clearly related to the pharmacology of fluoxetine. Chambers et al. (1996) did not list the frequency of findings for each type of observation, only the frequency for the combined grouping. It is not

clear if the statistical significance of this grouping would change if the observations were grouped differently.

Weaknesses of the study by Chambers et. al (1996) have been summarized by Wisner et al. (1999) and Robert (1996). Probably the most significant weakness of this study was that the control group does not allow separation of fluoxetine effects from those of the underlying maternal depression. The control group for the study by Chambers et al. (1996) consisted of normal pregnant women who called the California Teratogen Information Service and Clinical Research Program with questions about drugs and procedures not considered to be teratogenic. This control group had a lower incidence of confounding factors than the fluoxetine treatment groups. There was no assessment of the incidence of untreated depression in the control group. For the treatment groups, fluoxetine was used either early in pregnancy (exposed-early treatment group) or throughout pregnancy for most of the rest of the women (exposed-late treatment group). Comparison between the exposed-early group and the exposed-late group led Chambers et al. (1996) to conclude that exposure to fluoxetine in late pregnancy increases the risk of poor neonatal adaptation. This finding might also be explained by the fact that severe depression could have required treatment throughout pregnancy in the exposed-late group (Robert, 1996). The exposed-early group may have had a mild form of depression, allowing treatment with fluoxetine to be stopped. From this study, there is no way to determine if the results represented the effects of the underlying severity of the disease and maternal condition or long-term exposure to fluoxetine. Chambers et al. (1996) acknowledged that the "extent to which these findings may be due to the underlying maternal condition is unknown."

The risks of untreated depression can include poor nutrition, disrupted sleep patterns, difficulty following medical and prenatal care recommendations, suicide, worsening of co-morbid medical illness, and increased exposure to tobacco, alcohol, or drugs (Llewellyn et al., 1997). Maternal depressive symptoms can also lead to prematurity and low birth weight (Orr and Miller, 1995; Steer et al. 1992). Pre-term birth can be associated with postnatal complications. Clearly, without treatment for depression, pregnant women and their infants can face significant health risks. Even with antidepressant treatment, the severity of depression and maternal condition can complicate the interpretation of results from studies like those of Chambers et al. (1996).

Increased Incidence of Minor Anomalies

The expert panel states "Therapeutic fluoxetine exposure during early pregnancy may result in increased incidence of minor anomalies." Eli Lilly and Company does not believe that there is sufficient evidence for this statement.

The expert panel based their conclusion regarding minor anomalies on the work of Chambers et al. (1996). As previously noted, there are weaknesses in the design of the work by Chambers et al. (1996) that preclude any definitive conclusions regarding the potential for fluoxetine to cause minor anomalies in infants. These weaknesses include a

small number of subjects, multiple confounders due to factors such as smoking, alcohol consumption, use of other psychotheropeutic medications and differences in maternal age, and the inability to separate fluoxetine effects from those of the underlying maternal depression.

It is critically important to note that the statistical evaluation of minor anomalies utilized a univariate categorical analysis in which confounding risk factors were not controlled. The increase in minor anomalies reported by Chambers et al. (1996) was only statistically significant when all of the confounding risk factors were left uncontrolled. Acknowledging this significant design flaw and potential bias in outcome, Chambers et al. (1996) re-evaluated the minor anomaly data excluding infants from mothers that were also exposed to benzodiazepines. When infants exposed to benzodiazepines were excluded from analysis, Chambers et al. (1996) reported that the incidence of minor anomalies were not significantly different between the fluoxetine exposed and control groups. The change in study outcome when controlling for only one confounding risk factor highlights the importance of proper study design and controls. Given the lack of significant differences in minor anomalies when infants exposed to benzodiazepines were removed from the population and the lack of control for the other potential confounding risk factors, no definitive conclusions regarding increased minor anomalies can be drawn from the Chambers et al. (1996) work.

Multiple publications have addressed the characteristics, criteria and evidence that are needed to conclude that an environmental agent or a drug produces structural anomalies in humans (Brent 1978, Brent 1995 and Shepard 1998). Of primary importance in these criteria are A) proven exposure, B) consistent findings in two or more high quality epidemiology studies, including the control of confounding factors, C) careful delineation of the clinical cases identifying a specific defect or syndrome of effects, and D) similar positive findings in an animal model demonstrating a dose-response. The available data on fluoxetine does not fit the above criteria and does not support drawing the conclusion that therapeutic fluoxetine exposure during early pregnancy may result in an increased incidence of minor anomalies. This position is based on the following.

A. There was no quantitative proof of exposures in the study by Chambers et al. (1996). Analytical analysis of fluoxetine in maternal blood was not conducted in the Chambers et al. (1996) study. Fluoxetine exposure was assumed based on patient and physician response or records, however, quantitative exposure-response relationships could not be established.

B. There are no consistent findings in two or more high quality epidemiology studies demonstrating a significant effect on minor anomalies in humans. Only one human study deemed acceptable by the expert panel reported an increase in minor anomalies following fluoxetine exposure (Chambers et al. 1996). This study, however, cannot be considered of high quality relative to evaluating minor anomalies due to failure to control for multiple confounding factors. As previously described, when just one of the confounding risk factors (exposure to

benzodiazepines) was removed from the analysis, the authors reported that there was no significant difference in minor anomalies between the fluoxetine and control groups. Furthermore, the selection of infants for evaluation was apparently not conducted randomly or systematically with specific selection criteria. The lack of scientifically rigorous and acceptable selection criteria introduces additional bias and further compromises the validity of the study conclusions regarding minor anomalies.

C. There is no specific defect or syndrome of effects following fluoxetine exposure. Most, if not all agents known to cause malformations in humans are associated with a specific syndrome of effects or a specific defect (Shepard 1998). Chambers et al. (1996) concluded that no pattern of anomalies was recognized. Only when the data was grouped such that 132 different types of minor anomalies were considered, did the authors find statistical significance between the fluoxetine treated group and the control group. This difference was not present when controlling for one of the many risk confounders.

D. Similar findings of increased minor anomalies in animal studies have not been demonstrated. Two scientifically rigorous studies evaluating anomalies in rats and rabbits were conducted by Byrd and Markham (1994). Neither of these studies demonstrated an increase in fetal anomalies following exposure. Both of these studies were reviewed by the expert panel and the FDA and were found to be of high quality and sufficient to evaluate the effects of fluoxetine on embryo and fetal development in an animal model. Importantly, the expert panel also concluded that "The rat and rabbit data are assumed relevant to consideration of human risk."

Shortening of Gestation and Reduced Birth Weights

The expert panel suspected that exposure to fluoxetine could lead to shortening of gestation and reduced birth weights. Eli Lilly and Company does not believe that there is sufficient evidence to support this conclusion.

One of the two references quoted to support this assertion (Simon et al., 2002) was a study of SSRI and tricyclic antidepressants. The statistical evaluations were by class of compounds. The authors noted that any statistical significance disappeared when the effects of individual pharmaceuticals were evaluated. Minor, but statistically significant differences by class appeared to be as influenced by changes in the mean values for the randomly selected controls as they were by changes due to exposure to a class of chemicals. For example, the mean estimated gestational age for infants with mothers treated with tricyclic antidepressants was 38.8 ± 1.9 weeks. The mean estimated gestational age for infants with mothers treated with SSRI antidepressants was 38.5 ±1.8 weeks. The result for the tricyclic antidepressant was not different from the mean gestational age (39.1±1.7 weeks) of a randomly selected set of infants from a matched set of untreated mothers. Another random selection of infants to match the SSRI treatment

resulted in a gestational age of 39.4 ±1.5 weeks for infants from untreated mothers. The increase in the control gestational age and reduction in variability contributed at least as much as the slight decline in gestational age of the SSRI treatment group to allow the authors to claim a statistically significant difference for the SSRI treatment. This control bias is also present in the odds ratio evaluation for the percent of infants with gestational ages less than or equal to 36 weeks. Birth weight differences found in this study disappeared when normalized for gestational age. While the authors of this paper caution about the over interpretation of marginally significant results given the large number of comparisons, they go on to draw large conclusions from minor differences found for data sets that could have a control bias. They also suggest that the outcome of their research could be confounded by the underlying medical condition. Symptoms of maternal depression can themselves lead to prematurity and low birth weight (Orr and Miller, 1995; Steer et al. 1992). Based on this information, the study by Simon et al. (2002) should not be the basis for the NTP-CERHR Expert Panel to suspect shortened gestation or reduced birth weights for infants from mothers treated with fluoxetine.

The Expert Panel also referenced Chambers et al. (1996) for their suspicion about birth weight. The average birth weights listed by Chambers et al. (1996) for controls and treatments are within the normal range for infants at birth. There was a higher proportion of infants with gestational ages less than 37 weeks for women treated throughout their pregnancy for depression (10/70) than for women treated only in the first and second trimester (4/98) or for untreated controls (13/220). Chambers et al. (1996) did not determine whether this was due to treatment with fluoxetine or to the severity of the symptoms of the underlying depression. As already described, the symptoms of depression can lead to prematurity and low birth weight.

Since neither of the studies quoted by the NTP-CERHR panel evaluated the effect of depression as a variable on gestational age or birth weight, the panel has no real basis in humans to suspect that treatment of women with fluoxetine results in either of these effects in infants.

Reduced Size of Breastfed Infants

The NTP-CERHR Expert Panel indicated that there is suggestive evidence about the size of breastfed infants. This appears to be primarily based on the presence of fluoxetine in breast milk and the results from a different study by Chambers et al. (1999). Chambers et al. (1999) indicated "that infants who are breastfed by mothers who take fluoxetine track a growth curve significantly below that of infants breastfed without medication." Fluoxetine levels were not actually measured in breast milk by Chambers et al. (1999), but infant exposure was assumed to have occurred from breastfeeding. They did note, however, that the possibility of direct effects of fluoxetine on weight gain in nursing infants is not supported by the dose they could have received. Less than 10 percent of a maternal dose of fluoxetine is transferred to a nursing infant (Taddio et al., 1996). Chambers et al. (1999) also acknowledged "women with an underlying condition requiring a psychotherapeutic medication may breastfeed less often and engage in other

behaviors that influence postnatal weight gain in their infants." So the results of this study do not support the NTP-CERHR Expert Panel conclusions about the direct effects of fluoxetine on the size of breastfed infants. Chambers et al. (1999) even wrote that "there is no evidence from these data to indicate that mothers who breastfeed their infants while taking fluoxetine should be concerned about side effects attributable to the medication."

2. Reproductive Toxicity of Fluoxetine

According to the Expert Panel (Section 4.4 Summary),

> *…there is sufficient evidence in humans that fluoxetine produces reproductive toxicity in men and women manifested as impairment of sexual function, specifically orgasm.*

The Expert Panel also noted that effects on individual sexual performance are unpredictable. The Expert Panel acknowledged that, in many instances, it is not possible to differentiate drug-induced adverse effects from those induced by the disease process itself or the pharmacological action of the drug. Depression is associated with impaired sexual function, and successful treatment of depression may be associated with improvements in sexual function. Some evidence does suggest that SSRIs can cause untoward sexual experiences. At least one report indicates that improvement in sexual function (reversal of sexual dysfunction) occurs when the dose of an SSRI is diminished or the drug is withdrawn (Montejo-Gonzalez et. al, 1997). Reliable estimates of the incidence and severity of these experiences involving sexual desire, performance, and satisfaction are difficult to obtain. While it is difficult to know the precise risk of sexual dysfunction associated with the use of SSRIs, the product information for fluoxetine recommends that physicians should routinely inquire about such possible side effects with their patients. Important to note is that sexual function is not necessarily equivalent to reproductive performance. Since effects on sexual function from the use of SSRIs can be confounded by the disease state and appear to be based on reversible pharmacology, fluoxetine should not be classified as a reproductive toxin in humans.

3. Risk/Benefit Considerations in the Use of Fluoxetine

Subpopulation of patients: Children

The U.S. Food and Drug Administration recently approved Prozac® (fluoxetine hydrochloride) for the treatment of major depressive disorder and obsessive-compulsive disorder in the pediatric population. The clinical data reviewed by the FDA indicate that Prozac 20 mg has a comparable profile of safety and efficacy in both adults and children. Eli Lilly and Company submitted four pediatric studies: two clinical studies for depression, one for obsessive-compulsive disorder, and one pharmacokinetic study. The FDA based its decision on placebo-controlled clinical trials.

Lilly is working closely with the FDA on the design of a Phase IV post-marketing study to further evaluate whether there is any long-term effect on either weight or height of children who take Prozac. The FDA stated in its announcement that "the clinical significance of height and weight differences on long-term growth is unknown." (www.fda.gov) The Phase IV studies should help provide additional information in this area. Nonetheless, the FDA has found the product to be both safe and effective as seen in its approval.

On December 10, 2003 the MHRA (Medicines and Healthcare Products Regulatory Agency) in Europe announced, "On the basis of a review of the safety and efficacy of the SSRI class in the treatment of pediatric major depressive disorder undertaken by the Expert Working Group of the Committee on Safety of Medicines (CSM), the CSM has advised that the balance of risks and benefits for the treatment of major depressive disorder in under 18s is judged to be unfavorable for sertraline, citalopram and escitalopram and unassessable for fluvoxamine. Only fluoxetine (Prozac) has been shown in clinical trials to have a favorable balance of risks and benefits for the treatment of MDD in the under 18s" (www.mhra.gov.uk). The MHRA has asked Eli Lilly and Company to submit documents for the pediatric approval of Prozac in Europe.

General Patient Population

The stated purpose of the NTP-CERHR is to address the "widespread concern among health professionals, environmental scientists, and the public that environmental exposures may be contributing to human reproductive and developmental disorders". The chemicals previously reviewed by the NTP-CERHR reach the general population through environmental exposure. Some of these chemicals have the potential for a health risk, with no potential for a health benefit. The results of exposure of individuals to most environmental chemicals are only rarely monitored by a health care professional.

Unlike the chemicals that can reach entire populations by environmental exposure, fluoxetine is available only through prescription. Individuals are monitored by physicians qualified to evaluate the potential benefits and risks of fluoxetine to their patients. All of the approved uses of fluoxetine have been evaluated by the U.S. Food and Drug Administration.

Product information that describes potential benefits and risks of fluoxetine has been made available to physicians and the public. For example, the presence of fluoxetine in breast milk is noted in the product information and nursing is not recommended. Eli Lilly and Company also recently agreed to class labeling to include the published associations of poor neonatal adaptation in babies from mothers with depression who were treated with SSRIs. It is noted in the product information that fluoxetine should only be used during pregnancy, labor and delivery if the benefit justifies the potential risk. Determining the potential for certain developmental or reproductive hazards from the use of fluoxetine is complicated by the same hazards that result from maternal

depression. Improvement in depression from the use of fluoxetine might reduce the risk of these hazards. This risk/benefit consideration is best made on a case-by-case basis by physicians and their patients.

4. Communications about pharmaceutical safety in patients

The role adopted by the Expert Panel to review literature for a registered pharmaceutical product overlaps the statutory responsibility of the United States Food and Drug Administration (US FDA). The US FDA collects published information and proprietary data available for its safety and efficacy reviews and requires detailed explanations of the risks and benefits from the use of a pharmaceutical to be packaged with the product. While the report of the Expert Panel has no direct regulatory implications for the use of fluoxetine, generalized media reports already written about the findings of the "government report" have the potential to alarm people being treated for depression into discontinuing their treatment before they talk to their physicians. Patients could even decide to unnecessarily terminate healthy pregnancies.

The US FDA has the sole regulatory responsibility in the US government for evaluating the safety and efficacy of pharmaceuticals, and for determining appropriate communications to the public about them. Terminologies used in the NTP-CERHR report were developed for population exposures to environmental chemicals. The guidelines for the NTP-CERHR Expert Panel actually have no narrative criteria for observed effects that warrant categorizing a pharmaceutical as having reproductive or developmental toxicity. Nor do the guidelines provide direction for evaluating therapeutic drugs that may improve reproductive performance. The panel members had to partially develop and apply a new definition of reproductive toxicity in the report for fluoxetine, indicating that reversible pharmacology qualifies as toxicity to patients treated for depression, apparently even if reproductive function is improved. Even though the NTP-CERHR Expert Panel provides a thorough review of published information, interpreting the information for pharmaceuticals and communicating potential risks to patients should remain the responsibilities of the US FDA.

References

Byrd R, Markham J. 1994. Developmental toxicity studies of fluoxetine hydrochloride administered orally to rats and rabbits. Fund Appl Toxicol 22:511-518.

Brent RL. 1978. Editor's note. Teratology 17:183.

Brent RL. 1995. Bendectin: review of the medical literature of a comprehensively studied human nonteratogen and the most prevalent tortogen-litigen. Reproductive Toxicology 9(4):337-349.

Chambers CD, Anderson PO, Thomas RG, Dick LM, Felix, RJ, Johnson KA, Jones KL. 1999. Weight gain in infants breastfed by mothers who take fluoxetine. Pediatrics 104(5):e61, p 1-5.

Chambers CD, Johnson KA, Dick LM, Felix RJ, Jones KL. 1996. Birth outcomes in pregnant women taking fluoxetine. N Engl J Med 335:1010-1015.

Llewellyn AM, Stowe ZN, Nemeroff CB. 1997. Depression during pregnancy and the puerperium. J Clin Psychiatry 58(suppl 15):26-32.

Montejo-Gonzalez AL, Llorca G, Izquierdo JA, Ledesma A, Bousono M, Calcedo A, Carrasco JL, Ciudad J, Daniel E, De la Gandara J, Derecho J, Franco M, Gomez MJ, Macias JA, Martin T, Perez V, Sanchez JM, Sanchez S, Vicens E. 1997. SSRI-induced sexual dysfunction: fluoxetine, paroxetine, sertraline, and fluvoxamine in a prospective, multicenter, and descriptive clinical study of 344 patients. J Sex Marital Ther 23:176-194.

Orr ST, Miller CA. 1995. Maternal depressive symptoms and the risk of poor pregnancy outcome. Review of the literature and preliminary findings. Epidemiologic Rev 17:165-171.

Robert, E. (1996) Treating depression in pregnancy. N Engl J Med 335: 1056-1058.

Shepard TH. 1998. Catalog of teratogenic agents, 9th ed. Baltimore, MD: Johns Hopkins Univ. Press.

Simon GE, Cunningham ML, Davis RL (2002) Outcomes of prenatal antidepressant exposure. Am J Psychiatry 159 (12): 2055-2061.

Steer RA, Scholl TO, Hediger ML, Fischer RL. 1992. Self-reported depression and negative pregnancy out comes. J Clin Epidemiol 45:1093-1099.

Taddio, A., Ito, S., Koren, G. (1996) Excretion of fluoxetine and its metabolite, norfluoxetine, in human breast milk. J Clin Pharmacol 36: 42-47.

Wisner, K.L., Gelenberg, A.J., Leonard, H., Zarin, D., Frank, E. (1999) Pharmacologic treatment of depression during pregnancy. JAMA 282(13):1264-1269.