rilvm

research for
man and
environment

RIVM Reportnumber 650040 001

Developmental and testicular toxicity of butyl
benzyl phthalate in the rat and the impact of study design

AH Piersma, A Verhoef, JAMA Dormans, LH Elvers,
V de Valk, JD te Biesebeek, MN Pieters, W Slob

October 1999

Mailinglist

1.   Directeur Generaal Volksgezondheid, Dr. H.J.Schneider
2.   Directeur Gezondheidsbeleid, Dr.W.H.van Eck
3.   Voorzitter Gezondheidsraad, Prof.dr. J.J. Sixma
4.   Drs.G.E.H.Houben
5.   Dr.H.Roelfzema
6.   Drs.N.B.Lucas Luijkx
7.   Dr.ir.P.C.Bragt
8.   Dr.ir. G.Kleter
9.   Dr.J.A.van Zorge
10.   Directie RIVM
11.   Dr.ir.G.de Mik
12.   Dr.A.Opperhuizen
13.   Dr.W.H.Könemann
14.   Dr.W.C.Mennes
15.   Dr.A.K.D.Liem
16.   Dr.ir.E.H.J.M. Jansen
17.   Dr.P.W.Wester
18.   Ir.M.Hof
19. – 25 Auteurs
26.   Depot voor Nederlandse Publikaties en Nederlandse Bibliografie
27.   Voorlichting en Public Relations RIVM
28.   Bibliotheek RIVM
29.   Bureau Rapportenregistratie
30. – 45 Bureau Rapportenbeheer
46. – 70   Reserve exemplaren

Contents

Summary

The developmental toxicity of butyl benzyl phthalate was investigated in the rat in an alternative study design using ten treatment groups. The effect of exposure duration was studied and a comparison of pregnant versus nonpregnant females was made.The classical data analysis via no-observed-adverse-effect-levels (NOAEL) was compared to an alternative benchmark approach in which critical effect dosages (CED) are derived. In addition, a pilot study was carried out in young males in order to assess the testicular toxicity of the compound. This report describes the results obtained together with an initial analysis of the data. A more extensive data analysis will be presented in a consecutive report.

Comparison of the lowest NOAEL with lowest CED shows that in both analyses liver effects determine maternal toxicity, with liver weight NOAEL=450 mg/kg body weight/day (mkd) and serum ALAT CED20=456 mkd. For embryotoxicity, in both analyses fetal testicular effects appear critical, with fetal testicular weight NOAEL<270 mkd (lowest dose tested) and fetal testicular caudal migration CED01=164 mkd. As embryotoxicity was observed at dosages without observed maternal toxicity, BBP appeared as a specific embryotoxic compound. Most sensitive male toxicity appears at testis atrophy NOAEL=750 mkd and testis weight reduction CED10=576 mkd.

The effects reported in literature were generally reproduced. In addition, in the present study critical adverse effects were observed on maternal serum ALAT levels and on fetal testis weight and descent, both found only after long exposure. End points which are reversible or which are affected in the later phase of exposure appear more likely to be affected after the long exposure regimen. Liver enzyme changes and hematological parameters appear more sensitive in pregnant versus nonpregnant animals. The data set collected in this study will be instrumental for further development and discussion of the alternative benchmark approach for human risk assessment.

Samenvatting

De embryofoetale toxiciteit van butylbenzylftalaat is onderzocht in de rat in een studieopzet met tien doseringsgroepen. Het effect van blootstellingsduur is bestudeerd, en effecten in drachtige en niet-drachtige dieren zijn vergeleken. Bovendien is de klassieke data-analyse met no-observed- effect-levels (NOAEL) vergeleken met een alternatieve benchmark benadering waarin critische-effect-doseringen (CED) worden afgeleid. Daarnaast is een verkennende studie uitgevoerd met jonge mannetjesratten om de testistoxiciteit van de stof te bepalen. Dit rapport beschrijft de resultaten alsmede een initiële analyse van de data. Een uitgebreide data-analyse is voorzien voor een volgende rapportage.

Vergelijking van de laagste NOAEL met de laagste CED laat zien dat in beide analyses levereffecten de maternale toxiciteit bepalen, waarbij de NOAEL voor levergewicht 450 mg/kg lichaamsgewicht/dag (mkd) bedraagt en de CED20 voor serum ALAT 456 mkd. Voor embryotoxiciteit blijken in beide analyses testis-effecten het gevoeligst, met NOAEL voor testisgewicht < 270 mkd en de CED01 voor foetale caudale testismigratie 164 mkd. Aangezien embryotoxiciteit optrad bij doseringen waarbij geen maternale toxiciteit gezien werd, komt BBP uit deze studie naar voren als een specifiek embryotoxische stof. Gevoeligste toxiciteit bij mannetjes werd gezien als testis-atrofie met een NOAEL van 750 mkd en als testisgewicht reductie bij CED10 van 576 mkd.

De uit de literatuur bekende effecten werden in het algemeen gereproduceerd. Daarnaast werden in deze studie critische effecten gezien op maternale serum ALAT spiegels en op testisgewicht en -migratie, beide uitsluitend na lange blootstelling. Eindpunten die reversibel zijn of die gevoelig zijn in de latere fase van de blootstelling zijn veelal gevoeliger voor langere blootstelling. Lever enzym veranderingen en hematologische parameters blijken gevoeliger in drachtige dan in niet-drachtige dieren. De dataset van deze studie zal gebruikt worden voor verdere ontwikkeling en discussie van de alternatieve benchmark benadering voor humane risicoanalyse.

1. Introduction

Reproductive toxicity of chemicals is assessed on the basis of standardized test protocols, laid down in OECD guidelines, and performed according to strategies defined within EU guidelines. Hazard and risk assessment of new and existing chemicals is performed on the basis of the results obtained using these tests. Test protocols are regularly reconsidered to improve their utility for risk assessment. Experimental analysis and validation are essential for a proper evaluation of proposed amendments. In the present study the usefulness of three possible changes within the developmental toxicity study is evaluated experimentally as outlined below. Butyl benzyl phtalate is used as a model compound, also in view of the presently ongoing revision of its risk assessment. In relation to the latter aspect, a small 28-day study with young males was also performed to additionally assess the testicular toxicity of this compound.

Within the developmental toxicity study (OECD 414), exposure is classically done by daily gavage from gestation day 6 to 15 in the rat. Exposure covers the complete period of organogenesis roughly from one day after embryo implantation on gestation day five to closure of the palate on gestation day 15. Closure of the palate is considered as the last gross morphogenetic event in embryogenesis, which marks the onset of the fetal period. Fetal assessment is done on day 21 of gestation, one day prior to birth. A recent change in the developmental toxicity test guideline prescribes exposure from gestation day 6 to day 20 of gestation. Arguments for this change are the possibility of higher sensitivity of the test, together with an unchanged duration and little extra effort. Toxic effects induced in the fetal period can now be assessed within the developmental toxicity study. A possible consequence of the new method is that embryos with malformations may have a smaller chance of surviving the fetal period. This would alter the specific embryotoxic effect observed at necropsy, since conceptuses that would appear in the old protocol as malformed fetuses may now appear as late resorptions or dead fetuses. For most compounds the NOAEL may not change with the new protocol, as the organogenesis period is generally more sensitive to toxic insults than the fetal period. Within this study, both exposure protocols are compared.

The present guidelines for toxicity testing generally prescribe three dose groups and a vehicle control group. Evaluation of results leads to the determination of the NOAEL, which is extrapolated to the human situation using uncertainty factors. An alternative approach employs the complete dose-response curve (Slob and Pieters, 1998). In order to derive a more reliable dose-response relationship, a larger number of dose groups is preferable, with fewer animals per dose group. The total amount of animals used with the new approach can be similar or perhaps even lower than with the classical approach. Starting from a defined critical effect size (CES), the critical effect dose (CED) is derived. In contrast to the NOAEL, the CED is independent of the actual dosages chosen in the study design. A further advantage of this approach is the possibility to calculate the confidence interval around the CED. Both approaches are compared in this study.

The selection of dose levels for developmental toxicity studies is based to a large extent on the results of previous general toxicity studies, which are carried out with adult male animals. Moreover, the maternal toxicity in developmental toxicity studies is studied only superficially, which hampers comparison of effects in pregnant dams versus nonpregnant animals. The sensitivity of parameters may be clearly different between pregnant and nonpregnant animals

as profound physiological changes occur in pregnancy. In this study, extensive pathology of pregnant dams and in parallel groups of nonpregnant animals was performed in order to facilitate such a comparison.

The effects observed in developmental toxicity studies are classically assessed through weight parameters, morphology and histology. The interrelationship of these parameters with physiological and biochemical parameters such as enzyme activities and hormone levels are important for the determination of the CES for each effect parameter. In this study, in addition to the classical parameters several physiological and biochemical parameters are assessed that are pertinent to the effects expected with the model compound.

Butyl benzyl phthalate is chosen as the model compound for several reasons. Firstly, it is a developmental toxicant for which much toxicological information is available. Secondly, the compound was tested recently in this laboratory in a combined reproductive/developmental toxicity study (Piersma et al., 1995). From this study the relative sensitivity of our rat strain could be derived. With the above information the choice of dose levels was facilitated. Thirdly, the risk analysis of this compound especially regarding developmental toxicity was under re-evaluation in view of contaminations of phthalates recently found in canned baby foods and in plastic baby toys.

In view of the risk analysis of phthalates and their fertility effects we performed a pilot 28-day male toxicity study using the same dose groups as in the developmental study, with three young males per dose group. Also for this study a comparison is made for interpretation via NOAEL versus CED.

2. Materials and Methods

2.1 Test compound and dosing

Butyl benzyl phtalate (BBP), CAS 85-68-7, molecular weight 312.37, was obtained from Aldrich as a colorless transparent liquid. Infrared spectrometry carried out at RIVM-LOC revealed no contaminations (not shown). The compound was suspended in Corn Oil (Sigma) and administered daily by gavage. Dosage volume was 5 ml/kg body weight for all dose groups. The most recent body weight determination was taken as a reference. Pregnant animals were weighed on gestation days 0, 6, 11, 16, and 21. Animals in the male toxicity study were weighed on mondays, wednesdays, and fridays during the dosing period. Dosing was carried out in a predefined random order.

2.2 Test animals

Rats of the Harlan Cpb-WU strain were used. Females and males for the developmental toxicity study were 8 weeks of age on arrival and were acclimatized for 2 weeks before entering the study, animals for the male toxicity study were 3 weeks of age on arrival and acclimatized for one week before entering the study. Animals were housed in macrolon type 2 cages, firstly in groups, and after mating females were housed individually. They were supplied with RMH-GS feed in pellets (Hope Farms, Woerden) and tap water ad libitum. They were kept under a changed day/night regime with the light period between 12.00 noon and 24.00 midnight. Temperature was 20-24°C and relative humidity between 50 and 70%.

2.3 Experimental design

The experimental design was fully described in study plan nr. 650040 001 and in animal experiment nr. 199700114.

2.3.1 Developmental toxicity study

Females were mated with males, and dosed with the test compound daily by gavage either from gestation day 6 - 15 (short exposure, groups 1-10) or from gestation day 6 - 20 (long exposure, groups 11-20) (fig.1). Ten dose groups were started for both regimens, with dosages 0, 270, 350, 450, 580, 750, 970, 1250, 1600, or 2100 mg BBP per kg bodyweight per day. The 0, 450, 750, and 1250 mg/kg dose groups contained 25 mated females, representing the classical NOAEL-aimed design. The other groups contained 10 mated females. The complete dose range was to be used for dose-response curve fitting. To each dose group 3 nonmated females were added for comparison with pregnant females. Body weight and food consumption were determined on gestation days 0, 6, 11, 16, and 21. Necropsy was performed on gestation day 21. Before termination, blood was taken under CO_2/O_2 anaesthesia for hematological and biochemical analysis. After termination the weights of liver, kidneys, thymus, thyroid and spleen were determined, and a liver lobe was isolated for biochemical

analysis. The organs were then processed for histology. The uterus was explanted and the fetuses prepared free. Corpora lutea, implantation sites, early and late resorptions and fetuses with early and late death were counted, and external anomalies of living fetuses were noted. Odd fetuses were processed for skeletal staining and examination, even fetuses were fixed for morphological analysis according to Barrow and Taylor (1969).

2.3.2 Male toxicity study

Male rats of four weeks age were treated daily for 28 days by gavage with one of ten dosages as indicated above. Each dose group consisted of three animals. Necropsy was performed one day after the last dosage. Organs mentioned above and in addition the testes, were explanted, weighed and processed for histology.

2.4 Biochemical analyses

Blood cells were differentially counted and standard hematology performed. Serum ALAT (alanine aminotransferase), ASAT (asparagine aminotransferase), testosterone, progesterone, LH, FSH, PRL and estradiol were determined, and liver protein and PCO-activity (palmitoyl CoA oxidase) were assessed.

2.5 Statistical analyses

The data were analyzed in two ways: by comparing the dose groups to the control group using statistical tests of significance (the "NOAEL approach"), and by dose-response modeling (the "Benchmark approach"). In the NOAEL approach the treatment groups were compared to the control group by Dunnett's tests.

In the Benchmark approach a dose-response model is fitted to the data, then a Critical Effect Size (CES) is defined, and the associated Critical Effect Dose (CED) is derived from the fitted model. The uncertainty in the estimate of the CED is assessed by a bootstrap method (Slob and Pieters 1998), resulting in an uncertainty distribution from which any desired confidence interval can be derived. In this report the 5% and 95% confidence limits are not yet reported (i.e. 90% confidence intervals). Note that the 5% confidence limit can be considered as the Benchmark dose as originally defined by Crump (1984).
The choice of the model for deriving the CED follows from a procedure of applying likelihood ratio tests on the members of the following nested family of models:

model 1: $y = a$
model 2: $y = a \exp(b\, x)$
model 3: $y = a \exp(b\, x^d)$
model 4: $y = a\, (c - (c - 1) \exp(b\, x)$
model 5: $y = a\, (c - (c - 1) \exp(b\, x^d)$,

by choosing a more complicated model when the increase in number of parameters leads to a significantly better fit to the dose-response data. In these models the parameter a represents
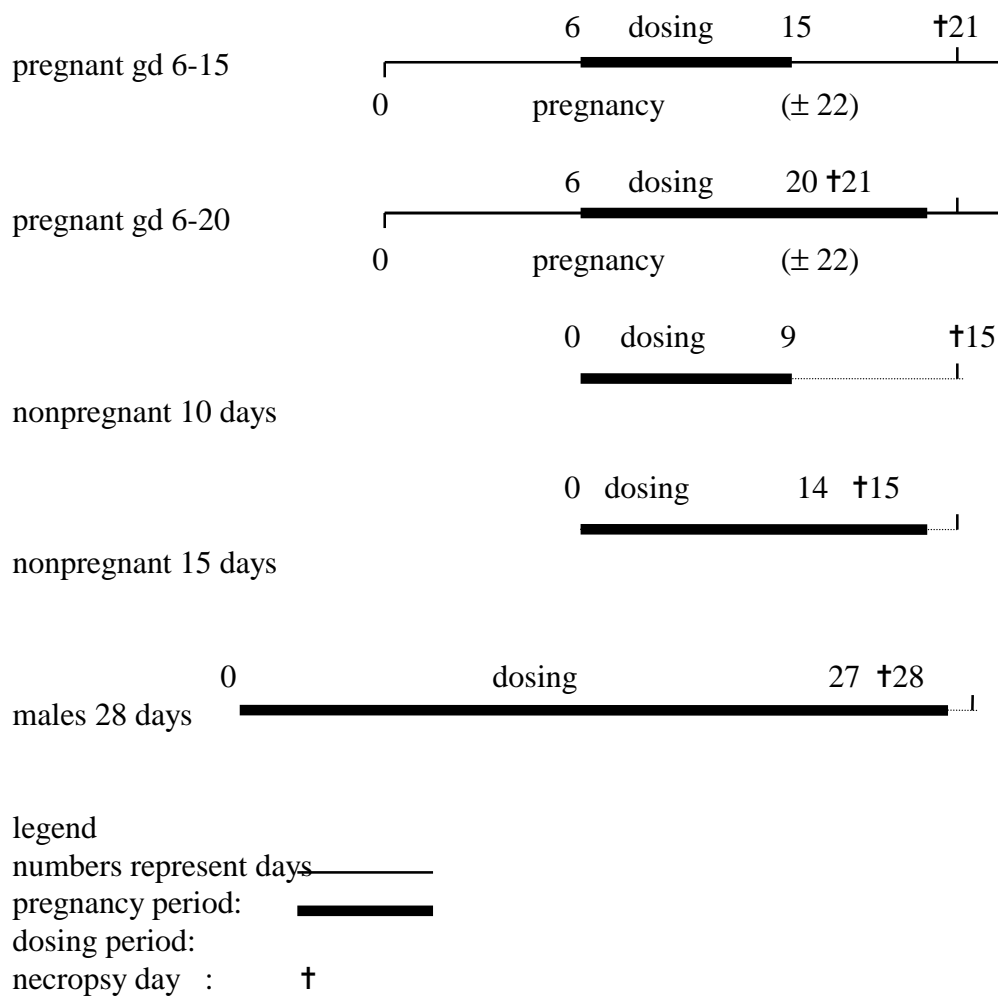
the background response; the parameter *b* reflects the 'slope' or the 'strength' of the response. These models are suitable for describing different (sub)populations by the same model. For example, when males and females are equally sensitive to the compound studied with respect to body weight, male and female body weights can be described by the same model, with only parameter *a* differing between males and females, to account for background body weights differing between sexes. When males and females are not equally sensitive, the parameter *b* should differ between sexes. The hypothesis that parameter *a* or parameter *b* (or both) differ between populations (e.g. sexes, or any experimental treatment) can be statistically tested by a likelihood ratio test.

The above family of models is also used for describing the histopathological data (i.e. ordinal data), by assuming that histopathological scores can be viewed as a discretization of an underlying continuous response, resulting in a limited number of response categories. The bounds between the categories (scores) are estimated by fitting the model tot the data.

For continuous endpoints the CES is defined as a change in average response at any given dose relative to the average response in the controls. For example the CED05 represents the dose associated with a 5% change in average response.

While for continuous data the CED associated with any chosen CES can be assessed, for ordinal data this choice is limited. Only those doses that are associated with the transitions between categories can be estimated. Therefore, in the case of ordinal data a number of CEDs is reported, representing the doses associated with the respective transitions from one response category (score) to the next.

Figure 1 Experimental design: dosing periods and time of necropsy

pregnant gd 6-15

6   dosing   15   †21
0   pregnancy   (± 22)

pregnant gd 6-20

6   dosing   20 †21
0   pregnancy   (± 22)

nonpregnant 10 days

0   dosing   9   †15

nonpregnant 15 days

0   dosing   14 †15

males 28 days

0   dosing   27 †28

legend
numbers represent days
pregnancy period:
dosing period:
necropsy day   :   †

3. Results

3.1 Developmental toxicity study

3.1.1 Maternal parameters

Pregnancy and Survival
Mating efficiency was 88% (Table 1). In the highest two doses (1600 and 2100 mkd) several animals died or were killed in extremis during the first five days of dosing due to toxicity of the compound. Two animals in the 1600 mkd dose group died at dosing day 10. One animal on 580 mkd appeared ill on dosing day 14 and died one day later. The remaining data given are for the surviving animals only.

Food consumption
Decreased food consumption was observed at the three highest doses (1250 mkd and up) during the first 5 days of dosing only (Table 2). Thereafter food consumption was not different between all dose groups.

Body weight
Pregnant rats showed a dose-related reduction in body weight gain at 750 mkd and higher (Table 3). Reduction occurred mainly in the later stages of pregnancy. The effects were more pronounced in the long than in the short exposure groups. Nonpregnant rats did not show dose-related differences in body weight between dose groups.
Dose-response curve fitting for pregnant body weights showed a CED10 of 830 (Figure 1a). However, after subtracting fetal weights at necropsy, CED10 moved to 1540 for long and 2053 mkd for short exposure (Figure 1b), respectively, indicating that the effects on the offspring explain a large part of the maternal body weight effects. For nonpregnant rats, body weight did not show a significant change with dose (Figure 1c).

Liver weight/PCO/ALAT/ASAT/pathology
A dose-related increase in relative liver weight was observed in pregnant animals from 350 mkd after long exposure and from 750 mkd after short exposure (Table 4). The maximal effect was reached earlier after longer exposure. Similar observations were made in nonpregnant animals.
Dose-response curve fitting for pregnant relative liver weights resulted in a CED10 of 516 mkd after both exposure durations (Figure 2a). In nonpregnant animals the comparable CED10 were 381 and 604 mkd for long and short exposure duration, respectively (Figure 2b). PCO activity in the liver showed a dose-related increase from the lowest dose tested (Table 5). In pregnant animals the maximal effect was a fivefold increase after long exposure, whereas after short exposure no more than a 50% increase was observed. In nonpregnant animals a similar pattern was found.
Dose-response curve fitting for pregnant liver PCO activity confirmed these findings with CED20 appearing at 64 and 825 for long and short dosing duration, respectively (Figure 3a). In nonpregnant rats, comparable CED20 amounted to 280 and 1058 respectively (Figure 3b), showing a lower sensitivity than in pregnant rats.

Liver enzyme ALAT and ASAT levels in the blood did not show changes after short exposure (Table 6), with the exception of increases in ASAT at 750 mkd and higher which did not show a clear dose-response relation. After long exposure, a dose-related increase in blood ALAT was found, from 580 mkd onward in pregnant rats and from 1250 mkd onward in nonpregnant rats. Blood ASAT increases after long exposure were found at 750 mkd and higher in pregnant rats, and at 1600 mkd and higher in nonpregnant rats.

Dose-response curve-fitting for blood ALAT after long exposure resulted in a CED20 of 456 mkd for pregnant rats and of 1205 mkd for nonpregnant rats (Figure 4a and b). Serum ASAT after long exposure resulted in CED20 of 756 mkd for pregnant rats and of 1739 mkd for nonpregnant rats (Figure 4d and e).

Histological examination of the liver revealed dose-related extensive peroxisome proliferation (Table 7). Peroxisome proliferation was observed only at the three higher doses after general staining for histopathology, which is not an optimal staining for peroxisome detection. A variable incidence of vacuolisation was observed in all groups. In addition, slight infiltrations of lymphoid aggregates were observed at 1600 and 2100 mkd. The routine histological staining employed was not ideal for determination of peroxisome proliferation.

Kidney weight/pathology
In pregnant animals, relative kidney weight increased dose-dependently from 750 mkd onward, reaching somewhat higher levels after long versus short exposure (Table 4). In nonpregnant animals, control relative kidney weight was around 40% higher than in pregnant animals. In addition, short exposure did not result in observed changes in relative kidney weight in nonpregnant animals. Long exposure resulted in an increase from 970 mkd onward with maximal levels comparable to those in pregnant rats. Kidney histopathology was unremarkable.

Dose-response curve fitting for pregnant relative kidney weight revealed a CED10 of 588 mkd irrespective of exposure duration (figure 5a). For nonpregnant females, CED10 was 1900 mkd for long exposure (Figure 5b), but it should be noted that these data showed a relatively large variation.

Thymus weight/pathology
Thymus relative weight did not change considerably after exposure (Table 4). Histopathology revealed increased medullary cellularity at 1600 mkd and 2100 mkd only in pregnant females (Table 7). In nonpregnant females this effect was observed at 970 mkd and higher.

Thyroid weight/pathology
Thyroid relative weights were dose-dependently increased in pregnant animals, from 750 mkd onward (Table 4). Effects were similar for both exposure durations. In nonpregnant rats no changes were observed. Histopathology of the thyroid was unremarkable.

Spleen weight and pathology
Relative spleen weights showed a dose-dependent increase in pregnant animals at 1250 mkd and higher (Table 4). At the highest doses, relative spleen weight had doubled over controls. In nonpregnant animals, no effects were found.

Histopathology revealed a dose-dependent increase in the extent of extramedullary hemopoiesis in the spleen (Table 7). Pregnant controls already showed extramedullary hemopoiesis, which was further increased after exposure in all dose groups. Nonpregnant control animals had no extramedullary hemopoiesis, and increases were observed after 580 mkd and above.

The results of curve fitting of splenic extramedullary hemopoiesis are shown in Figure 6. The CED's represented in the figures represent the doses associated with the transitions from one category to the next, e.g. from "slight" to "moderate". The differences in control levels and in sensitivity between pregnant and nonpregnant are evident. Moderate hematopoiesis occured in 50% of pregnant animals at 94 mkd after long and at 479 mkd after short exposure (Figure 6a). In nonpregnant animals, the moderate effect level was reached only after long exposure at 1847 mkd (Figure 6b).

Hematology
In pregnant animals at doses 750 mkd and above, hematocrit was increased irrespective of dosing duration (Table 8). Both increased white and red blood cell counts contributed to this increase, and hemoglobin was concomittantly increased. White blood cell increases were evenly distributed over the various subsets of hemopoietic cell types. In nonpregnant females no consistent effects on hematologic parameters were seen.
Curve fitting of hematocrit values showed a CED05 of 756 for both exposure durations in pregnant rats (Figure 7a). In nonpregnant rats the variation in the data interferes with interpretation, although a decrease in hematocrit rather than an increase seems to occur (Figure 7b).

Endocrinology (Table 9)
In pregnant animals after long exposure, progesterone was dose-relatedly reduced from the lowest dose onward. Dose-response curve fitting resulted in CED20 of 707 mkd for long and 967 mkd for short exposure, respectively (figure 8). Estradiol did not show changes, whereas LH was increased at 450 mkd and FSH was increased at 1600 mkd and higher. Nonpregnant animal sex hormone data are not relevant due to oestrus cycle variability.

3.1.2 Foetal parameters

Corpora lutea, implantations and survival
Corpora lutea numbers were comparable in all dose groups (Table 10). Early resorptions were increased in the highest two dose groups only, whereas late resorptions were increased at 750 mkd and higher. Total resorptions were curve fitted and showed a CED01 of 252 mkd (Figure 9a). No differences were observed with exposure duration. Hardly any dead fetuses were found at necropsy. Live fetus numbers decreased from 750 mkd onward, and curve fitting showed a similar but inverse situation as for total resorptions, with CED01 at 394 mkd (Figure 9b).

Fetal weights
Fetal weights were statistically significantly decreased at 450 mkd and higher after both exposure regimens (Table 11). As male fetuses have higher body weights than female fetuses, both sexes were also analysed separately. Curve fitting resulted in CED05 of 534 mkd (Figure 10), irrespective of sex and exposure duration.

Skeletal parameters
At 750 mkd and higher doses skeletal anomalies were observed in increased incidences (Table 12). Reduced rib size and fusion of two ribs and incompletely ossified and fused sternebrae were observed at these doses. The occurrence of a 13th lumbar rib appeared more sensitive to the compound, as it was found in somewhat increased incidence already at 270 mkd and

showed a dose-relatedly increased incidence. Curve fitting revealed CED05 of 259 and 222 mkd for short and long exposure, respectively (Figure 11). There did not appear to be a difference in skeletal parameters between short and long exposure.

Internal organs
Specific soft tissue malformations observed are shown in Table 13. Cleft palate was found in two cases of 750 mkd and one of 1250 mkd, both after long exposure. Anophthalmia was found in several cases after 750 and 970 mkd both after short and long exposure. Two cases of exencephaly occurred in the 750 mkd group after long exposure. A low incidence of fetal ovary malformations was found at 580 mkd and higher, mainly after long exposures. Rather than the spheric shape in unaffected specimens, these ovaries had an elongated shape reminiscent of an earlier developmental stage, suggesting a retardation of morphologic development. The incidence of retarded fetal testicular caudal migration was dose-relatedly increased, the incidence being higher after long as compared to short exposure. Curve fitting revealed CED01 of 275 and 164 mkd for short and long exposure, respectively (Figure 12). Relative fetal testis weight was dose-relatedly reduced after long exposure (Table 14 and Figure 13). Curve fitting revealed a CED05 of 182 mkd (!) for long exposure.

3.2 Male toxicity study

Food consumption
No changes in food consumption were observed after dosing with BBP in this study (Table 15).

Body weight
Body weight gain was reduced at the three highest dosages as measured over the complete dosing period (Table 16).

Liver weight/PCO/ALAT/ASAT/pathology
Average relative liver weights were dose-relatedly increased from 750 mkd onward (Table 4). PCO showed a similar dose-response (Table 5). Blood ALAT showed increases from the lowest dose onward (Table 6) albeit with considerable variation between animals. Curve fitting resulted in CED20 of 1574 mkd for circulating ALAT (Figure 4c). Blood ASAT did showed a statistically significant but small response. Pathological analysis revealed peroxisome proliferation in the liver at 1250 mkd and higher (Table7).

Kidney weight/pathology
Average relative kidney weights showed a dose-related increasing trend from the 580 mkd dose onward (Table 4). Kidney pathology was unremarkable.

Thymus weight/pathology
Average relative thymus weight tended to be decreased in all dose groups (Table 4), and thymus pathology was unremarkable.

Thyroid weight/pathology
Average relative thyroid weight tended to be increased in all dose groups (Table 4), and thyroid pathology was unremarkable.

Spleen weight/pathology
Average relative spleen weight did not show changes (Table 4), and splenic pathology was unremarkable.

Testis weight/pathology
Average relative testis weight was dose-relatedly decreased at 750 mkd and higher, with statistical significance at 1250 mkd and higher (Table 17). Curve fitting resulted in a CED10 of 576 mkd (Figure 14). Testis pathological analysis revealed severe atrophy at 970 mkd and above (Table 7).

Hematology
No effects on hematology were observed in the male toxicity study (Table 8).

Endocrinology LH/FSH/testosterone
LH and FSH were increased at 1250 mkd and higher (Table 9). Testosterone was decreased at 450 mkd and higher. The data did not allow a reliable dose-response analysis, but effects appear to occur below the lowest dose applied (Figure 15).

4. Discussion

The present study was designed to serve multiple aims as outlined in the introduction. This discussion will focus mainly on the developmental toxicity of BBP as analysed in the classical way using all available data. Furthermore, the impact of exposure duration and the effects of pregnancy on adult toxicity will be discussed. Finally, the comparison of the NOAEL approach and the CED approach will be discussed only briefly, leaving room for more detailed analysis and discussion of the present data in subsequent reports.

4.1 Maternal toxicity of BBP

In pregnant animals at the lowest dose tested in the long exposure regimen, circulating progesterone, extramedullary hemopoiesis, and liver PCO level are statistically significantly affected.

Progesterone levels in the circulation are significantly lower than controls at all dose levels. LH increases at 450 mkd and FSH increases at 1600 mkd and higher. At lower doses, hormone level changes do not appear to interfere with the ability to maintain pregnancy, as resorptions occur only at 750 mkd and higher. In addition, short exposure shows effects on progesterone only at 1250 mkd and higher, indicating that the effects at lower doses may be reversible, as short exposure ended 6 days before necropsy. Therefore, although the effects on progesterone level are significant even at the lowest dose, they are not associated with observed effects on pregnancy below the dose of 750.

Extramedullary hemopoiesis is dose-dependently increased at 270 mkd and higher, whereas increases in blood cellularity occur at 750 mkd and higher, and relative spleen weight increases at 1250 mkd and higher. In view of the fact that the rat normally has a variable extent of splenic extramedullary hemopoiesis during pregnancy, these changes are considered to be within physiological limits, and are therefore considered not adverse.

Liver PCO is dose-dependently increased at 270 mkd and higher.The stimulating effect on PCO is accompanied by a dose-dependent liver weight increase which is statistically significant at 580 mkd and higher and by a dose-dependent increasing trend in circulating ALAT and ASAT. PCO is a biomarker for peroxisome proliferation, which is one of the most sensitive effects of phthalates in mammals. Whereas PCO increases are not necessarily adverse per se, the effects on liver weight accompanied by leakage of hepatic enzymes such as ALAT and ASAT into the circulation does indicate liver toxicity. Therefore, at 580 mkd and higher, an adverse effect on the maternal liver is observed in this study.

Other dose-related effects observed are increased relative thyroid weight and increased relative kidney weight at 750 mkd and higher. In addition, body weight gain is reduced at 750 mkd and higher, food consumption is decreased during the first six dosing days at 1250 mkd and higher, and maternal death occurs at 1600 mkd and higher.

Taken together, the LOAEL for maternal toxicity is 580 mkd, based on the dose-dependent liver weight increase.

4.2 Fetotoxicity of BBP

The incidence of an extra lumbar rib is dose-relatedly increased from 270 mkd both after short and long exposure. Extra lumbar ribs are generally not considered as a malformation but as a

variation, which may occur in untreated animals as well. Relative fetal testis weight is decreased only after long exposure, also in a dose-related way from 270 mkd. Testis migration and ovary morphologic development are retarded from 580 mkd onward, especially after long exposure. Fetal weight is decreased at 450 mkd and higher. At 750 mkd, the incidence of resorptions is increased and the number of live fetuses is likewise decreased.

All observations taken together, the most sensitive specific fetotoxic effect is fetal testicular relative weight reduction, with a LOAEL of 270 mkd, a dose at which fetal weight is unaffected. The morphologic effects on testis and ovary at 580 mkd and higher are indicative of retardation of fetal sex organ development. General effects appear dose-dependently as extra lumbar ribs at 270 mkd, with lower fetal weights at 450 mkd and increased numbers of resorptions at 750 mkd. In view of the maternal LOAEL of 580 mkd based on liver toxicity, the effects observed in fetal testis development at 270 mkd and higher are probably specific adverse effects of the test compound. The increased incidence in extra lumbar ribs, albeit a variation, further illustrates that BBP is fetotoxic in the absence of observed maternal toxicity.

4.3 Male toxicity of BBP

In young males, exposed for 28 days starting at 4 weeks of age, trends in relative liver weight and PCO level, thyroid weight, and testosterone level were observed.

Testosterone level, although variable within dose groups, showed a statistically significant decrease at 450 mkd and higher. At 970 mkd and higher, testis weights decreased, histological atrophy was noted, and LH and FSH increased over controls.

Taken together, testis atrophy at a LOAEL of 970 mkd appears as the critical adverse reproductive effect in males in this study, in the presence of toxicity to the liver observed as an increased relative liver weight of around 20%.

4.4 Comparison with effects in published literature

The oral LD50 of BBP in the rat has been reported to be 20.4 g/kg bw (Hammond et al 1987). Substantial body weight effects have been observed at 1500 mkd. Increased liver and kidney weights have been described at around 1000 mkd. Histopathologic effects in the liver have been observed at 960 mkd in the Wistar rat, but not in SD and F334 rats. BBP, like many phthalate esters, has been shown to induce peroxisome proliferation and related PCO-induction in the liver (Barber et al., 1987). Internal haemorrhaging, depressed bone marrow cellularity and reduced thymus weights have been reported in males after subchronic exposure to high doses of BBP (Agarwal et al., 1985)

Oral 14 day studies in the rat showed strain-dependent sensitivity of the testis (NOAEL's 160/480/625 mkd for Sprague Dawley/Wistar/F344 rats, respectively), resulting in reduced weights and histological changes (Agarwal et al 1985, Hammond et al 1987). In a teratogenicity study (Price et al 1990, abstract) with SD rats, exposed via the diet between gestation days 6 and 13, reduced body weight gain and 'fetal variations' were observed at around 1000 mkd. Resorptions and 'malformations' were increased at around 1600 mkd. More specific data were not given. Wistar rats, dietarily exposed between gestation days 0 and 20, showed decreased body weight gain in dams at around 650 mkd, and complete litter resorption at around 1000 mkd (Ema et al 1990) In that study no structural malformations were found.

In the present study, the effects reported in literature were generally reproduced. In addition, in the present study critical adverse effects were observed on fetal testis weight and descent, both found only after long exposure. We have not found these end points studied in published literature.

4.5 Short versus long exposure

Differences as well as similarities were observed in effects after long versus short exposure. Differential effects may be caused by the difference in exposure duration (gd 6-15 versus gd 6-20) resulting in a 1.5 fold difference in the total amount of compound administered. The (partial) development of an effect during the later phase of pregnancy (gd16-20) may also affect the outcome. Furthermore, the moment of necropsy in relation to the exposure period, 6 days after the last dosage in the short exposure regimen, versus 1 day after the last dosage in the long exposure regimen, may play a role for effects that are reversible within days. In addition, in contrast to animals in the long exposure regimen, animals with short exposure were not handled and gavaged with corn oil during gd 16-20.

For several maternal parameters the effects are more pronounced after long exposure. Striking is the effect of exposure duration on PCO and ALAT, which are almost unchanged after short exposure, but dose-relatedly affected from the lowest dose level onward after long exposure. Rather than being unaffected after short exposure, it is likely that these parameters have returned to normal values during the 6 days without exposure before necropsy. Similarly, the differential effect of exposure duration on maternal progesterone seems to imply that reversibility decreases with higher doses.

Similarities and differences in fetal effects with exposure duration may partly be explained by considering the period of sensitivity for each individual parameter. Numbers of resorptions and live fetuses and extra lumbar rib frequency are not different with exposure duration, as these parameters are determined between gd 6 and 15, which are covered in both regimens. Although fetal weight may be affected by exposure in any phase of the prenatal period, the absence of compound-mediated differences with exposure duration in this study imply that the gd 6-15 period is more important for determination of fetal weight on gd 21 than the gd 16-20 period. The differences in foetal weights between control groups indicate that exposure duration influenced fetal weight. Fetal testicular weight and dislocation is clearly different with exposure duration. This can be explained by the fact that growth and intraperitoneal descent of the testis occurs mainly in the later phase of pregnancy, which is covered only by the long exposure period.

As discussed above, the change of exposure period from gd 6-15 to gd 6-20 (and the concomittant change in timing of necropsy relative to exposure) may affect the outcome of several parameters in the developmental toxicity test. In the present study, with a LOAEL of 270 mkd fetal testis weight reduction is the most sensitive fetal adverse effect parameter. This effect, however, only occurs after long exposure. After short exposure, the effects on testicular migration and ovary morphogenesis provide a LOAEL of 580 mkd for adverse specific developmental toxicity. These findings show that the parameter determining the LOAEL as well as the level of the LOAEL may change with the protocol. Furthermore, in the present study the overall sensitivity of the test appears to increase as LOAEL's decrease with longer exposure duration.

4.6 Pregnant versus nonpregnant animals

When comparing maternal effects with effects on nonpregnant animals, differences occur in all parameters measured, and in most cases pregnant animals appear more sensitive than nonpregnant animals. The differences in effects on body weight gain and food consumption are readily explained by the growing conceptuses and the extra need for food consumption in pregnancy, as well as the untoward effects of BBP on growth and survival of the conceptuses. In addition, relative organ weights of liver, kidney, spleen and thyroid are more sensitive to BBP in pregnant versus nonpregnant animals. These effects are superimposed on the effects of pregnancy itself, which results in decreased relative organ weights for all these organs except the liver (table 4). Rather than considering organ weights relative to total pregnant body weight, as is commonly done, organ weights relative to corrected maternal body weight (pregnant body weight minus the weight of uterus and its contents) would be more informative when comparing organ weight effects between pregnant and nonpregnant groups. However, uterine and placental weights were not recorded in the present study. Whereas liver PCO readily responds at the lowest dose irrespective of pregnancy, blood ALAT and ASAT increasing trends appear at lower doses in pregnant versus nonpregnant animals. Therefore, the liver appears to be more sensitive to BBP in pregnant animals.
Most remarkable are the differential effects on hematological parameters. Splenic extramedullary hematopoiesis does not occur in normal adult rats, but does occur in pregnancy in the rat as a physiologic response to the increased need for hematopoiesis. In pregnant rats, hematopoiesis is stimulated by BBP at all dose levels. This results in higher blood cellularities only at 750 mkd and higher. The observation of increased extramedullary hematopoiesis at low doses in pregnant animals in the absence of an increasing effect on blood cellularity implies that BBP affects blood cell survival. The effect on the spleen can thus be explained as a compensatory response to repair hematologic damage. In nonpregnant animals, blood cellularity decreases at 1250 mkd and higher, whereas splenic extramedullary hematopoiesis shows substantial increases at these high doses, which is likely a compensatory response.

4.7 NOAEL versus CED

In the above discussion sections, the classical approach was used for determination of the NOAEL in the most sensitive end point as the basis for the final conclusions on the effectiveness of the compound tested. This approach has several disadvantages. The NOAEL is dependent on preselected dose levels and does not take into account the full dose-response curve. Furthermore, often statistical significance of the difference of the effect as compared to controls is taken as the determinant of the presence of a real effect. In that case the variability of parameters and the study design (dose group size) can be of great influence on the outcome, whereas the assessment of compound toxicity per sé is the aim of the study.
In an alternative approach a dose-response curve on each effect parameter can be derived and used as a basis for further analysis (Slob and Pieters, 1998). For each parameter a critical effect size (CES) is predefined, which indicates the lowest response level at which an effect should be considered adverse. The critical effect dose (CED) is defined as the dose at which the CES occurs. The CED is used in further risk assessment, instead of the NOAEL in the classical approach. The CED is independent of preselected dose levels. The CED is

determined by all data and more importantly, the (point) estimate of the CED is not affected by the variability of the data or the number of animals used.

The fundamental toxicological problem encountered with this approach is the predetermination of the CES for each parameter. As discussed above, the CES for each parameter should not be derived using classical statistical information of an experiment, but should be based upon physiological and toxicological knowledge, and defined independent of experimental outcome. An attempt to define CES for the parameters in this study is given below. For effects on body weight, a 10% effect is usually taken as the cutoff for adversity, although more extreme effects may still be well within physiological limits. As a starting point for further discussion a CES of 10% may be taken as a default CES for all weight parameters. For enzyme and hormone levels, considerable physiological variation occurs within individuals. Liver enzymes may vary considerably, and may therefore require a CES of 20% or higher. Similarly, progesterone may be 20% decreased without compromising the continuation of pregnancy. CES for testosterone, LH and FSH are probably also high in view of the substantial variation in controls. Hematologic parameters did not show adverse effects in the present study, and considerable variation may occur within physiological limits. For developmental parameters such as fetal weight and skeletal variations, which indicate general embryotoxic effects which do not compromise normal functionality, a CES of 5% may be acceptable. Specific parameters however, which may affect the viability of the conceptus, such as morphological malformations (e.g. of testis and ovary in this study) and embryonic death (resorptions), would require a CES of less than 5%, and 1% could be taken as a default, although for such severe effects an even lower setting may be required. As stated above, the CES levels proposed here are given as a starting point for further discussion. In addition, the adversity of changes in parameter values may have to be discussed in the context of changes in related parameters, such as can be done in the present study for related hepatic parameters, various hematologic parameters, or reproductive parameters.

Based on the above CES settings, the lowest CED for maternal toxicity are PCO (CED20=64 mkd), extramedullary hematopoiesis (CED50 for intermediate=94 mkd), ALAT (CED20=456 mkd), relative liver weight (CED10=516 mkd and relative kidney weight (CED10=588 mkd). As PCO is less relevant for the human situation, and substantial extramedullary hematopoiesis is normal in pregnant rats, the CES levels chosen for these end points are probably too low. As a consequence, in view of risk analysis the ALAT CED20=456 mkd should be regarded as the relevant CED for maternal toxicity.

For specific embryotoxicity, fetal testicular caudal migration (CED01=164 mkd) and relative fetal testicular weight (CED05=182 mkd) are the most sensitive parameters, followed by 13th lumbar rib incidence (CED05= 222 mkd), incidence of resorptions (CED01=252 mkd) and live fetus numbers (CED01=394 mkd). For fetal weight CED05 = 534 mkd. Therefore, specific embryotoxicity occurs at 164 mkd, a dose well below the critical dose for maternal toxicity. On these grounds, in the present study BBP appears as a specific embryotoxic compound.

For male toxicity the testis weight reduction gives the lowest CED of CED10=576mkd. In view of its large variability, testosterone level changes are not taken into account here.

Comparison of the lowest NOAEL with lowest CED shows that in both analyses liver effects determine maternal toxicity, with liver weight NOAEL=450 mkd and ALAT CED20=456 mkd. For embryotoxicity, in both analyses fetal testicular effects appear critical, with fetal testicular weight NOAEL<270 mkd and fetal testicular caudal migration CED01=164 mkd. Male toxicity appears at testis atrophy NOAEL=750 mkd and testis weight reduction CED10=576 mkd. It is noteworthy that in both analyses the critical effect parameter is found in the same organ in all cases. In addition, the general conclusions in terms of the specific

embryotoxicity of BBP are the same with both approaches. However, a strict comparison of NOAEL versus CED approach requires the use of 4 doses at 20 animals versus 10 doses at 8 animals, which has yet to be performed with the present data set. Furthermore, the interpretation of these findings in terms of extrapolation to the human situation and risk asessment is subject to further analysis.

Acknowledgments

References

Agarwal DK, Maronpot RR, Lamb JC, Kluwe WM. Adverse effects of butyl benzyl phthalate on the reproductive and hematopoietic systems of male rats. Toxicology 35: 189-206, 1985.

Barber ED, Astill BD, Moran EJ, Schneider BF, Gray TJB, Lake BG, Evans JG. Peroxisome induction studies on seven phthalate esters. Toxicol. Industrial Health 3: 7-24, 1987.

Barrow MV, Taylor J. A rapid method for detecting malformations in rat fetuses. J. Morphol. 127: 291-306, 1969.

Ema M, Murai T, Itami T, Kawasaki H. Evaluation of the teratogenic potential of the plasticizer butyl benzyl phthalate in rats. J. Appl. Toxicol. 10: 339-343,1990.

Hammond BG, Levinskas GJ, Robinson EC, Johannsen FR. A review of the subchronic toxicity of butyl benzyl phthalate. Toxicol. Industrial Health 3:79-98,1987.

Piersma AH, Verhoef A, Dortant PM. Evaluation of the OECD 421 reproductive toxicity screening test protocol using butyl benzyl phthalate. Toxicology 99: 191-197, 1995.

Price CJ, Field EA, Marr MC, Myers CB, Morissey RE. Developmental toxicity of butyl benzyl phthalate (BBP) in mice. Teratology 41: 586 (abs),1990.

Slob W and Pieters MN. A probabilistic approach for deriving acceptable human intake limits and human health risks from toxicological studies: general framework. Risk analysis 18: 787-798, 1998.

Apendix 1   Tables

Table 1. Pregnancy and Survival

dose (mkd)	females gd 6-15					females gd 6-20				
	no. mated	no. pregnant	no. dosing ended	no. died	no. not pregnant	no. mated	no. pregnant	no. dosing ended	no. died	no. not pregnant
0	25	24	0	0	4	25	24	0	0	4
270	10	9	0	0	4	10	10	0	0	3
350	25	21	0	0	7	25	24	0	0	4
450	10	10	0	0	3	10	10	0	0	3
580	10	8	0	0	5	10	9	0	1	3
750	25	23	0	0	5	25	23	0	0	5
970	10	9	0	0	4	10	10	0	0	3
1250	10	8	0	0	5	10	10	0	0	3
1600	25	20	1	2	5	25	20	3	1	4
2100	10	6	4	1	2	10	4	5	3	1

Appendix 2   Figures

figure legends
- one curve plots (fig 1cd, 2c, 3c, 4cdf, 5c, 7ac, 9ab, 14, 15)
  small dots: individual data points
  circles: mean of all data per dose
- two curve plots (fig 1ab, 2ab, 3ab, 4abe, 5ab, 7b, 8, 11, 12, 13)
  small dots: individual data points
  triangles: mean per dose after long exposure
  circles: mean per dose after short exposure
- four curve plots (fig 10)
  circles: short exposure males
  triangles: short exposure females
  plusses: long exposure males
  exes: long exposure females
- cathegoric plots (fig 6ab)
  curves represent transitions from each category to the following.

figures
1a   Pregnant body weight
1b   Corrected pregnant body weight
1c   Nonpregnant body weight
1d   Male body weight
2a   Relative liver weight pregnant females
2b   Relative liver weight nonpregnant females
2c   Relative liver weight males
3a   PCO activity pregnant females
3b   PCO activity nonpregnant females
3c   PCO activity males
4a   serum ALAT pregnant females
4b   serum ALAT nonpregnant females
4c   serum ALAT males
4d   serum ASAT pregnant females
4e   serum ASAT nonpregnant females
4f   serum ASAT males
5a   Relative kidney weight pregnant females
5b   Relative kidney weight nonpregnant females
5c   Relative kidney weight males
6a   Extramedullary splenic hematopoiesis pregnant females
6b   Extramedullary splenic hematopoiesis nonpregnant females
7a   Hematocrit pregnant females
7b   Hematocrit nonpregnant females
7c   Hematocrit males
8   Serum progesterone pregnant females
9a   Resorptions
9b   Live fetuses
10   Fetal weight
11   Extra lumbar rib 13

12   Testicular dislocation
13   Relative fetal testis weight
14   Relative adult testis weights
15   Serum testosterone