AIMS For a better birth...

Ultrasound - weighing the propaganda against the facts

By Beverley Lawrence Beech

AIMS Journal, Winter 1999, Vol 11 No 4

An abridged version appears in this Journal, this is the full paper, which was first published in Midwifery Today, Autumn 1999, Number 51.

Women believe that they can ensure the well-being of their babies by having ultrasound scans and that early detection of a problem is beneficial for these babies. That is not necessarily so, and there are a number of studies which show that early detection of some problems with ultrasound might itself be harmful.

In response to women's desire for information about the implications of routine ultrasound examinations, Jean Robinson and I wrote Ultrasound Unsound? in which we reviewed the research evidence and drew attention to some of the hazards (Beech and Robinson, 1996). Since then more evidence has accumulated.

Miscarriage and perinatal death

It is ironic that women who have had previous miscarriages, or who have conceived after infertility treatment, often have additional ultrasound examinations in order to "reassure" them that their baby is developing properly. Would they consent if they knew of possible increased risks of miscarriage or premature labour or birth?

Obstetricians in Michigan (Lorenz et al, 1990) studied 57 women who were at risk of giving birth prematurely. Half were given a weekly ultrasound examination, the rest simply had pelvic examinations. Preterm labour was more than doubled in the ultrasound group - 52% - compared with 25% in the controls. Although the numbers were small the difference was statistically significant and unlikely to have emerged by chance.

In a large randomised controlled trial from Helsinki (Saari-Kemppainen et al, 1990) which randomly divided over 9,000 women into a group who were scanned at 16-20 weeks, compared with those who were not, there were 20 miscarriages after 16-20 weeks in the screened group and none in the controls.

A later study in London (Davies et al, 1993) randomised 2,475 women to routine Doppler ultrasound examination of the umbilical and uterine arteries at 19-22 weeks and 32 weeks compared with women who received standard care without Doppler ultrasound. There were 16 perinatal deaths of normally formed infants in the Doppler group compared with 4 in the standard care group.

We cannot say whether these studies represent greater risks for everyone, since ultrasound exposure varies so much from one machine to another.

It is not, however, only antenatal clinic patients who may be at risk. Physiotherapists use ultrasound to treat a number of conditions. A study done in Helsinki (Taskinen et al, 1990) found that if the physiotherapist was pregnant, handling ultrasound equipment for at least 20 hours a week significantly increased her risk of spontaneous abortion. Also, the risk of spontaneous abortions occurring after the tenth week was significantly increased for deep heat therapies given for more than 5 hours a week and ultrasound more than 10 hours a week.

Diagnosis of placenta praevia

The Saari-Kemppainen study also revealed the lack of value of early diagnosis of placenta praevia. This is a condition where the placenta is low in the womb and obstructs the opening into the cervix. Of the 4,000 women who were scanned at 16-20 weeks, 250 were diagnosed as having placenta praevia. Because the womb stretches during pregnancy, when it came to delivery there were only four who actually had the condition. Interestingly, in the unscanned group there were also four women found at delivery to have placenta praevia. All the women were given caesarean sections and there was no difference in outcomes between the babies. Indeed, there are no studies which demonstrate that early detection of placenta praevia improves the outcome for either the mother or the baby.

The researchers did not investigate the possible effects on the 246 women who presumably spent their pregnancies unnecessarily worrying about having to have a caesarean section and the possibility of a sudden haemorrhage. We now know that stress alone may restrict the baby's growth in the womb.

Since the publication of *Ultrasound? Unsound* further studies have raised questions about the value of routine ultrasound scanning.

Babies with serious defects

The only way in which mass antenatal ultrasound screening has been proved to reduce perinatal mortality statistics is by detection of serious problems at 18-20 weeks and mothers choosing termination. Babies which would have died anyway from lethal conditions die sooner and do not appear in birth statistics. The perinatal mortality rate is reduced but babies' lives are not saved.

Almost all babies receive a dose of ultrasound, but even at the best centres, wide variations occur in detection rates for babies with major abnormalities and even in the best centres, antenatal detection

rates for heart problems are not high. Both local and international detection rates differ widely in published studies (which are usually undertaken in centres of excellence), but the majority of mothers will be exposed to older machines in ordinary hospitals and clinics.

Low detection rates (either from poor equipment or unskilled operators) means all the babies get the ultrasound dose but few of them get the 'benefits' of accurate diagnosis. The skill of the operators will vary (everybody has to learn sometime) but even with the best machines and the best operators misdiagnoses occur. AIMS has had a number of problems with inaccurate diagnosis from portable hand held scanners used on the ward - usually by inexperienced junior doctors.

A study from Oslo (Skari H et al, 1998) looked at how many babies born with serious defects had been diagnosed by antenatal scans, and whether the early diagnosis made any difference to the outcomes. Women in Norway have a scan at 17-21 weeks done by trained midwives, who refer to obstetricians if an abnormality is suspected.

Knowing about the problem in advance did not benefit these babies; more of them died. They got delivered sooner, when they were smaller, a choice that could have long-term effects. All 12 babies with abdominal wall defects survived. But for the 6 detected on the scan, their length of hospital stay was longer and they spent longer on ventilators, though the numbers are too small to be significant. They were operated on sooner (4 hours rather than 13 hours) but the short term outcomes were the same. The survival rate in the diagnosed group was 77% and 96% in those who were not detected until after birth.

As well as false negatives (undetected abnormalities) there are also false positives (mistaken diagnosis of abnormalities which don't exist). Such screening carries the risk of termination of a normal pregnancy. Some of these are reported in the literature but we know of other cases.

In 19 months, 36 babies were referred from a population of 2.5 million. They had diaphragmatic hernias, abdominal wall defects, bladder extrophy (the bladder protrudes through the abdominal wall) or meningomyelocele (a defect in the spinal column where the tissue protrudes through it). Only 13 of the 36 defects had been detected before birth (36%). They found that only 2 of 8 congenital diaphragmatic hernias were picked up on ultrasound, half the cases of abdominal wall defects (6 out of 12), 38% of the meningomyelocele (5 out of 13) and none of the three cases of bladder extroversion. The mothers had an average of 5 scans (from 1 to 14); but those where an abnormality was detected had an average of 7.

3 out of the 13 babies diagnosed antenatally died. There was 1 death in the 23 undiagnosed. All 13 babies with antenatal diagnosis were delivered by caesarean, and in all but 3 of these babies the caesarean was elective. 19 of the 23 undiagnosed babies had an uncomplicated vaginal delivery. The diagnosed babies had lower birth weight and two weeks shorter gestation. Although the babies with pre-diagnosed abdominal wall defects got surgery more quickly (4 hours versus 13 hours) the outcomes were the same in both groups.

Although small, this is an important study because pregnant women often automatically assume that antenatal detection of serious problems in the baby means that lives will be saved or illness reduced.

Growth Retarded Babies

One of the promises held out by antenatal scanning is that obstetricians will be able to identify the baby with problems and do something to help it and great attention is paid to whether the baby is growing normally.. A German study from Wiesbaden hospital (Jahn A et al, 1998) found that out of 2,378 pregnancies only 58 of 183 growth retarded babies were diagnosed before birth. Forty-five fetuses were wrongly diagnosed as being growth retarded when they were not. Only 28 of the 72 severely growth-retarded babies were detected before birth despite the mothers having an average of 4.7 scans.

The babies diagnosed as small were much more likely to be delivered by elective caesarean - 44.3% compared with 17.4% for babies who were not small for dates. If the baby actually had intrauterine growth retardation (IUGR) the section rate varied hugely according to whether it was diagnosed before birth (74.1%) or not (30.4%).

So what difference did diagnosis make to the outcome for the baby? Pre-term elective delivery was 5 times more frequent in those whose IUGR was diagnosed before birth than those who were not. 77% of these were the result of medical interventions because of suspected fetal distress and not related to premature labour or rupture of membranes. The average diagnosed pregnancy was 2-3 weeks shorter than the undiagnosed one. The admission rate to intensive care was 3 times higher for the diagnosed babies.

The long-term emotional impact

The effects of screening on both parents can be profound. For example, women waiting for the results of tests may try not to love the baby in case they have to part with it. The medical literature has little to say about the human costs of misdiagnosis unless the baby was mistakenly aborted , and even then it tends to focus on legal action. However, a letter in the British Medical Journal revealed how a diagnosis of a minor anomaly can have serious long-term implications for the family:

A woman was referred for amniocentesis during her second pregnancy on the grounds of maternal age (35 years) and anxiety, she was accompanied by her husband. Their normal three year old son played happily during the consultation. The husband confided to the doctor that they had opted for amniocentesis to avoid having another 'brain damaged' child. On questioning it became apparent that an ultrasound examination before their son's birth had shown a choroid plexus cyst. Despite having a healthy child, the husband remained convinced that this cyst had caused his son to be disabled. (Mason G and Baillie C, 1997). This is one of many cases where an abnormal test - even if it turns out to be wrong - can do long term damage to relationships between parent and child (Robinson J, 1997)

As Theresa Marteau, an expert in psychological problems in screening, warned 10 years ago: "Unless the

emotional and behavioural consequences of such screening programmes are properly monitored, we will not know their true costs and hence net benefits. We will be uncertain of their value for many and even of whether they should be provided at all" (Marteau T, 1989).

Evaluating the risks

When ultrasound was first developed researchers suggested that "the possibility of hazard should be kept under constant review" (Donald I, 1980) and they said that it would never be used in early pregnancy under three months, this initial caution was ignored. As soon as vaginal probe ultrasound was developed, which could get good pictures in early pregnancies (and get nearer to the baby giving it a bigger dose) it became widely used for dating pregnancy and is claimed to be more accurate. However, a recent article from Denmark (Olsen O and Clausen J, 1997) challenged the widely held view that scan dating of pregnancies is more accurate than counting days from the last menstrual cycle.

The problem is, the authors say, that the time-honoured method has always estimated length of pregnancy as 280 days from last period, whereas a study of over 400,000 pregnancies where dates were sure, showed that the actual length was most likely to be 283 days. So this method often led to an under-estimate. Scanning, on the other hand, was more likely to overestimate pregnancy length by two days. The most accurate assessment would be made by either adding 3 days to the old calculations of time since the last period, or taking 2 days off the estimated scan date. As a result of this study it will no longer be possible to claim that an ultrasound scan is more accurate than using a calendar and the date of your last period.

AIMS has had a number of complaints from women who were certain of their dates but doctors insisted they were wrong because the scan must be right. They believed the technology rather than the woman - sometimes with tragic results.

Research by Liebeskind revealed "the persistence of abnormal behaviour ... in cells exposed to a single dose diagnostic ultrasound ten generations after insonation". She concluded, "If germ cells were ... involved, the effects might not become apparent until the next generation" (Liebeskind, 1979). When asked what problems should be looked for in human studies, she suggested: "Subtle ones. I'd look for possible behavioural changes, in reflexes, IQ, attention span (Bolsen B, 1982).

Because ultrasound has been developed rapidly without proper evaluation of its risks and benefits it is now extremely difficult to prove that ultrasound exposure causes subtle effects or indeed any effects, because we simply do not have a sizeable population of pregnant women whose babies did not get ultrasound for comparison. (After all, it took over ten years to prove that the gross abnormalities found in some newborn babies were caused by thalidomide). However, there are a number of ultrasound studies which raise serious questions that still have to be addressed.

The first evidence we saw of possible damage to humans came in 1984 when American obstetricians publised a follow-up study of children aged 7 to 12 years born in three different hospitals in Florida and

Denver, who had been exposed to ultrasound in the womb (Stark et al, 1984). Compared with a control group of children who had not been exposed they were more likely to have dyslexia and to have been admitted to hospital during their childhood, but no other differences were found. Because this was not a randomised study, it was too easily dismissed. It did not prove there was a problem but gave early warning that there might be.

In 1993 a study in Calgary, Alberta, which examined the antenatal records of 72 children with delayed speech of unknown cause were compared with those of 142 controls who were similar in sex, date of birth and birth order within the family. The children were similar in social class, birthweight and length of pregnancy. The children with speech problems were twice as likely as controls to have been exposed to ulrasound in the womb. 61% of cases, and only 37% of controls, had had at least one exposure. Once again, because it was not a randomized study but was looking back at children after the event its results were not accepted as proved and were dismised. That does not mean, however, that the results were not valid.

AIMS pressed the Medical Research Council to do a randomised controlled trial without success but the Norwegians, to their credit, did a large scale prospective trial where half the women had two low dose ultrasound scans and half did not.

A Norwegian study (Salvesen K, 1993) showed an increase in left handedness, but no increase in dyslexia. While the increase in left handedness was not large, it was significant and it does suggest that ultrasound may have an effect on the development of the brain. It should be noted, however, that the scanners used in this study emitted very low doses of ultrasound - lower than exposures from many machines nowadays - the women had only two exposures and it was real time, not Doppler, a more powerful form of ultrasound.

Assessing the risks

"Present day ultrasonic diagnostic machines use such small levels of energy that they would appear to be safe, but the possibility must never be lost sight of that there may be safety threshold levels possibly different for different tissues, and that with the development of more powerful and sophisticated apparatus these may yet be transgressed" (Donald I, 1979).

Donald's foresight was remarkable. The machines in use today are far more powerful than the machines used a decade or more ago, and new variants are being developed all the time. The number of routine scans for even a normal low-risk pregnancy are increasing, and the duration of each scan appears to be longer than ten years ago. Yet, there has been inadequate research into the potential long-term effects of these developments.

Measuring the impact of any intervention in pregnancy is very complicated because there are so many things to look at. The child's intelligence, personality, growth, sight, hearing, susceptibility to infection, allergies and subsequent fertility, are but a few factors which, if affected, could have serious long-term

implications, quite apart from emotional effects on families when babies have a false positive or false negative diagnosis of some sort of abnormality.

Because a baby grows rapidly, exposing it to ultrasound in utero at 8 weeks can have different effects to exposure at for example, 10, 18 or 24 weeks (this is one of the reasons the potential effects of exposure are so difficult to study). Undertaking an ultrasound examination at 8 weeks can have different implications to one undertaken at 24 weeks.

Women are now exposed to so many different types of ultrasound: Doppler scans, real-time imaging, triple scans, external fetal heart-rate monitors, hand held fetal monitors. Unlike drugs, whereby every new drug is required to be tested and shown to have benefits that outweigh any possible harmful adversed effects, the rapid development of each new variation of ultrasound machine has not been accompanied by similar careful evaluation by controlled large scale trials prior to being let loose on the unsuspecting public.

Despite decades of ultrasonic investigation, no-one can demonstrate whether ultrasound exposure has a particular adverse effect at a particular gestation, whether any ill-effects of ultrasound might be cumulative, or whether they might relate more directly to the output of a particular machine or the length of the examination. How many exposures are too many? What is the mechanism by which fetal growth may be affected? A large randomised study (Newnham et al, 1993) of women exposed to five Doppler examinations during their pregnancies (compared with a control group which had single ultrasound imaging at 18 weeks) showed decreased birthweight, although a later study showed the children caught up later.

There is, in effect, a presumption of safety in the use of this technology (rather as there was for 20-30 years when pregnant women were told that x-rays were absolutely safe, subsequent research revealed an increased rate of cancer in the exposed children). The published studies raise serious concerns which are widely ignored and more and more women and babies are exposed to ultrasound routinely and often for the most trivial of reasons. No-one asks the question "We have the technology, we can do it, but should we do it? or "Should we introduce it more slowly and in a way which allows us to monitor any possible adverse effects?"

It should not be forgotten that numerous studies on rats, mice and monkeys over the years have found reduced fetal weight in babies that had ultrasound in the womb compared with controls. Nor should it be forgotten that in the monkey studies quite striking behavioural differences were noticed in the babies after birth (Tarantal et al, 1993). The ultrasound babies sat or lay around the bottom of the cage, whereas the little control monkeys were up to the usual monkey tricks. Long term follow up of the monkeys has not been reported. Do they reproduce as successfully as the controls? And, let us also bear in mind, as Jean Robinson has noted: "Monkeys do not learn to read, write, multiply, sing opera, or play the violin". Human children do, and perhaps we should consider seriously whether the increases in children with dyslexia, learning difficulties and behavioural problems are related to ultrasound exposure in the womb.

Furthermore, when a woman is scanned her baby's ovaries and life-time supply of eggs are also scanned. So if the woman had seven scans during her pregnancy, when her daughter eventually turns up years later at the antenatal clinic, her developing baby will already have had seven scans.

Since the introduction of ultrasound this technology has developed rapidly. Because of the wide variation in machines and the different types of ultrasound risks or claims of safety from different studies, at different times and different places are not comparable. It is not like giving a standard dose of a drug in different trials. Despite almost thirty years of vigorous ultrasound promotion very few follow-up studies have been done on the children who were exposed in the womb and the maximum age is nine years. By the time that research is available the old equipment to which those children were exposed are no longer in use and have been replaced by machines which are much more powerful.

Unfortunately an unexposed population no longer exists throughout the developed world. So we have no point for comparison.

Do women really know what they consent to when they rush to hospital to have their first routine ultrasound scan, then trustingly agree to yet more? Before consenting to any scan the woman should be told what is the purpose of this particular scan and we suggest a copy of the AIMS consent form should be completed and signed every time an ultrasound examination is undertaken.

Note: Copy of Ultrasound? Unsound which is no longer in print can be found here