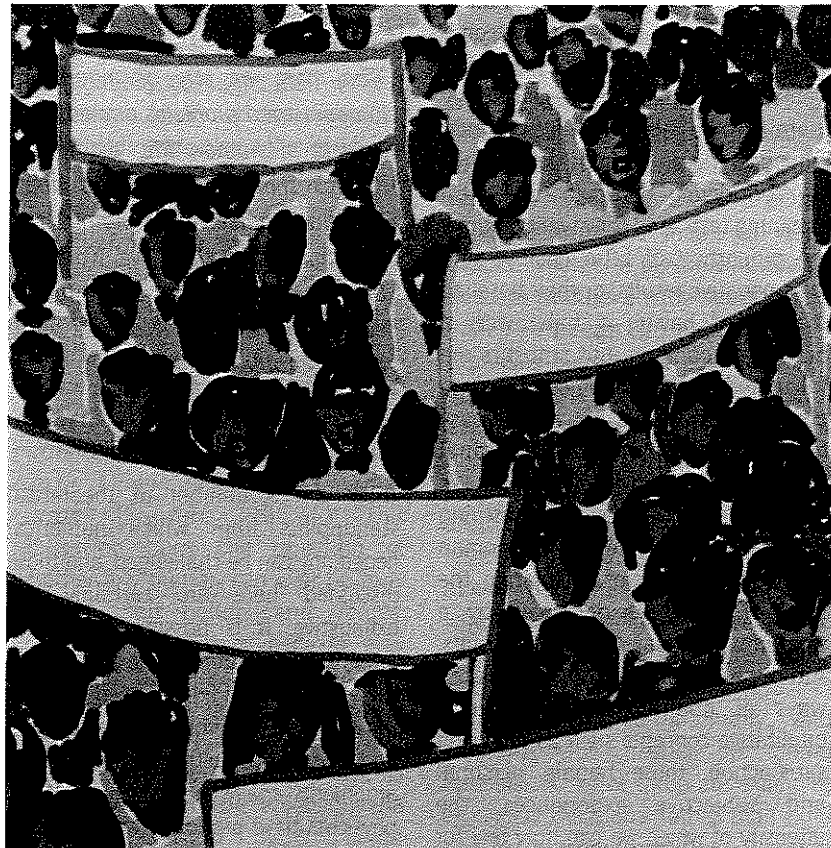


## Seeing women in the numbers



© BSI/ Eand/ Science Photo Library

This article explores some of the issues around the dominance of mathematical, statistical and scientific ways of knowing in Western society and the impact this kind of evidence has on women's choices and midwifery practice. It is suggested that many of the decisions women need to make cannot be informed by statistical and pseudo-scientific evidence, and that other ways of knowing are essential in midwifery practice. The example of intention-to-treat analysis is examined as an illustration of how even the results generated by large-scale clinical trials may not always be as useful to women and midwives as one might assume from the emphasis which is placed on these as the 'gold standard' of knowledge. Some of the different ways in which numerical and scientific hegemony can be challenged are considered, along with the issue of the contextualisation of different forms of knowledge.

## Introduction

Love them or hate them, numbers surround us. And most of us have been taught to believe they hold a greater degree of truth than our feelings and emotions, whether we like a particular piece of art or what our intuition tells us. But to what extent are numbers – and the mathematical and scientific approaches that they serve – helpful to midwives who need to find evidence which can assist women to make choices? In a world dominated by science, research and numerical evidence, it can sometimes be difficult to make sense of numbers in ways that are useful for the women with whom we work. Is there a need to consider non-numerical forms of midwifery knowledge alongside the statistical and numerical ways of knowing which

underpin the majority of policies and guidelines dictating midwifery practice today? If so, from what power and knowledge bases can we challenge the current statistical hegemony in order to make this a mainstream possibility? This article will consider these issues, and explore the concept of intention-to-treat analysis in order to illustrate some of the problems which can be caused by an over-reliance on numbers without consideration of the broader context of knowledge.

Given that I am a strong believer in the desirability of stating one's own bias rather than expecting the recipients of any information offered to figure this out for themselves, I would like to make two things clear at the outset of this article. Firstly, I am not opposed to the use of maths and

statistics, either in research or everyday life; my view is that numbers – and what we can learn from them – can be fascinating. On the other hand, I am concerned that, by placing an undue value or status on the numerical, we are missing out on so much that we could learn and know about the world in other ways.

I would argue that there are questions which can be answered more effectively by using quantitative research and statistical approaches than any other method. If we discover a new drug which we think will treat a certain pathological condition, there is no better test of this at the current time than to conduct a randomised controlled trial. It can sometimes be incredibly useful to be able to offer women a statistical estimation of the chances that action A will lead to

outcome B. Where the question is 'black and white', then we can usually (although not always) find some research findings which are pretty much 'black and white' to help inform a woman's choices, with the inevitable caveat that we should carefully appraise the research to determine whether the findings are reliable and/or relevant to the situation in question. Yet very few issues are this cut and dried, and there exists a massive grey area, where the knowledge women need does not relate to such specific options. When faced with the day-to-day dilemmas of midwifery practice, we often have no quantitative research evidence to which we can turn for the answers.

### Exploring the non-numerical

*"As long as our so-called thinkers take the unwarranted position that the road to truth must be through certain approved and pre-determined channels and no other, many vital truths will be missed and ignored by men (sic) who are supposed to be leaders in their fields of knowledge".* (Shelton 1995:p6)

It is not difficult to find examples of the need for non-numerical knowledge in midwifery. Imagine a midwife is working with a woman who is seeking an active labour and physiological birth in a birth centre or midwifery-led unit, and who would prefer not to have pharmaceutical pain relief unless this is absolutely necessary. She might ask her midwife about non-pharmacological ways of coping with the sensations of labour, an area that is relatively abundant in quantitative data about the effectiveness of different tools to help women cope with labour. Yet no numbers can tell this woman exactly what her labour will be like, even if she has experienced labour before, or what ways of coping will be best for her at the time. Her midwife may well need to draw upon her own experience and intuition in finding ways to help this woman once she is in labour, and the woman's own body knowledge may

well help her in choosing positions which ease her labour and help her achieve a successful birth.

Equally, there may be plenty of audits which tell us what percentage of women have a forceps delivery in a given unit, but statistical analysis is of little use in analysing the complex personal, political, systemic, and other factors which may render this outcome more or less likely for a particular woman. Again, it can be incredibly useful to be able to give women numerical data, perhaps when they are considering particular antenatal screening tests, and they want to know the proportion of false positive and false negative results, or their baseline chances of having a baby with a particular condition. But statistics are calculated on the outcomes of a given population, and it is not always easy to relate these to the personal situation of an individual woman. We have no effective way of measuring the impact of bad news on a woman and her family, or of calculating the level of anxiety experienced by a woman who is waiting for test results. Although researchers have created ways of quantitatively measuring qualitative concepts such as pain and satisfaction, these remain personal and subjective experiences, the intensity of which we can only feel for ourselves, and the relevance of such quantification is unclear.

Where proponents of quantitative research continue to seek objective and absolute truth, those people whose goals are different are only beginning to start thinking about how other forms of knowledge, such as experience or intuition, are evaluated. On whose standards do we judge this: the birthing woman's, the midwife's, or from the perspective of an external observer, who will bring their own personal biases to the situation? The midwife working with the woman seeking a normal birth in the scenario above might suggest something which has — in her experience — generally worked before, but which might not be a positive experience for this woman. Or the midwife might make a decision that, on reflection, she felt was not the best she could have made; yet the

woman achieved exactly the outcome she desired and is delighted with her experience. If we are to make progress in finding any kind of solutions to these questions, even if those solutions are simply improved ways of asking the questions, we will need to find ways other than the numerical in which to explore them. As David Boyle (2000) suggests:

*"The more we rely on numbers to understand problems or measure aspects of human life, the more it slips through our fingers and we find ourselves clinging to something less than we wanted. Because every person, every thing, every event is actually unique and unmeasurable."* (p14)

### Challenging numerical hegemony

There are a number of theoretical standpoints from which numerical hegemony can be challenged, and those discussed here are certainly not exhaustive. It is possible to choose arguments both from inside and outside the scientific paradigm; indeed, some of those at the forefront of mathematics and physics provide the most compelling arguments against reductionism. While physicists have previously assumed that the way to understand the world is to isolate it into smaller and smaller particles until the 'fundamentals' are reached and understood, the dawning of the 'new physics' in the early twentieth century led some scientists and mathematicians to adapt their views of whether cause does lead to effect in the way we have previously thought. As a consequence, free thinkers such as Gary Zukav are drawing parallels between quantum mechanics and spirituality (Zukav 1991), suggesting that cutting-edge science and the so-called esoteric ways of knowing are not nearly as far apart as we might imagine. Given that some physicists are now trying to explore the whole rather than measure the individual parts, we might look to their work both to justify the questioning of our reliance on reductionist number crunching and to explore the possibilities for the future.

We might choose to accept the pragmatic view that science itself is a useful way – or at least the best way we currently have – of measuring some aspects of the world around us, while acknowledging that not everything can be measured, and concurrently looking for alternative means of exploration. Having said that, if we want to take the stance that science is a useful tool for investigation, we might then want to look more closely at the theoretical and ethical principles which underpin such activity, especially where 'malestream' principles are being used to underpin the study of women (Oakley 1981).

Mathematician Ian Stewart claims that the use of numbers "is the easiest – and consequently the most dangerous – method for finding patterns. It is easy because anybody can do it, and dangerous for the same reason" (Stewart 1996:p5). This realisation begs the question of whether any human can ever be truly free of their beliefs and values. Proponents of Western medicine may claim that this, too, is rational and objective, but it is also deeply political and ideological (Coney 1995), and influenced by economic concerns (Begley 2001). If the tenets of the scientific method were being effectively applied to the data gathered by research into Western medicine, many of the interventions which women experience routinely would have been discarded years ago (Wickham 2003). It is, then, highly debatable whether Western medicine could justifiably be described as a science (Taylor 2001), enabling some people to take the stance that science itself may be a useful – albeit still limited – way of measuring the physical world, but without the same value when it comes to measuring the experience of people.

One of the challenges to the tenet of objectivity, a principle esteemed by those who would claim numerical knowledge to be king, is that viewpoint is all. A constellation of stars seen from Earth would appear as a completely different pattern to a passenger on a spaceship (Stewart

1996). Why then do we assume that the effect of interventions on women's pregnant bodies can be usefully measured in a fixed way, rather than in relation to the psycho-socio-spiritual environment around them? The results of research studies may, then, be relative only to the viewpoint of those carrying out the study.

In several well-cited research studies, including the Canadian term breech trial (Hannah *et al* 2000) and the Bristol third stage trial (Prendiville *et al* 1988), the outcomes measured were not absolute, or 'hard'. Examples of 'hard' outcome measures would include whether a person was alive or dead, or whether a baby was female or male. Although in both of these examples there is the possibility of a grey area (the person who is alive only because of a life support machine; the baby whose gender is uncertain), the extent of this grey area is very small, and unlikely to radically affect study results. Yet in the studies mentioned above, outcome measures included whether a woman was deemed to need a caesarean section, whether she was deemed to need to have her placenta removed manually, and whether she needed a blood transfusion. All of these are examples of clinical decisions, rather than rigid or fixed outcome measures. Different practitioners might make different decisions given the same scenario, depending on factors such as their personal philosophy, current mood and/or whether they had other people urgently needing their attention. As will be explored further below, such studies might tell us more about the attitudes and workload of the clinicians in the trial than about any fixed truths about how women's bodies react in different situations. This is well illustrated by consideration of the 'intention-to-treat' method of analysing numerical data.

### Intention-to-treat

Intention-to-treat (ITT) is a specific form of quantitative data analysis used in clinical trials, which supposedly gives superior results in comparison with other methods

(Newell 1992). Its basic premise is that the data gathered from researching people (in our examples, women) should be analysed according to the group to which the woman was originally randomised to, rather than according to what treatment she actually received. For instance, in a trial evaluating foot massage in a labour ward, a woman may be randomised on admission to receive a foot massage, but her labour might progress so quickly that she gives birth before the massage therapist arrives. Or, having seen that the woman's labour was progressing quickly, the midwife might have made the decision to ask the therapist not to come into the room, in case her presence interfered with the progress of the woman's second stage. ITT analysis will continue to treat the woman's data (perhaps including how satisfied she was with her care) in the same way as the other women who were randomised to the 'foot massage' group; the majority of whom (we hope) actually received their massage.

One of the key words in the ITT debate is *pragmatic*. Use of ITT enables researchers to calculate a pragmatic estimate of the benefit – or detriment – of an intervention rather than telling us what would happen if everybody received the treatment to which they were allocated. In other words, it acknowledges that in research, as in the real world and the example above, not everybody will get – or take – the treatment as allocated (Hollis & Campbell 1999). To illustrate this point, proponents of ITT often cite the example of people who are given drugs to take, but who do not comply with the treatment.

A good illustration of the effective use of ITT is Dallal's example of a medical study which randomises severely ill people to either surgery or drug treatment (Dallal 1998). As the people are severely ill to begin with, it is likely that deaths would occur early on in both groups. But while those who have been allocated to the 'drugs' group have already started taking their drugs, some of the people who died in the 'surgery' group might not yet have been operated upon. To

remove those people from the analysis altogether changes the baseline level of health in the people who are entering each group; the surgery group has a larger number of healthy people – because those who died early, and so were probably quite unhealthy to begin with, have now been removed from the sample completely. Ultimately, doing this might falsely inflate the level of deaths in the surgery group and bias the outcomes of the research. So the outcome data is analysed in the groups to which the people were originally allocated, giving a fairer comparison of the pharmaceutical and surgical treatments being tested.

The issue of non-compliance, one of the factors taken into account by the use of ITT, has a significant application in public health: by using ITT, researchers can offer policymakers an idea about what will happen if a recommendation is made to the public about a certain health behaviour. An ITT analysis acknowledges the fact that some people will ignore the advice given to them, and will lead to results which combine the behaviours of those who do everything the practitioner tells them, and those who throw their pills away and go out and drink beer instead.

Yet this is also the basis of one of the problems with ITT. Individual women are not populations, nor is any individual woman 'the general public'. There is a fundamental incompatibility between the mass use of ITT in trials evaluating pregnancy and childbirth interventions, and the principles of woman-centred care. It might well be useful for policymakers to be able to determine what the level of need for and potential outcomes of particular interventions are for the population as a whole, but midwives work with individual women. We want to be able to give these women personalised information about what the chances of a particular outcome are if they follow a given course of action. If no data is available on what happens if advice or suggestions are followed 'perfectly', what information can we offer the woman who truly wants to do the best thing in relation to the

evidence and knowledge we have, and who is willing to follow any advice offered in order to secure the best outcome? Why should the data on which her choices could be based be contaminated by the desire to achieve results which relate across the spectrum of the population and include those people who did not follow the advice they were given?

The second aspect of the compliance issue – and perhaps the most important in relation to some of the large studies of birthing women which are informing policies and guidelines – is that ITT not only attempts to eradicate non-compliance by people receiving treatment; it also allows for the problems of non-compliance, deviance from protocol and bias in practitioners as well. As above, researchers understand that not all of the 'subjects' in each group will get the treatment as planned, but they also know that this is, in part, because clinicians have their own inherent biases about what is 'best'. For example:

- In the Canadian term pregnancy trial, which compared induction of labour at 41 weeks of pregnancy with serial monitoring, 32.5% of women who were assigned to the 'monitoring' group had their labours induced (Hannah *et al* 1992).
- In the Canadian term breech trial, which compared caesarean section with assisted vaginal breech delivery, 43.3% of women who were randomised into the 'vaginal delivery' group actually had a caesarean (Hannah 2000).
- In the Bristol third stage trial, less than half of the women in the physiological third stage group had what the researchers considered a physiological third stage (Prendiville 1998) and it should be borne in mind midwives who are experienced in supporting physiological third stage might judge that even fewer of these third stages were truly physiological.

One explanation for each of these figures is that, in each case, the

'expectant', physiological or less-interventive option was, in practice, truly detrimental to significant numbers of women, who ended up in genuine need of intervention. It is also possible that practitioners were, for whatever reason, more comfortable with the use of the intervention, be it induction of labour, caesarean section or managed third stage, or more likely to worry about the women in the control groups, causing them to revert to technological management more quickly. A midwife or obstetrician who has practised for a few years in a system which fears postpartum haemorrhage and learned that active management of the third stage of labour is safer than physiological third stage might then be more nervous when looking after a woman who has been allocated to the physiological group in a trial comparing different types of third stage. It is not difficult to see how that fear could translate into some degree of interference (in what the practitioner perceives to be the best interests of the woman), or a faster decision to do a manual removal than might be made on behalf of a woman who had the 'benefit' of early oxytocic drugs.

Another example concerns the woman who is in the 'watchful waiting' or 'monitoring' group of a trial evaluating induction at a specified point in pregnancy. It is possible that a practitioner might be more worried about a non-reassuring CTG trace in a woman who was in spontaneous labour at 42 weeks of pregnancy than if the same woman, with the same non-reassuring trace, was being induced at 41 weeks. Those caring for the woman in the 'spontaneous labour' arm of the trial might be quicker to suggest a caesarean than those caring for the woman who is being induced at what is perceived to be the appropriate time. This is even more likely to be the case where the woman whose labour is being induced is experiencing the kind of care which was the norm in that setting before the trial began.

It should also be noted that, where the trial has been carried out in one kind of setting, or even in just one unit, the

results of the trial then only apply to that kind of setting. The vast majority of clinical trials are conducted in hospitals, which makes it difficult to be able to offer data to women who are choosing out-of-hospital births. This has been pointed out a number of times in relation to the Canadian term breech trial. If, as Michel Odent commented, the only thing we learned from this trial is that, "a breech birth in a conventional hospital and in the presence of an obstetrician is dangerous" (Culpin & Odent 2003), then surely we need to investigate the alternatives to the hospital setting and the obstetrician as the accoucheur of choice in breech birth?

Proponents of ITT feel that, because all of these different kinds of biases would occur in the real world, analysis using ITT is the most appropriate way of presenting the data gathered. Again, the results are pragmatic: they take into account biases in practitioner's attitudes which affect women in the real world just as much as in trials. The bottom line is that the results in each of the studies mentioned, which are given as relating to either the managed approach (caesarean, induction, actively managed third stage) or the less-managed (assisted breech delivery, serial monitoring, physiological third stage) include a significant proportion of women in each group who did not have that particular treatment. If results are to be given according to ITT, then at the very least we also need to see concomitant on-treatment analysis; the results according to the treatment the women actually received. Is it fair to say to women, "if you choose option A, then you have a better chance of a good outcome than if you choose option B", when we know perfectly well that nearly half of women who chose option A in the relevant study ended up actually experiencing option B anyway?

This becomes even more confusing (especially for the women having to make these decisions) when we have to admit that such advice is based on the premise that a significant proportion of women will not get what they choose, often because of the

prevailing professional attitudes or the constraints on the unit that day. Perhaps we should conduct a study to ask women whether they would rather be given data which is pragmatic, and based on effectiveness — acknowledging that, for lots of reasons, the ideal does not always happen in the real world — or data which is based on efficacy: the measurement of an intervention under ideal conditions. When, as in the examples above, so many women are not getting the treatment as planned, it may be better to arm women with knowledge about outcomes in both ideal and pragmatic contexts, along with tools which can help them find their way through the system towards the treatment or care that they would like to receive.

Intention-to-treat has often been used improperly, with inconsistency between trials using this method of analysis (Hollis & Campbell 1999) and, even when used as intended, it may not be the magical solution which some people suggest it to be. On balance, it should be said that there are times when ITT is the most appropriate form of analysis. There are also situations when its use seems bizarre (Dallal 2003). As far as carrying out the actual research is concerned, there are occasions where it would be more helpful to provide results from both ITT and on-treatment analysis, and occasions where the use of on-treatment analysis alone is appropriate. Just as midwives understand the need to tailor care for individual women, researchers need to tailor methods of analysis to the specific needs and issues of each particular research study. There is often room for development of a specific method of analysis which fits both the question and its context. In the meantime, there is a need for midwives to develop a reasonable level of working knowledge and confidence in the philosophy and the maths behind such studies in order to be able to 'translate' these into information which becomes useful and relevant for individual women.

## Contextualising knowledge

Given that philosophers have spent hundreds of years trying — and generally failing — to come up with a justification for the existence of absolute, fixed truth, an observer of Western thought might be surprised to find that we are still talking about 'evidence'-based practice, in a social context where the term evidence historically relates to the legal concept of proof (Kirkham 2003). Perhaps it is time for midwives to become more involved in epistemological debates around the kind of approach which is most appropriate for contemporary midwifery practice.

Most of the major clinical trials of medical interventions in birth have been conducted in (generally) male-oriented philosophical and systemic contexts in which childbirth has been perceived as danger-filled and the idea of routine medical intervention has not been seen as especially problematic. When attempting to use the results of these trials to inform decision and practice, it becomes necessary for midwives and women to take into account all sorts of contextual factors; such as the personal and professional philosophy, and the degree of trust or fear that was held by the practitioners in the study. This may be the only way in which data derived from models of practice that are based on a general fear of birth or a specific dread of litigation can be made helpful to the woman who has trusting midwives or obstetricians with a different philosophy from their colleagues who were involved in the study.

There is almost always a need for knowledge — or evidence — to be considered in context by the person who is using it to inform decisions. This may, for some, entail the adoption of a relativist viewpoint — taking the stance that all knowledge is relative to time and place; or, as Thomas Kuhn put it, one era's 'science' is another era's 'error' or 'superstition' (Kuhn 1996). Yet the idea that knowledge is tentative, relative and different for each of us seems incompatible with the current emphasis on national guidelines of so-called 'excellence', which recommend courses of action based on the interpretation of

research and statistical knowledge by a relatively small group of stakeholders. This is especially the case when we realise that it is not as simple as taking the guidelines and applying them to each individual midwife's area of practice, but that some of the guidelines produced are in fundamental opposition to the views of midwives who have explored the same evidence from a midwifery model perspective.

If we want to redress the balance and attempt to capture and validate some of the 'softer' forms of midwifery knowledge, such as experience and intuition, there remains the same need to apply some kind of contextual awareness to these more personal ways of knowing. We need to be able to differentiate between truly useful knowledge gathered from reflecting upon a lifetime of midwifery experience and the tendency to justify entrenched practice on the grounds that, 'this has been my experience for 20 years and it has always worked for me'. Whether or not individual midwives feel more comfortable with the use of research data, with more personal ways of knowing, or seek to combine as many forms of knowledge as they can, there may well be a need for midwives to collectively decide where we stand,

philosophically, on the question of knowledge.

The reality of the problems we have created by focusing on the scientific and the numerical seems fairly stark. Some of the midwives who have qualified in the last few years have told me they can analyse a research paper to pieces, but they were never taught any of the 'old tricks' for getting women through long labours. Some women who simply don't want the 'routines' or interventions or places of birth that are offered to them, are told that none of their local midwives are trained in the particular art or skill required. Whichever way one manipulates the statistics, the majority of women experience uncomplicated pregnancies, and it is these women for whom we so desperately need to recapture those ways of knowing and practising midwifery which are about the normal and the natural and only occasionally about the numerical.

#### References

- Begley CM (2001). 'Giving midwifery care': student midwives' views of their working role. *Midwifery* 17(1):24-34.
- Boyle D (2000). *The tyranny of numbers: why counting can't make us happy*. London: Harper Collins.
- Coney S (1995). *The menopause industry*. 2nd ed. London: The Women's Press.
- Culpin E, Odent M (2003). Home breech birth. *Practising Midwife* 6(1):11.

Dallal GE (1998). Intention-to-treat analysis. <http://www.tufts.edu/~gdallal/it.htm> [Accessed 15 July 2003].

Hannah M, Hannah WJ, Hellmann J *et al* (1992). Induction of labor as compared with serial antenatal monitoring in post-term pregnancy. A randomized controlled trial. *New England Journal of Medicine* 326(24):1587-92.

Hannah ME, Hannah WJ, Hewson SA *et al* (2000). Planned caesarean section versus planned vaginal birth for breech presentation at term: a randomised multicentre trial. *Lancet* 356(9239):1375-83.

Hollis S, Campbell F (1999). What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ* 319(7211):670-4.

Kirkham M (2003). Personal correspondence [unpublished].

Kuhn TS (1996). *The structure of scientific revolutions*. 3rd ed. Chicago: University of Chicago Press.

Oakley A (1981). Interviewing women: a contradiction in terms? In: Roberts H ed. *Doing feminist research*. London: Routledge and Kegan Paul, pp 30-61.

Prendiville WJ, Harding JE, Elbourne DR *et al* (1988). The Bristol third stage trial: active versus physiological management of third stage of labour. *BMJ* 297(6659):1295-300.

Shelton HM (1995). *The myth of medicine*. Boca Raton: Cool Hand Communications.

Stewart I (1996). *Nature's numbers: discovering order and pattern in the universe*. London: Phoenix.

Taylor M (2001). Thoughts on science, RCTs and midwifery knowledge. *Midwifery Matters* 91:3-8.

Wickham S (2003). Feminism and ways of knowing. In: Stewart M ed. *Pregnancy, birth and maternity care: feminist perspectives*. Oxford: Elsevier [in press].

Zukav G (1991). *The dancing wu li masters: an overview of the new physics*. London: Ryder.

Wickham S. MIDIRS Midwifery Digest, vol 13, no 4, Dec 2003, pp 439-444.

Original article written for MIDIRS by Sara Wickham, midwifery tutor.

#### ADVERTISEMENT

## Christmas at the MIDIRS book shop!

**With the beginning of the Christmas season now upon us, there is no better way to save time than to visit our online bookshop. With over 200 midwifery titles available you will be sure to find the perfect present. Just visit [www.midirs.org](http://www.midirs.org) and click on 'shop' option.**

