anti-inflammatory drugs. Dieppe *et al.*² together with Brandt and Bradley,³ reviewed all identifiable controlled studies, which showed that "paracetamol reduced pain and tenderness significantly" compared with placebo; "Voltarol significantly improved mobility score" compared with Distalgesic; "ibuprofen improved pain at rest significantly" compared with paracetamol; "diclofenac 100 mg daily produced significantly greater improvement in pain and mobility than ... dextropropoxyphene 180 mg and paracetamol 1.95 g daily"; and there was "strong patient preference for ketoprofen rather than paracetamol/propoxyphene." Brandt and Bradley concluded that paracetamol was as good as ibuprofen in osteoarthritis since scores obtained with the health assessment questionnaire showed no difference even though pain on walking ($P=0.10$) and at rest ($P=0.05$) did.⁴

Because of these data I believe that it was wrong to conclude that "the assertion that [non-steroidal anti-inflammatory drugs] are more effective ... than paracetamol is not yet substantiated."⁵ The inadequacy of paracetamol is further supported by a recent survey in Nottingham of 1137 patients receiving non-steroidal anti-inflammatory drugs in general practice; 51% named their current non-steroidal anti-inflammatory drugs as the most effective treatment they had received, compared with only 6% favouring other analgesics that were not non-steroidal anti-inflammatory drugs.⁶

While anxiety about the gastrointestinal toxicity of non-steroidal anti-inflammatory drugs is legitimate, this accounts at most for one admission to hospital per 45 patient years in elderly people.⁶ Paracetamol is a relatively poor analgesic with imperfectly assessed gastrototoxicity that is likely to make no more than a minor contribution to reducing the problems of complications associated with non-steroidal anti-inflammatory drugs.

C J HAWKEY*
Professor of gastroenterology

Department of Medicine,
Division of Gastroenterology,
University Hospital,
Queen's Medical Centre,
Nottingham NG7 2UH

*The author does research on non-steroidal anti-inflammatory drugs, some of which has been sponsored by pharmaceutical companies.

- 1 March L, Irwig L, Schwarz J, Simpson J, Chock C, Brooks P. n of 1 trials comparing a non-steroidal anti-inflammatory drug with paracetamol in osteoarthritis. *BMJ* 1994;309:1041-6. (22 October.)
- 2 Dieppe PA, Frankel SJ, Toth B. Is research into the treatment of osteoarthritis with non-steroidal anti-inflammatory drugs misdirected? *Lancet* 1993;341:353-4.
- 3 Brandt KD, Bradley JD. Simple analgesics versus NSAIDs for osteoarthritis. *Lancet* 1993;341:770-1.
- 4 Bradley JD, Brandt KD, Katz BP, Kalasinski LA, Ryan SI. Comparison of an antiinflammatory dose of ibuprofen, an analgesic dose of ibuprofen, and acetaminophen in the treatment of osteoarthritis of the knee. *N Engl J Med* 1991;325:87-91.
- 5 Cullen DJ. Suitability of ibuprofen for widespread ulcer sparing use in the community. *Gastroenterology* 1994;106:A6.
- 6 Silverstein FE, Geis GS, Struthers BJ, and the MUCOSA study group. NSAIDs and gastrointestinal injury: clinical outcome, the musosa trial. *Gastroenterology* 1994;106:A180.

Caution in interpretation needed

EDITOR,—Considerable caution must be exercised in interpreting the results of Lyn March and colleagues' n of 1 trials in osteoarthritis.¹ For example, the authors observe that "seven of he eight patients who changed treatment from baseline changed to paracetamol ($P=0.07$, exact binomial)." Of the 20 patients analysed, 16 were taking a non-steroidal anti-inflammatory drug and only three were taking paracetamol. Therefore, at most three patients could have changed their treatment from paracetamol.

Furthermore, the authors implicitly assume that

failure to find a significant difference in favour of diclofenac in the n of 1 trials considered constituted adequate grounds for preferring paracetamol. In drug development generally it is many years since failure to find a significant difference has been accepted as "proof" of equivalence. For fixed effects an n of 1 trial with three pairs of treatment periods has roughly the same power as a two period crossover with three patients; this is inadequate to establish treatment effects in osteoarthritis. We have reconstituted the variances of the individual estimates of the differences between treatments from the authors' table and find an average variance of 33 for pain scores. If the clinically relevant difference (between paracetamol and diclofenac) was 10 mm the power to detect a significant difference would be only 33%; thus, using a significance test under such circumstances one would, on average, incorrectly designate two thirds of patients as "adequately controlled" with paracetamol.

It is true that when the data are examined with a random effects model there is clear evidence of heterogeneity. Whether a random or fixed effects model is used, however, the overall treatment effect is significant and in favour of diclofenac ($P=0.04$ and $P<0.0001$, respectively).

"Peering at P values" is not an appropriate way to use data from n of 1 trials.² Strategies for combining information from individuals with information from groups are needed.² In this series of n of 1 trials flaws in the methodology will have biased downwards the estimate of the proportion of patients responding better to diclofenac.

We also consider that, whereas a small series of trials in selected patients (essentially a "convenience" sample) might be suitable for showing the capability of one drug to prove superior to another (which is the case here for diclofenac), it could never, and particularly if analysed in this way, support the generalisation that "in osteoarthritis many patients currently receiving or being considered for non-steroidal anti-inflammatory drugs may achieve adequate control with paracetamol."

STEPHEN SENN
Biometrician

RAJESH BAKSHI
Central medical adviser

NATHALIE EZZET
Biometrician

Medical Department,
Ciba-Geigy AG,
Basle CH-4002,
Switzerland

- 1 March L, Irwig L, Schwarz J, Simpson J, Chock C, Brooks P. n of 1 trials comparing a non-steroidal anti-inflammatory drug with paracetamol in osteoarthritis. *BMJ* 1994;309:1041-6. (22 October.)
- 2 Senn SJ. Suspended judgment: n-of-1 trials. *Control Clin Trials* 1993;14:1-5.

Authors' reply

EDITOR,—We regard n of 1 trials as a useful tool to help a patient and practitioner decide whether non-steroidal anti-inflammatory drugs offer appreciably better relief than paracetamol. Only then do we think it worth using non-steroidal anti-inflammatory drugs, given their risk of serious side effects. The conceptual bias discussed by C J Hawkey is not a bias but rather our view of sensible decision making. Nor is statistical bias an issue because the decision about treatment was made not by "peering at P values" but rather by the patient and general practitioner viewing their individual plots of pain and stiffness scores and side effect profiles. We agree with Stephen Senn and colleagues that few patients could have changed from paracetamol. The real issue, however, was that seven of 16 who were previously taking non-steroidal anti-inflammatory drugs thought that these drugs did not offer appreciable relief over paracetamol and six of these seven thought that

paracetamol continued to relieve symptoms adequately three months later.

We agree that failure to find a difference should not be regarded as proof of equivalence, and we at no stage claimed that it was. We were not addressing drug development, nor were we making statements about which drug in general is better for the treatment of osteoarthritis. Rather, as we made explicit in the evaluation of efficacy in our paper, in a setting in which one of the possible treatments (non-steroidal anti-inflammatory drugs) may have hazardous side effects we asked the patients and their general practitioner to decide whether there was any appreciable difference in response. The table in our paper shows clearly that most subjects' responses were more favourable during the diclofenac periods, yet many elected to continue taking paracetamol.

Our primary aim was to analyse individual patients' responses, whereas an overall effect seems to be the primary focus of interest of all the authors commenting on the article, including K B Saunders.¹ We agree with Senn and colleagues that there is heterogeneity, as shown in the plot of individual treatment effects in our paper, and consider that this makes n of 1 trials particularly important. To evaluate overall group effects there is no substitute for randomised controlled trials, and we were surprised to find so few comparing non-steroidal anti-inflammatory drugs with paracetamol alone. We stand by our claim that n of 1 trials are acceptable and clinically useful in osteoarthritis.

LYN MARCH
Research fellow
JENNIFER SCHWARZ
Research assistant

Florance and Cope Professorial Department of Rheumatology,
Royal North Shore Hospital,
St Leonards, NSW 2065,
Australia

LES IRWIG
Associate professor

NSAHS Public Health Unit and Department of Public Health,
University of Sydney,
NSW 2006

JUDY SIMPSON
Senior lecturer in biostatistics
CATHERINE CHOCK
Research assistant

Department of Public Health,
University of Sydney,
NSW 2006

PETER BROOKS
Professor of medicine

St Vincent's Hospital,
Darlinghurst,
NSW 2010

- 1 Saunders K. n of 1 trials. *BMJ* 1994;309:1584. (10 December.)

Childhood antecedents of schizophrenia

EDITOR,—Brian Martindale presents selective evidence to suggest that the environmental contribution to schizophrenia is considerable,¹ neglecting to mention recent strong evidence supporting genetic transmission of schizophrenia based on the biological relatives of schizophrenic adoptees.² The diagnoses and analysis of this supportive study have, furthermore, been independently replicated.³

Moreover, a more recent report of the Finnish adoptive family study than that cited by Martindale has been published.⁴ This states that schizophrenia (according to the definition in the *Diagnostic and Statistical Manual of Mental Disorders*, third edition, revised) was diagnosed in nine adoptees, eight of whom were in the group whose biological mothers had either schizophrenia or "spectrum psychoses" (n=155) while just one was in the control group with normal biological mothers (n=186). Martindale's assertion that adopted offspring of schizophrenic mothers had

more severe psychopathology than control adoptees when they were reared in disturbed adoptive families is relevant only to a blanket of diagnoses in the adoptees. The diagnoses include the relatively frequent diagnosis of personality disorder as well as borderline syndrome in addition to psychosis. The ratings of the adoptive family environment do not relate to the diagnosis of schizophrenia alone in the adoptees.

McGuffin *et al* recently modelled the genetic contribution to schizophrenia, producing the best fit with a genetic contribution of 89%.⁵ They also state that "While there is good evidence that psychosocial factors, such as high expressed emotion at home or life events, can hasten relapses or precipitate onsets, few would now argue that such factors are likely to be causal." In conclusion they state that "the environmental component is small and there is a lack of consensus concerning the identity of specific environmental factors." Martindale seems to fall outside this apparent consensus.

With such a lack of evidence that family environment has a role in the aetiology of schizophrenia, it is questionable whether psychotherapeutic intervention will have any influence on the expression of this disease. Research directed towards elucidating the genetic and epigenetic aetiology of schizophrenia seems potentially much more rewarding since this may allow the development of more acceptable and effective drug treatment and also the possibility of gene therapy.

ANDREW D PATERSON
Medical student

Medical Faculty,
University of Glasgow,
Glasgow G12 8QQ

- 1 Martindale B. Childhood antecedents of schizophrenia. *BMJ* 1995;310:57. (7 January.)
- 2 Kety SS, Wender PH, Jacobsen B, Ingraham LJ, Jansson L, Faber B, *et al*. Mental illness in the biological and adoptive relatives of schizophrenic adoptees—replication of the Copenhagen study in the rest of Denmark. *Arch Gen Psychiatry* 1994;51:442-55.
- 3 Kendler KS, Gruenberg AM, Kinney DK. Independent diagnoses of adoptees and relatives as defined by DSM-II in the provincial and national samples of the Danish adoption study of schizophrenia. *Arch Gen Psychiatry* 1994;51:456-68.
- 4 Tienari P, Wynne LC, Moring J, Lahti I, Naarala M, Sorri A, *et al*. The Finnish adoptive family study of schizophrenia—implications for family research. *Br J Psychiatry* 1994; 164(suppl 23):20-6.
- 5 McGuffin P, Asherson P, Owen M, Farmer A. The strength of the genetic effect—is there room for an environmental influence in the aetiology of schizophrenia? *Br J Psychiatry* 1994;164:593-9.

Rates of episiotomy

Data on parity are not given

EDITOR,—Tine Brink Henriksen and colleagues report that the use of feedback to midwives produced a welcome decrease in the incidence of episiotomy.¹ Unfortunately, data on the parity of the parturients are not shown. It is now widely accepted that the incidence of perineal trauma is an important index of care, and this information is therefore reported in our annual clinic report. Our audit, however, highlights the importance of analysing such information by parity (table), and we believe strongly that such an analysis is an

Perineal trauma in all deliveries by parity. Figures are numbers (percentages) of women

	Primigravidas	Multigravidas	Total
Labial laceration	92 (4.2)	116 (2.9)	214 (3.4)
Tear:			
1st degree	295 (12.4)	995 (25.3)	1290 (20.5)
2nd degree	243 (10.3)	369 (9.4)	612 (9.7)
3rd degree of muscle	23 (1.0)	5 (0.1)	28 (0.4)
3rd degree of anal mucosa	12 (0.5)	6 (0.2)	18 (0.3)
Episiotomy	988 (42.0)	411 (10.4)	1399 (22.2)

essential element of any feedback to individual practitioners. Our results also show that third degree tears are more likely to occur in primigravidas. Thus any analysis that examines the impact of decreasing the rate of episiotomy on the incidence of third degree tears must make allowance for parity.

MICHAEL TURNER
Master

CATHERINE CASEY
Assistant master

Coombe Women's Hospital,
Dublin 8,
Republic of Ireland

- 1 Brink Henriksen T, Møller Bek K, Hedegaard M, Jørgen Secher N. Methods and consequences of changes in use of episiotomy. *BMJ* 1994;309:1255-8. (12 November.)

Conclusions and validity of data cannot be judged

EDITOR,—Tine Brink Henriksen and colleagues fail to give adequate clinical information so that we can judge the validity of their data and their conclusions on changes in the use of episiotomy.¹ First and foremost is their failure to analyse their results for primigravidas and multigravidas separately.

As most parturients who require episiotomy are primigravidas, combining the results for both groups may mask important underlying differences and outcomes. The high incidence of pudendal anaesthesia (43.1% and 33.4% before and after the distribution of the management awareness profiles) seems to be at variance with modern obstetric practice. What other forms of analgesia were available, and, in particular, what was the incidence of regional epidural analgesia? Other factors that have not been taken into account in the analysis of the data are differences in rates between night and day and position for delivery. For primigravidas the rate of intact perineum is 58% in the lateral position compared with 33% in the dorsal position.²

We have been auditing our midwifery practice for the past 14 years to ensure that every mother having a normal delivery is delivered with the minimum perineal trauma possible. The standards aspired to for primigravidas are an episiotomy rate of 20% and an intact perineum in at least 38%, while for multigravidas the standards are an episiotomy rate of 5% and an intact perineum in 60%.^{3,4} In a review of perineal management in 1990 awareness profiles similar to those used by Brink Henriksen and colleagues were distributed to every midwife. Six months later an audit showed that the above standards had been met for multigravidas and that, while target rates for primigravidas had been achieved in the first three months, the episiotomy rate in the second three months had risen by 6%. This suggests that continual feedback is necessary, particularly for primigravidas, if standards are to be maintained.⁵

Today many women prefer to have home births to avoid what is seen as unnecessary medical intervention in a natural process. Episiotomy has a place in modern obstetric care, and we must ensure that it is used appropriately in the interests of safe childbirth for both the mother and her newborn infant.

JOHN F STRATTON
Registrar
H GORDON
Head of department
M LOGUE
Midwifery sister

Northwick Park Hospital,
Harrow,
Middlesex HA1 3UJ

- 1 Brink Henriksen T, Møller Bek K, Hedegaard M, Jørgen Secher N. Methods and consequences of changes in use of episiotomy. *BMJ* 1994;309:1255-8. (12 November.)
- 2 Logue M. Putting research into practice: perineal management during delivery. In: Robinson S, Thomson AM, eds. *Midwives, research and childbirth. II*. London: Chapman and Hall, 1991:252-71.

- 3 Sleep J, Grant A, Garcia J, Elbourne D, Spencer J, Chalmers I. West Berkshire perineal management trial. *BMJ* 1984;289: 587-90.
- 4 Welch A, Chapple J, Beard R, eds. Intrapartum care. In: *North West Thames region annual maternity figures*. London: St Mary's Maternity Information System, 1990:57-9.
- 5 Logue M. Midwives' perineal management. In: *Nursing quality and audit activities in midwifery, neonatal, and gynaecology*. Harrow: Northwick Park Hospital Family Services Directorate, 1991:16-20.

Authors' reply

EDITOR,—Both letters about our paper emphasise that parity is an important variable to consider when dealing with perineal trauma and obstetrics in general. We agree. Parity was not an element of our feedback to the midwives but will be in the future. We, however, found no difference in parity between different time periods (categorised in three month intervals) or between different groups of midwives (categorised as in table III in our paper). Although more episiotomies and perineal tears were found in primiparous women in our study, our results were the same in relation to the changes regardless of stratification by parity.

The table shows that the drop in episiotomies among primiparous women corresponded to 11.4% and among multiparous women to 3.5%.

Perineal lesions in 3919 spontaneous vaginal deliveries assisted by 30 midwives before and after distribution of management awareness profiles in September 1990, by parity. Figures are numbers (percentages) of women

	Primiparous women	Multiparous women
Episiotomy:		
Before	384 (54.2)	231 (24.3)
After	423 (42.8)	262 (20.8)
Intact perineum:		
Before	180 (25.3)	353 (37.0)
After	298 (30.1)	494 (39.3)
Complete tear of anal sphincter:		
Before	10 (1.4)	7 (0.7)
After	14 (1.4)	9 (0.7)
Incomplete tear of anal sphincter:		
Before	9 (1.3)	6 (0.6)
After	21 (2.1)	6 (0.5)

The relative decreases (21.0% and 14.4% respectively), however, were more comparable because of the higher initial rate among primiparous women. Thus the observed overall changes were reflected in a decrease among all women, not just primiparous women. Small numbers after stratification by parity produce some statistical uncertainty and the possibility of type II errors when the occurrence of anal sphincter lesions is interpreted. It follows that our recommendations are not specific to parity as the guidelines put forward by John F Stratton and colleagues are. Thus our recommendations apply to populations having spontaneous vaginal deliveries with a ratio of primiparous to multiparous women of about 45% to 55%.

We agree with Stratton and colleagues that some details about obstetric practice and characteristics related to the population should perhaps be considered before discussions of the need for episiotomies. Some of the details on our population that they request were as follows. There was virtually no difference in the frequency of epidural analgesia before and after the feedback (2.6% and 2.5% of all deliveries respectively). The same was true for caesarean sections (8.0% and 9.7%) and vacuum or forceps deliveries (11.7% and 10.2%). Furthermore, we found no implications for our results when data were stratified by the time of the day: the incidence of episiotomy during the night was 33.1% and during the day 33.0%. Unfortunately, our data on the presence of students and the position for delivery are incomplete, but a test sample of around 800 deliveries showed that the frequency of deliveries in the lateral position was 6.0% during our observation period and 5.1% during the period after the feedback. Stratton and