Open Access: free full text available online
For the full versions of these articles see bmj.com

Early head injury and attention-deficit/hyperactivity disorder: retrospective cohort study

Heather T Keenan,[1] Gillian C Hall,[2] Stephen W Marshall[3]

EDITORIAL by Zwi and Clamp

[1]Department of Pediatrics, University of Utah, Salt Lake City, UT 84158, USA
[2]Grimsdyke House, Ravenscroft Park, Barnet, Hertfordshire EN5 4ND
[3]Department of Epidemiology, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC 27699, USA
Corresponding to: H T Keenan
heather.keenan@hsc.utah.edu

Cite this as: *BMJ* 2008;337:a1984
doi:10.1136/bmj.a1984

ABSTRACT

Objective To explore the hypothesis that medically attended head injury in young children may be causal in the later development of attention-deficit/hyperactivity disorder.

Design Retrospective cohort study.

Setting Health improvement network database (1988-2003), a longitudinal UK general practice dataset.

Participants All children registered in the database from birth until their 10th birthday.

Main outcome measures Risk of a child with a head injury before age 2 developing attention-deficit/hyperactivity disorder before age 10 compared with children with a burn injury before age 2 and children with neither a burn nor a head injury.

Results Of the 62 088 children who comprised the cohort, 2782 (4.5%) had a head injury and 1116 (1.8%) had a burn injury. The risk of diagnosis of attention-deficit/hyperactivity disorder before 10 years of age after adjustment for sex, prematurity, socioeconomic status, and practice identification number was similar in the head injury (relative risk 1.9, 95% confidence interval 1.5 to 2.5) and burn injury groups (1.7, 1.2 to 2.5) compared with all other children.

Discussion Medically attended head injury before 2 years of age does not seem to be causal in the development of attention-deficit/hyperactivity disorder. Medically attended injury before 2 years of age may be a marker for subsequent diagnosis of attention-deficit/hyperactivity disorder.

INTRODUCTION

Studies have suggested both that injury is more common in children with attention-deficit/hyperactivity disorder (ADHD) and that moderate to severe traumatic brain injury in school age children results in the development of ADHD (secondary ADHD).[1-4] We hypothesised that children with a medically attended head injury before 2 years of age would be more likely to be subsequently diagnosed as having ADHD than children who did not have a medically attended head injury. To investigate this, we selected two comparison groups: children who had a medically attended burn or scald injury without a head injury and children with neither a head injury nor a burn injury.

METHODS

Data came from the health improvement network,[5] a longitudinal dataset from primary care practices in the United Kingdom that contains information collected from 308 practices. Medical events are automatically coded at entry by use of the READ coding system. The study cohort included children who had a date of birth recorded in the health improvement network database since the inception of computerised data collection and continuous enrolment until they received a diagnosis of ADHD or until their 10th birthday. The study period was 1988 to 2003. We excluded children with a diagnosis of ADHD before their second birthday.

We divided the study cohort into three groups. The first was children with early head injury—children whose record contained any READ code for head injury, except for "minor head injury" and "nursing advice for head injury," before the child's second birthday. We extracted information on the source of the referral for patients with early head injury and dichotomised the data according to whether or not the child was referred by a physician to a higher level of care or admitted to hospital. The second group, children with early burn or scald injury, comprised all children who had a READ code for burn or scald recorded before their second birthday. Children from the cohort without head injury or burn injury before age 2 years made up the third group (comparison group).

We extracted baseline data for all children, including codes and dates of head injuries (up to three injuries after 2 years of age) and burn or scald injury, and any code indicating prematurity (<37 weeks gestational age) or child abuse. For each child, we extracted the date of the first diagnostic code for ADHD and the date of the first prescription for drugs commonly used for ADHD. We used the Townsend deprivation index to measure socioeconomic status.

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2008;337:a1984

Data analysis—We used frequencies and percentages for categorical variables and medians with interquartile ranges for non-normally distributed continuous variables. We calculated the relative risk of developing ADHD for each injury group compared with the comparison group. We explored associations between predictor variables and the outcome. We did multivariate modelling, to account for clustering by practice, to examine the association of injury type with the odds of developing ADHD after adjustment for Townsend index, sex, and prematurity. We used Poisson regression to calculate the unadjusted rate of diagnosis of ADHD.

RESULTS

Study population

A total of 62 088 children in the health improvement network database met the study criteria (see bmj.com). Of these, 2782 (4.5%) had a medically attended head injury before 2 years of age and 1116 (1.8%) had a medically attended burn or scald injury before age 2. Study follow-up time was similar for both injured groups and the comparison group (P=0.08). Children with head injuries were slightly younger when injured (median 13.9 (interquartile range 9.0-18.4) months) compared with the burn injury group (14.8 (11.0-18.8) months) (P<0.001). Injured children were more likely to be male than non-injured children (P<0.001). The Townsend index showed greater deprivation among injured children than in the comparison group (P<0.001). The frequency of diagnosis of child abuse was low in all three groups (<0.1% in each group). A higher percentage of children who had had a head injury or burn injury before age 2 had medically attended head injuries after age 2 (P<0.001).

Diagnosis of ADHD

A diagnosis of ADHD after the second birthday and before the 10th birthday occurred in 1.5% (n=934) of the entire study cohort. The incidence of diagnosis of ADHD for the entire cohort was 1.2 (95% confidence interval 1.1 to 1.3) per 1000 person years before age 10 and 1.4 (1.3 to 1.5) per 1000 person years over the entire study period. Most children (769; 70.4%) with a diagnosis of ADHD had a diagnosis of primary ADHD; 284 (26.0%) had a diagnosis of hyperkinetic disorder, and 37 (3.4%) had a diagnosis of overactivity; only 2 (0.2%) children were diagnosed as having hyperkinetic conduct disorder. Compared with the comparison group, children in the two injury groups had a similarly increased likelihood of being diagnosed as having ADHD after the age of 2 (table). These two groups accounted for 6.3% (n=3898) of the study cohort and 11.3% (126) of the children in the cohort who went on to be diagnosed as having ADHD.

Treatment for ADHD was prescribed to 45.5% (425) of the entire cohort diagnosed as having ADHD before 10 years of age. Children with a history of burn or head injury were not more likely than their uninjured comparison counterparts to receive treatment once diagnosed. Most children (545; 95.3%) were treated with methylphenidate; relatively few were treated with atomoxetine (17; 3.0%) or dexamfetamine (10; 1.8%).

Children with head injury before age 2 were diagnosed as having ADHD at a slightly younger median age (6.2 (interquartile range 4.0-8.0) years) than children with burn injury (6.8 (4.8-9.0) years) or the comparison group (6.9 (4.6-8.8) years); however, this difference did not reach statistical significance (P=0.07). Male sex (relative risk 4.8, 4.0 to 5.7), a history of prematurity (1.6, 1.1 to 2.4), and a more deprived Townsend index (Cochran-Armitage test for trend, P<0.001) were all positively associated with diagnosis of ADHD.

The adjusted risk of diagnosis of ADHD before age 10 remained higher in both the head injury group and the burn injury group compared with the comparison group (table). In addition, the risk of diagnosis of ADHD increased among children who had a head

Unadjusted and adjusted associations of early injury with diagnosis of attention-deficit/hyperactivity disorder (ADHD)

	Relative risk (95% CI) of ADHD diagnosis before age 10	Relative risk (95% CI) of ADHD diagnosis ever	Comparison
Comparison group			
Unadjusted	1.0 (referent)	1.0 (referent)	Referent
Adjusted*	1.0 (referent)	1.0 (referent)	Referent
Head injury group			
Unadjusted	2.0 (1.6 to 2.5)	1.9 (1.5 to 2.4)	Referent
Adjusted*	1.9 (1.5 to 2.5)	1.8 (1.4 to 2.2)	Referent
Burn injury group			
Unadjusted	2.3 (1.6 to 3.1)	2.0 (1.5 to 2.8)	Referent
Adjusted*	1.7 (1.2 to 2.5)	1.8 (1.3 to 2.5)	Referent

*Adjusted for sex, prematurity, Townsend index, and practice identification number.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Mild traumatic brain injury is thought to be associated with behavioural changes in children

Disentangling pre-existing behaviour problems from the sequelae of injury has been difficult

WHAT THIS STUDY ADDS

Medically attended head injury before age 2 does not seem to be causal in the development of attention-deficit/hyperactivity disorder

Any medically attended injury before age 2 may be a marker for subsequent diagnosis of attention-deficit/hyperactivity disorder

injury after age 2 in all three groups in the cohort (see bmj.com).

DISCUSSION

Children with a head injury before age 2 were twice as likely to be diagnosed as having ADHD as a population based comparison group but not more likely to be diagnosed as having ADHD than another medically attended injury group (burn injury). Thus, contrary to our hypothesis, the head injury itself does not seem to be causal in the development of ADHD. Rather, some other factor seems to be associated generally with early injury and the development of ADHD. These results indicate that medically attended injury before age 2 may be an early marker for behavioural traits that lead to diagnosis of ADHD.

We find it plausible that children who go on to develop clinical ADHD exhibit more risk taking behaviours as young children, and are therefore more likely to be injured before age 2. Consistent with this, children with a head injury or a burn injury before age 2 were also more likely to have medically attended head injuries after age 2 than were comparison children.

Several studies have documented that children who have been diagnosed as having ADHD are more likely to be injured.[6 7] This is important, as children with ADHD before a head injury tend to fair more poorly in both cognition and adaptive behaviour after serious traumatic brain injury than do children without pre-injury behavioural concerns.[2 8] Additionally, children with ADHD may be more vulnerable to post-concussive syndrome after a mild traumatic brain injury.[9]

A limitation imposed by the dataset is the lack of detail about the seriousness of the injury event. The health improvement network database codes whether the child was admitted to hospital or seen in the emergency department for the injury; however, this field was not coded for 50% of the patients. Severity of injury was also unavailable from the diagnostic codes. Secondary ADHD has been associated more strongly with severe and moderate traumatic brain injury than with mild injury.[10] Thus, inclusion of some children with moderate to severe injury in this dataset would have tended to bias the results towards a higher risk of ADHD in the children with head injuries than in the controls with burns, which we did not find. Also, children may not have received formal psychiatric evaluation and formal diagnosis of ADHD. This is true for all children in the dataset, however. The overall prevalence of ADHD in this dataset is low compared with the prevalence estimates of ADHD in US children (3-7%) and similar to that reported in British Columbia (1.6%),[6 11] so the condition was probably not overdiagnosed. Strengths of these data include the ability to follow children over time, to use a second injured control group, and to adjust for factors commonly related to both injury and ADHD.

We thank Mary Thompson and the team from EPIC for providing the health improvement network data and the general practitioners for their continued support of the dataset.

Contributors: See bmj.com.

Funding: None.

Competing interests: None declared.

Ethical approval: Oxfordshire Multicentre Research Ethics Committee (ID# 06/Q1605/32) and the Institutional Review Board of the University of Utah gave approval for use of the health improvement network data.

Provenance and peer review: Not commissioned; externally peer reviewed.

1 Levin H, Hanten G, Max J, Li X, Swank P, Ewing-Cobbs L, et al. Symptoms of attention-deficit/hyperactivity disorder following traumatic brain injury in children. *J Dev Behav Pediatr* 2007;28:108-18.

2 Yeates KO, Armstrong K, Janusz J, Taylor HG, Wade S, Stancin T, et al. Long-term attention problems in children with traumatic brain injury. *J Am Acad Child Adolesc Psychiatry* 2005;44:574-84.

3 Massagli TL, Fann JR, Burington BE, Jaffe KM, Katon WJ, Thompson RS. Psychiatric illness after mild traumatic brain injury in children. *Arch Phys Med Rehabil* 2004;85:1428-34.

4 Yeates KO, Taylor HG. Neurobehavioural outcomes of mild head injury in children and adolescents. *Pediatr Rehabil* 2005;8:5-16.

5 Bourke A, Dattani H, Robinson M. Feasibility study and methodology to create a quality-evaluated database of primary care data. *Inform Prim Care* 2004;12:171-7.

6 Brehaut JC, Miller A, Raina P, McGrail KM. Childhood behavior disorders and injuries among children and youth: a population-based study. *Pediatrics* 2003;111:262-9.

7 Pastor PN, Reuben CA. Identified attention-deficit/hyperactivity disorder and medically attended, nonfatal injuries: US school-age children, 1997-2002. *Ambul Pediatr* 2006;6:38-44.

8 Gerring JP, Brady KD, Chen A, Vasa R, Grados M, Bandeen-Roche KJ, et al. Premorbid prevalence of ADHD and development of secondary ADHD after closed head injury. *J Am Acad Child Adolesc Psychiatry* 1998;37:647-54.

9 Yeates KO, Luria J, Bartkowski H, Rusin J, Martin L, Bigler ED. Postconcussive symptoms in children with mild closed head injuries. *J Head Trauma Rehabil* 1999;14:337-50.

10 Max JE, Lansing AE, Koele SL, Castillo CS, Bokura H, Schachar R, et al. Attention deficit hyperactivity disorder in children and adolescents following traumatic brain injury. *Dev Neuropsychol* 2004;25:159-77.

11 Rappley MD. Attention deficit-hyperactivity disorder. *N Engl J Med* 2005;352:165-73.

Accepted: 8 September 2008

Impact of micronutrient supplementation during pregnancy on birth weight, duration of gestation, and perinatal mortality in rural western China: double blind cluster randomised controlled trial

Lingxia Zeng,[1] Michael J Dibley,[2,3] Yue Cheng,[1] Shaonong Dang,[1] Suying Chang,[4] Lingzhi Kong,[5] Hong Yan[1]

EDITORIAL by Hiremath

[1]Department of Public Health, Xi'an Jiaotong University College of Medicine, PO Box 46, Xi'an, Shaanxi 710061, China
[2]School of Public Health, University of Sydney, Room 307A, Edward Ford Building (A27), University of Sydney, NSW 2006, Australia
[3]George Institute for International Health, PO Box M201, Sydney, NSW 2050, Australia
[4]Chinese Centre for Disease Control and Prevention, Beijing 100050, China
[5]Bureau for Disease Prevention and Control, Ministry of Health, Beijing 100044, China
Correspondence to: H Yan
xjtu_yh.paper@yahoo.com.cn

Cite this as: *BMJ* 2008;337:a2001
doi:10.1136/bmj.a2001

ABSTRACT

Objective To examine the impact of antenatal supplementation with multiple micronutrients or iron and folic acid compared with folic acid alone on birth weight, duration of gestation, and maternal haemoglobin concentration in the third trimester.

Design Cluster randomised double blind controlled trial.

Setting Two rural counties in north west China.

Participants 5828 pregnant women and 4697 live births.

Interventions Villages were randomised for all pregnant women to take either daily folic acid (control), iron with folic acid, or multiple micronutrients with a recommended allowance of 15 vitamins and minerals.

Main outcome measures Birth weight, length, and head circumference measured within 72 hours after delivery. Neonatal survival assessed at the six week follow-up visit.

Results Birth weight was 42 g (95% confidence interval 7 to 78 g) higher in the multiple micronutrients group compared with the folic acid group. Duration of gestation was 0.23 weeks (0.10 to 0.36 weeks) longer in the iron-folic acid group and 0.19 weeks (0.06 to 0.32 weeks) longer in the multiple micronutrients group. Iron-folic acid was associated with a significantly reduced risk of early preterm delivery (<34 weeks) (relative risk 0.50, 0.27 to 0.94, P=0.031). There was a significant increase in haemoglobin concentration in both iron-folic acid (5.0 g/l, 2.0 to 8.0 g/l, P=0.001) and multiple micronutrients (6.9 g/l, 4.1 to 9.6 g/l, P<0.001) groups compared with folic acid alone. In post hoc analyses there were no significant differences for perinatal mortality, but iron-folic acid was associated with a significantly reduced early neonatal mortality by 54% (relative risk 0.46, 0.21 to 0.98).

Conclusion In rural populations in China antenatal supplementation with iron-folic acid was associated with longer gestation and a reduction in early neonatal mortality compared with folic acid. Multiple micronutrients were associated with modestly increased birth weight compared with folic acid, but, despite this weight gain, there was no significant reduction in early neonatal mortality. Pregnant women in developing countries need sufficient doses of iron in nutrient supplements to maximise reductions in neonatal mortality.

Trial registration ISRCTN08850194.

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2008;337:a2001

INTRODUCTION

One of the major causes of low birth weight in developing countries is the poor nutritional status of the mother before and during pregnancy, resulting in restricted fetal growth especially during the third trimester.[1] Unicef has proposed the use of multiple micronutrient supplements in pregnancy that provide the recommended intake for pregnant women.[2] The only antenatal supplement promoted by the Ministry of Health in China is folic acid.[3] There are no specific policies or programmes for the distribution of multiple micronutrient or iron-folic acid supplements during pregnancy, even to disadvantaged women. To provide evidence in China for formulation of public health policy on nutrient supplementation in pregnancy we conducted a community based cluster randomised controlled trial in a disadvantaged rural population.

METHODS

Experimental design

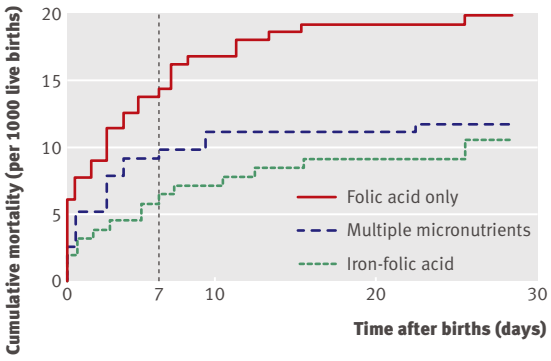
The trial took place in two poor rural counties in Shaanxi Province. A pilot study found a low birth rate of 13% and 57% of women with anaemia (haemoglobin <110 g/l) in the third trimester. We allocated the same treatment to all pregnant women in a given village. In 2001 in the first county there were 14 townships and 234 villages, and in the second county there were 20 townships and 327 villages. Randomisation of villages was stratified by county with an approximately equal distribution of treatments per township.

Study population and sample size

The study sample consisted of all women resident in the counties who became pregnant between August 2002 and January 2006 and who fulfilled trial selection criteria. Limitations in funding meant that we recruited and randomised 5828 eligible women—fewer women than anticipated. In consequence we estimated that our sample of about 1900 women per treatment would provide 80% power to detect a 50 g difference in birth weight between either iron-folic acid or multiple micronutrients and folic acid (control) groups. See bmj.com for further details of power calculation.

Enrolment and pregnancy surveillance procedures

Village doctors, with support from the township maternal and child healthcare workers, recruited women by active surveillance for pregnancy in women of reproductive age. At the start the village doctor conducted a mini-survey of all women of reproductive age living in their village to identify those who were likely to become



Cumulative mortality curves for infants from birth to 28 days by treatment. Log rank test of difference: P=0.055 for three groups; 0.032 for iron-folic acid *v* folic acid; 0.077 for multiple micronutrients *v* folic acid; 0.708 for multiple micronutrients *v* iron-folic acid

pregnant. Pregnancies in women resident in the studied townships or counties were passively detected at antenatal clinics in local health facilities.

Newly identified pregnant women were interviewed and once recruited received three free antenatal care checks, at which they were asked about pregnancy complications and underwent a physical examination. Trained maternal and child healthcare staff with monitoring by project staff collected and recorded all the information collected during the pregnancy and the follow-up visit six weeks after delivery in a pregnancy care record book, which served both as a clinical record and data capture instrument.

Interventions

Villages were randomly assigned for women to receive one of three daily antenatal supplements: multiple micronutrients, iron-folic acid, or folic acid alone (control). The multiple micronutrient supplements contained approximately the WHO/Unicef recommended dietary allowances for each of 15 minerals or vitamins[4] including 30 mg iron, and 400 µg folate. Iron-folic acid supplements contained 60 mg iron and 400 µg folic acid. The folic acid supplement contained 400 µg of folic acid.

Measurement of outcomes

In the six county hospitals and the three largest township hospitals (78% of birth weights), hospital nursing staff measured birth weight within one hour of delivery with an electronic scale. For births in the 31 smaller township hospitals (10% of birth weights) and for home births (12% of birth weights) birth weight was measured with a baby scale. For home deliveries, township maternal and child health staff visited the women at home within 72 hours of delivery to measure the baby and gather information on delivery.

Maternal haemoglobin concentration was measured in capillary blood collected early in the third trimester (gestation 28-32 weeks) from a subsample of 599 pregnant women with a birth outcome, who were consecutively enrolled from 6 July 2004 to 28 October 2005.

Village doctors or hospital staff reported fetal losses during pregnancy, birth outcome, delivery information, and neonatal and maternal deaths; maternal and child health staff recorded data with precoded structured forms. Neonatal survival was assessed at the six week follow-up visit.

Statistical analysis

To assess the effectiveness of randomisation we examined the baseline characteristics of the clusters and the individual pregnant women across treatment groups. A wealth index was constructed from an inventory of 16 household assets or facilities. The mean number of supplements consumed and treatment compliance rates were also examined.

We estimated mean differences for birth weight and gestation and adjusted for the effect of randomisation by villages. The adjusted mean differences in birth weight, birth length, head circumference, and gestation were computed relative to the folic acid group. Similarly, we calculated adjusted mean differences for newborn anthropometry and gestation at birth to compare the multiple micronutrient and the iron-folic acid groups. We adjusted for cluster randomisation. We analysed perinatal mortality only in singleton births. We calculated rates of stillbirth and perinatal death using the number of pregnancies at 28 weeks' gestation as the denominator. Neonatal death rates were calculated with the number of live births as the denominator. Kaplan-Meier survival analysis was used to estimate the survival probabilities of liveborn singleton infants from birth to 28 days and to compare survival across treatment groups. We adjusted for the cluster randomisation. See bmj.com for more details of statistical analysis.

RESULTS

Over the 3.5 year period, there were 7144 confirmed pregnancies from the monitored population. We enrolled and randomised 5828 women, of these there were 4421 (94%) live births with birth weight available for analysis and 222 perinatal deaths. See bmj.com.

The sociodemographic characteristics and the anthropometric measurements at enrolment and the cluster and individual level baseline characteristics were balanced by treatment groups. The reproductive history of the pregnant women was similar across groups and reflected China's "one child policy," with 3585 (61.5%) women having their first pregnancy. The percentage of women delivering at home was balanced across treatment groups.

There was a high level of compliance with the supplementation in all treatment groups. The mean number of doses of supplements consumed per women during pregnancy was 165, and this was similar in each group. See bmj.com.

Impact of nutrient supplementation on anthropometry and gestation at birth

The intracluster correlation coefficient for birth weight was 0.03 (95% confidence interval 0.015 to 0.052).

There was no evidence of an effect of iron-folic acid on mean birth weight (P=0.17) but birth weight was significantly higher (42 g, 7 to 78 g, P=0.019) in the multiple micronutrients group compared with the folic acid alone. The increase in birth weight observed in the iron-folic acid and the multiple micronutrient groups corresponded with reductions in the risk of low birth weight (<2500 g) compared with folic acid alone of 19% (P=0.20) and 22% (P=0.14), respectively.

The intracluster correlation coefficient for gestation at birth was 0.02 (0.004 to 0.036). There was a significant increase in the duration of gestation of 0.23 weeks (0.10 to 0.36 weeks, P=0.001) in the iron-folic acid group and 0.19 weeks (0.06 to 0.32 weeks, P=0.004) in the multiple micronutrient group compared with folic acid. The increase in the mean duration of gestation in the iron-folic acid group corresponded with a 21% reduction in the risk of preterm delivery (<37 weeks) (P=0.13) and a significant 50% reduction in the risk of early preterm delivery (<34 weeks) (relative risk 0.50, 0.27 to 0.94, P=0.031). Compared with folic acid alone, the increase in the mean duration of gestation observed in the multiple micronutrient group corresponded with non-significant reductions in the risk of preterm delivery (<37 weeks) and early preterm delivery (<34 weeks) of 14% (P=0.29) and 30% (P=0.26), respectively.

The intracluster correlation coefficient for birth length was 0.03 (0.007 to 0.047) and for head circumference was 0.08 (0.050 to 0.101). In the iron-folic acid group there was a significant increase in birth length (0.24 cm; 0.02 to 0.46 cm, P=0.03) compared with folic acid alone. There was no significant effect of iron-folic acid on mean head circumference (P=0.21). There were no significant effects of the multiple micronutrients on either mean birth length (P=0.12) or mean head circumference (P=0.80).

Compared with folic acid alone, the difference in mean birth weight, adjusted for multiple births, gestation at delivery, and cluster randomisation, was 12.6 g for the iron-folic acid and 31.0 g for the multiple micronutrients, indicating that the extension of the duration of gestation contributed to a similar increment in mean birth weight for iron-folic acid (11.7 g) and multiple micronutrients (11.3 g) compared with the folic acid.

Impact of nutrient supplementation on maternal haemoglobin concentration

Haemoglobin samples were available for 599 women. The baseline characteristics for these women were balanced across the treatment groups, and there were no significant differences between the women with and without haemoglobin measurements. There was a significant increase in haemoglobin concentration with both iron-folic acid and multiple micronutrients compared with folic acid (P=0.001 and P<0.001, respectively). In both the iron-folic acid and the multiple micronutrient groups, however, more than 40% of the women were still anaemic in the third trimester.

Impact of supplementation on perinatal mortality

Post hoc analyses showed a reduction in the risk of early neonatal mortality among infants born to women randomised to receive either iron-folic acid or multiple micronutrients compared with folic acid. The reduction was significant for the iron-folic acid group compared with folic acid but not for the multiple micronutrients group (table).

The figure illustrates cumulative mortality curves, with lower mortality rates for neonates whose mothers received iron-folic acid and multiple micronutrients. The overall differences were of borderline significance (P=0.055), but the difference between iron-folic acid and folic acid alone was significant (P=0.032). The other individual treatment comparisons (multiple micronutrients *v* folic acid and multiple micronutrients *v* iron-folic acid) were not significant (P=0.077 and P=0.708, respectively).

On the basis of our results, the number of women who would need to be treated with iron-folic acid from early in pregnancy would be 6.2 (6.1 to 6.3) to prevent one case of maternal anaemia in the third trimester, 83.3 (83.28 to 83.32) to prevent one preterm delivery <37 weeks, 122 (121.99 to 122.01) to prevent one early preterm delivery <34 weeks, and 125 (124.99 to 125.01) to prevent one early neonatal death. With multiple micronutrients the numbers needed to treat would be 5.6 (5.5 to 5.7), 111.1 (111.08 to 111.12), 188.7 (188.69 to 188.71), and 222 (222.19 to 222.21), respectively.

Mortality outcomes (in singleton births)

	Folic acid		Iron-folic acid		Multiple micronutrients		Relative risk (95% confidence interval)*	
	No in group	Rate/ 1000	No in group	Rate/ 1000	No in group	Rate/ 1000	Iron-folic acid *v* folic acid	Multiple micronutrients *v* folic acid
Pregnancies with single live or stillbirth:								
Total	1688	—	1546	—	1532	—	—	—
Stillbirths (≥ 28 weeks)	52	30.8	47	30.4	63	41.1	1.01 (0.67 to 1.51)	1.39 (0.95 to 2.04)
Live births	1636	—	1499	—	1469	—	—	—
All neonatal deaths	33	20.2	16	10.7	18	12.3	0.53 (0.29 to 0.97)	0.61 (0.34 to 1.10)
Early neonatal deaths	24	14.7	10	6.7	15	10.2	0.46 (0.21 to 0.98)	0.71 (0.36 to 1.39)
Perinatal deaths	76	45.0	57	36.9	78	50.9	0.84 (0.59 to 1.19)	1.18 (0.85 to 1.63)

*Adjusted for cluster randomisation with general estimating equation binomial model.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Neonatal deaths remain a major cause of child mortality in those aged under 5 in developing countries and are associated with low birth weight

One of the major causes of low birth weight in developing countries is the poor nutritional status of the mother before and during pregnancy

WHAT THIS STUDY ADDS

Antenatal iron-folic acid increases the duration of gestation, reduces early preterm delivery, and is associated with a significant reduction in early neonatal mortality compared with folic acid alone

Antenatal multiple micronutrients modestly increase birth weight compared with folic acid but, despite this weight gain, are not associated with reduced early neonatal mortality

An adequate dose of iron is required in micronutrient supplements for pregnant women to maximise reductions in neonatal mortality in developing countries

DISCUSSION

In this trial the response to nutrient supplementation in pregnancy varied by outcome. Overall, we have shown that antenatal nutrient supplementation is associated with increased maternal haemoglobin concentration in late pregnancy, mean birth weight, and duration of pregnancy and reduced preterm delivery and early neonatal mortality.

The effects of iron-folic acid and multiple micronutrients on the duration of gestation and neonatal mortality seem to be related to the iron in these supplements. The largest impact on neonatal mortality was with the supplement with the highest dose of iron, and we saw a similar pattern for duration of gestation and effects on preterm delivery. A large scale community based trial from Indonesia found no significant differences for neonatal mortality between daily antenatal supplements with iron-folic acid and multiple micronutrients, which both had the same dose of iron (30 mg).[5] These findings from Indonesia suggest that the higher neonatal mortality rate we observed in China with multiple micronutrients compared with iron-folic acid cannot be explained by the additional micronutrients but by the different dose of iron in the supplements.

Evidence of an important role of antenatal iron supplements in reducing preterm delivery has been reported from the United States.[6] Longitudinal studies in Brazil[7] and Bangladesh[8] indicate that preterm infants have much higher neonatal mortality than babies with retarded intrauterine growth or those born at term. These observational findings suggest a plausible reason as to why the increased mean duration of gestation and reduced preterm delivery (especially early preterm delivery) we observed could be related to a large reduction in neonatal mortality.

Study strengths and limitations

Even though randomisation was by village clusters, there were a large number of clusters (531) with about 180 per treatment group and a relatively small number of births in each cluster. Population recruitment of pregnancies allowed complete tracing of fetal losses, and pregnancy outcomes in both hospital and home deliveries. Only 2.3% of the enrolled women were lost to follow-up, and 94% of the liveborn infants were weighed within 72 hours of delivery. There was no evidence of under-enumeration of neonatal deaths.

A limitation was the low power to detect changes in the prevalence of low birth weight, preterm delivery, and perinatal mortality. The mortality analyses were post hoc rather than a primary hypothesis and we did not have enough data to fully examine the effects of the different supplements on mortality.

Comparison with other studies

Our results help to explain the apparent increase in neonatal and perinatal mortality reported from other trials using the same or similar formulation of multiple micronutrient supplements compared with standard iron-folic acid supplements.[9-11] Our findings showed a significant reduction in neonatal mortality for iron-folic acid (47%) and a non-significant reduction for multiple micronutrients (39%) compared with folic acid. The greater effect on mortality of iron-folic acid compared with multiple micronutrients could account for why comparisons of multiple micronutrients with iron-folic acid give the appearance of an increased risk of neonatal mortality.

There are similarities between our results and the results of other earlier controlled trials of iron supplementation in pregnancy.[12 13] Our findings were also consistent with those from observational studies that reported an increased risk of preterm delivery in women with iron deficiency anaemia in the first trimester of pregnancy.[14-18]

We observed a larger birth weight response in the multiple micronutrients group than in the iron-folic acid group. The total increase in birth weight in our multiple micronutrients group compared with the iron-folic acid group was less than that reported in other studies.[10 19 20] This did not translate into lower neonatal mortality, although it implies improved fetal growth and development, which might confer other health benefits to the newborn.

An adequate dose of iron is required in micronutrient supplements for pregnant women in developing countries to maximise reductions in neonatal mortality.

We thank all the women and their families who participated in the trial. We also thank the following members of the Shaanxi study team: Kang Yijun, Fang Bo and Liang Weifeng (field supervisors); and the area coordinators of the Maternal and Child Healthcare Stations and Health Bureaus in the study counties. Ray Yip, Mary Cogswell and Zuoguo Mei provided useful comments about the study design and field procedures. Li Qiang helped with data management. Wang Quanli, Xiao Shengbin, Xing Yuan, Zhou Xiaoyan, Wang Bei, and Xie Hong helped with data management. We also thank David Hipgrave, Sandy Huffman, Rachel Huxley, and Robyn Norton for valuable comments about the draft manuscript.

Contributors: See bmj.com.

Funding: This work was supported by the United Nations Children's Fund (grant No YH101-H12/03) through a cooperative agreement between Unicef and the Centers for Disease Control and Prevention, Atlanta, US, and the National Natural Science of Foundation of China (grant No 30271131), Beijing, China. Sight and Life and DSM Nutritional Products, China, provided the nutrient supplement, which was manufactured by Beijing Vita Nutritious Products, Beijing, China.

Competing interests: MJD was a consultant for Unicef China Unicef Pyongyang during the conduct of the trial. SC was a nutrition consultant for

Unicef China from 2001 to 2002, and is now the liaison officer for Unicef with the Ministry of Health.

Ethical approval: This study was approved by the Human Research Ethics Committee of the College of Medicine, Xi'an Jiaotong University (No 2002001) and the Ministry of Health, China. All women gave informed verbal consent.

Provenance and peer review: Not commissioned; externally peer reviewed.

1 Kramer M, Victora C. Low birth weight and perinatal mortality. In: Semba RD, Bloem MW, eds. *Nutrition and health in developing countries*. Clifton, NJ: Humana Press, 2001.
2 Huffman S, Baker J, Shumann J, Zehner E. *The case for promoting multiple vitamin/mineral supplements for women of reproductive age in developing countries*. Washington: Academy for Educational Development, 1998.
3 Berry RJ, Li Z, Erickson JD, Li S, Moore CA, Wang H, et al. Prevention of neural-tube defects with folic acid in China. China-US collaborative project for neural tube defect prevention. *N Engl J Med* 1999;341:1485-90.
4 Unicef. *Composition of a multi-micronutrient supplement to be used in pilot programs among pregnant women in developing countries. Unicef/WHO/UNU Workshop*. New York: Unicef, 1999.
5 Shankar AH, Jahari AB, Sebayang SK, Aditiawarman, Apriatni M, Harefa B, et al. Effect of maternal multiple micronutrient supplementation on fetal loss and infant death in Indonesia: a double-blind cluster-randomised trial. *Lancet* 2008;371:215-27.
6 Siega-Riz AM, Hartzema AG, Turnbull C, Thorp J, McDonald T, Cogswell ME. The effects of prophylactic iron given in prenatal supplements on iron status and birth outcomes: a randomized controlled trial. *Am J Obstet Gynecol* 2006;194:512-9.
7 Barros FC, Huttly SR, Victora CG, Kirkwood BR, Vaughan JP. Comparison of the causes and consequences of prematurity and intrauterine growth retardation: a longitudinal study in southern Brazil. *Pediatrics* 1992;90:238-44.
8 Yasmin S, Osrin D, Paul E, Costello A. Neonatal mortality of low-birth-weight infants in Bangladesh. *Bull World Health Organ* 2001;79:608-14.
9 Christian P, West KP, Khatry SK, Leclerq SC, Pradhan EK, Katz J, et al. Effects of maternal micronutrient supplementation on fetal loss and infant mortality: a cluster-randomized trial in Nepal. *Am J Clin Nutr* 2003;78:1194-202.
10 Osrin D, Vaidya A, Shrestha Y, Baniya RB, Manandhar DS, Adhikari RK, et al. Effects of antenatal multiple micronutrient supplementation on birthweight and gestational duration in Nepal: double-blind, randomised controlled trial. *Lancet* 2005;365:955-62.
11 Christian P, Osrin D, Manandhar DS, Khatry SK, de L Costello AM, West KP Jr. Antenatal micronutrient supplements in Nepal. *Lancet* 2005;366:711-2.
12 Preziosi P, Prual A, Galan P, Daouda H, Boureima H, Hercberg S. Effect of iron supplementation on the iron status of pregnant women: consequences for newborns. *Am J Clin Nutr* 1997;66:1178-82.
13 Cogswell ME, Parvanta I, Ickes L, Yip R, Brittenham GM. Iron supplementation during pregnancy, anemia, and birth weight: a randomized controlled trial. *Am J Clin Nutr* 2003;78:773-81.
14 Scanlon KS, Yip R, Schieve LA, Cogswell ME. High and low hemoglobin levels during pregnancy: differential risks for preterm birth and small for gestational age. *Obstet Gynecol* 2000;96:741-8.
15 Scholl TO, Reilly T. Anemia, iron and pregnancy outcome. *J Nutr* 2000;130(2 suppl):443-7S.
16 Scholl TO, Hediger ML, Schall JI. Questions about adolescent gestational age and birth weight. *Pediatrics* 1994;94:243-4.
17 Zhou LM, Yang WW, Hua JZ, Deng CQ, Tao X, Stoltzfus RJ. Relation of hemoglobin measured at different times in pregnancy to preterm birth and low birth weight in Shanghai, China. *Am J Epidemiol* 1998;148:998-1006.
18 Scholl TO, Hediger ML, Fischer RL, Shearer JW. Anemia vs iron deficiency: increased risk of preterm delivery in a prospective study. *Am J Clin Nutr* 1992;55:985-8.
19 Kaestel P, Michaelsen KF, Aaby P, Friis H. Effects of prenatal multimicronutrient supplements on birth weight and perinatal mortality: a randomised, controlled trial in Guinea-Bissau. *Eur J Clin Nutr* 2005;59:1081-9.
20 Christian P, Khatry SK, Katz J, Pradhan EK, LeClerq SC, Shrestha SR, et al. Effects of alternative maternal micronutrient supplements on low birth weight in rural Nepal: double blind randomised community trial. *BMJ* 2003;326:571.

Accepted: 22 August 2008

Effect of social deprivation on blood pressure monitoring and control in England: a survey of data from the quality and outcomes framework

Mark Ashworth,[1] Jibby Medina,[2] Myfanwy Morgan[2]

EDITORIAL by Lester

[1]King's College London, Department of General Practice and Primary Care, Division of Health and Social Care Research, London SE11 6SP
[2]King's College London, Department of Public Health Sciences, Division of Health and Social Care Research, London SE1 3QD
Correspondence to:
Mark Ashworth
mark.ashworth@kcl.ac.uk

Cite this as: *BMJ* 2008;337:a2030
doi:10.1136/bmj.a2030

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2008;337:a2030

ABSTRACT

Objective To determine levels of blood pressure monitoring and control in primary care and to determine the effect of social deprivation on these levels.

Design Retrospective longitudinal survey, 2005 to 2007.

Setting General practices in England.

Participants Data obtained from 8515 practices (99.3% of all practices) in year 1, 8264 (98.3%) in year 2, and 8192 (97.8%) in year 3.

Main outcome measures Blood pressure indicators and chronic disease prevalence estimates contained within the UK quality and outcomes framework; social deprivation scores for each practice, ethnicity data obtained from the 2001 national census; general practice characteristics.

Results In 2005, 82.3% of adults (n=52.8m) had an up to date blood pressure recording; by 2007, this proportion had risen to 88.3% (n=53.2m). Initially, there was a 1.7% gap between mean blood pressure recording levels in practices located in the least deprived fifth of communities compared with the most deprived fifth, but, three years later, this gap had narrowed to 0.2%. Achievement of target blood pressure levels in 2005 for practices located in the least deprived communities ranged from 71.0% (95% CI 70.4% to 71.6%) for diabetes to 85.1% (84.7% to 85.6%) for coronary heart disease; practices in the most deprived communities achieved 68.9% (68.4% to 69.5%) and 81.8% (81.3% to 82.3%) respectively. Three years later, target achievement in the least deprived practices had risen to 78.6% (78.1% to 79.1%) and 89.4% (89.1% to 89.7%) respectively. Target achievement in the most deprived practices rose similarly, to 79.2% (78.8% to 79.6%) and 88.4% (88.2% to 88.7%) respectively. Similar changes were observed for the achievement of blood pressure targets in hypertension, cerebrovascular disease, and chronic kidney disease.

Conclusions Since the reporting of performance indicators for primary care and the incorporation of pay for performance in 2004, blood pressure monitoring and control have improved substantially. Improvements in achievement have been accompanied by the near disappearance of the achievement gap between least and most deprived areas.

Table 1 | Characteristics of general practices in the least deprived and most deprived fifths of the "super output areas"* in England during the first three years of the quality and outcomes framework

Characteristic	Least deprived areas			Most deprived areas		
	Year 2004-5	Year 2005-6	Year 2006-7	Year 2004-5	Year 2005-6	Year 2006-7
No of practices	1096	1228	1170	2391	2373	2225
No of full time equivalent general practitioners	3974	4660	5617	6036	6342	7427
Mean list size per full time equivalent general practitioners	2098	2063	1694	2285	2284	1982
% of practices that were single handed	14.4	13.4	8.9	36.2	35.1	28
% of practices that were training practices	38.0	38.2	38.2	19.7	19.1	19.8

*See text and National Statistics[5] for explanation of super output areas.

INTRODUCTION

In 2004 the contract for all UK general practitioners was revised to include a pay for performance system accounting for up to a quarter of total annual income. This system, the quality and outcomes framework (QOF), has provided new performance data aggregated at practice level.[1] Achievement of blood pressure recording for all registered patients aged ≥45 years is monitored annually. So too is the achievement of blood pressure target levels for all patients with any of five chronic conditions included in the QOF—hypertension, coronary heart disease, stroke and transient ischaemic attacks, diabetes, and chronic kidney disease.

The linkage of financial incentives to performance targets was intended to drive up the standards of primary care. It has also been observed that health inequalities between different population groups may be diminished as overall collective achievement increases. This phenomenon has been termed the inverse equity hypothesis.[2] In general terms, this hypothesis predicts that new public health interventions initially reach the wealthier sector of the population and then later begin to benefit the poor. In consequence, inequity ratios are initially increased, only to be diminished once the poor gain access to the intervention and a ceiling effect is reached in the richer population.

We used the data available through the QOF to provide a current perspective on two measures—national rates of blood pressure monitoring in primary care patients and the achievement of blood pressure control targets in patients with chronic conditions. We aimed to describe changes in these two measures over the three years since QOF data have been available. In addition, we describe the effect of any changes over time on health inequalities between general practices in deprived and less deprived communities.

METHODS

Quality and outcomes framework data

We obtained QOF data covering the three years from April 2004 to March 2007 for all general practices in England. Data for the preceding year were collected in February (prevalence data) and March (performance indicator data) of years 2005, 2006, and 2007.[3] We confined our analysis to the following:

- The proportion of all registered patients aged ≥45 years who had had their blood pressure taken within the preceding five years
- The prevalence of five chronic conditions (hypertension, coronary heart disease, stroke and transient ischaemic attacks, diabetes, and chronic kidney disease) and the achievement of blood pressure targets for each of these conditions.

The denominator for the analysis of all QOF data was the total population of patients on the respective disease registers for each practice. This method of analysis precludes the process of "exception reporting," whereby general practitioners are permitted to omit certain patients from their performance data on the grounds of unsuitability. For research purposes, exception reporting introduces bias, since different general practitioners have different exception reporting thresholds.[4] All data presented in this study are the values reported by GPs before they had excluded any patients using the mechanism of exception reporting.

Practice and population characteristics

We obtained details of practice characteristics including practice list size, a breakdown of the registered population by age and sex, the number of full time equivalent general practitioners, and training practice status.

We obtained data from the 2001 national UK census based on the lower layer super output area (SOA) for each practice.[5] Each such area consists of about 1500 people within a defined geographical locality. We used the super output area in which a general practice was located as a proxy for the registered population at each practice. The super output area forms the basis for calculating the index of multiple deprivation. We added a measure of ethnicity from the 2001 national census to our own dataset, again aggregated at the level of super output area.

Statistical methods

We constructed a dataset for all general practices in England containing data from the QOF, the practice,

and census data for the surrounding super output area. After exclusions (see bmj.com), the final dataset consisted of 8515 practices (99.3% of the total) in year 1, with 8480 of these practices linked to the 2004 index of multiple deprivation; 8264 practices (98.3% of the total) in year 2, all linked to the deprivation index; and 8192 (97.8% of the total) in year 3, with 7831 linked to the deprivation index.

Firstly, we explored the relationship between social deprivation and the achievement of the six QOF indicators related to blood pressure by comparing the achievement of practices located in the most deprived fifth of super output areas in the country with the achievement of practices in the least deprived fifth. We calculated mean values for each of the blood pressure related variables in all practices and then recalculated them as weighted means (based on the number of registered patients) in order to adjust for the effect of practice size. We then searched for possible confounding variables using multivariate analysis.

RESULTS

Practice characteristics

Table 1 shows the characteristics of the practices.

Blood pressure recording in the adult population

We looked at the proportions of all registered patients aged ≥45 years with a record of a blood pressure reading in the preceding five years. We found the small discrepancy between values for blood pressure recording in the least and most deprived fifths of super output areas in 2005 was dwarfed by the overall improvement in values over the three years. By 2007, the difference between the general practices in the least and most deprived areas had all but disappeared. See bmj.com.

When we analysed data for all practices in England, the mean number of patients with an up to date blood pressure recording in 2007 (year 3 of the QOF) was 88.3% (n=53.21 million); after weighting for practice list size, the mean was 87.9%. Comparable values for 2005 were 81.6% and 82.3% (n=52.78 million), respectively.

Achievement of blood pressure targets for five chronic conditions

Table 2 summarises the success of the practices in the least and most deprived areas at achieving the blood pressure targets incorporated into the QOF. Modest shortfalls in blood pressure control by practices in more deprived areas have largely disappeared by the third year of the QOF, even though the small residual differences were significant.

Association between blood pressure indicators, social deprivation, and possible confounding variables

We constructed a series of regression models to explore the role of possible confounding variables, and 6.5% of the variation in the proportion of adult patients with up to date blood pressure monitoring was explained by four factors: practices found to be performing less well on multivariate analysis (in 2007) were those with higher proportions of black or black British residents in the local population, were situated in less deprived areas, had larger numbers of general practitioners, and had larger list sizes per general practitioner. See bmj.com. The training status of the practice, its singlehanded status, and disease prevalence values were not significant predictor variables.

DISCUSSION

Main findings of this study

Blood pressure monitoring levels of all adult patients (aged ≥45 years) registered at general practices in England have risen by 5% over the three years during which QOF data have been available such that, by 2007, 88% of the adult population have had their blood pressure measured in the preceding five years.

When blood pressure monitoring data first became available through the QOF dataset, there was a difference between most and least deprived areas. However, as overall blood pressure monitoring levels increased, this inequality has narrowed such that, by 2007, the difference was negligible (0.2%).

In the third year of QOF, social deprivation was no longer having a negative effect on blood pressure

Table 2 | Achievement of blood pressure targets set in the quality and outcomes framework for five chronic conditions among adult patients (≥45 years) registered with general practices in the least deprived and most deprived fifths of the "super output areas"* in England. Values are mean (95% CI) percentages of patients at each practice whose blood pressure was within the target value†

	Least deprived areas			Most deprived areas		
Condition	Year 2004-5	Year 2005-6	Year 2006-7	Year 2004-5	Year 2005-6	Year 2006-7
Hypertension	72.4 (71.9 to 72.9)	75.9 (75.5 to 76.3)	78.0 (77.7 to 78.4)	69.1 (68.6 to 69.6)	74.7 (74.3 to 75.1)	77.4 (77.1 to 77.7)
Coronary heart disease	85.1 (84.7 to 85.6)	87.5 (87.2 to 87.9)	89.4 (89.1 to 89.7)	81.8 (81.3 to 82.3)	86.1 (85.8 to 86.4)	88.4 (88.2 to 88.7)
Stroke and transient ischaemic attacks	82.8 (82.2 to 83.3)	85.5 (85.0 to 85.9)	87.5 (87.2 to 87.9)	78.5 (77.9 to 79.1)	83.7 (83.3 to 84.1)	86.2 (85.8 to 86.5)
Diabetes	71.0 (70.4 to 71.6)	74.9 (74.4 to 75.4)	78.6 (78.1 to 79.1)	68.9 (68.4 to 69.5)	74.5 (74.1 to 74.9)	79.2 (78.8 to 79.6)
Chronic kidney disease‡	—	—	89.9 (89.2 to 90.7)	—	—	88.7 (88.1 to 89.4)

*See text and National Statistics[5] for explanation of super output areas.
†Target blood pressure values were ≤150/90 mm Hg for hypertension, coronary heart disease, and stroke and transient ischaemic attacks; ≤145/85 mm Hg for diabetes; and ≤140/85 mm Hg for chronic kidney disease.
‡Chronic kidney disease data were included in the quality and outcomes framework only in 2006-7.

WHAT IS ALREADY KNOWN ON THIS TOPIC

High blood pressure is the single most powerful risk factor for the development of cardiovascular disease

Successful control of blood pressure has the potential to reduce greatly the incidence of cardiovascular disease but is less likely to be achieved in more socially deprived areas

WHAT THIS STUDY ADDS

Levels of blood pressure monitoring and control in England have increased substantially over the years 2005 to 2007

This increase has been accompanied by narrowing of the small but real social gradients in blood pressure monitoring and control, which were found to have improved more rapidly in more deprived areas

The introduction of a series of primary care performance indicators related to blood pressure control may have contributed to this success

monitoring and, once corrected for confounding, had a weakly positive effect. The strongest confounding effect (based on the standardised regression coefficients) was the proportion of black or black British people in the local community.

For the five chronic conditions included in our study, the greatest increase over time was observed for blood pressure control in diabetics in the most deprived practices.

The improvements in achievement of blood pressure targets in chronic conditions have resulted in almost complete disappearance of the differences between least and most deprived areas. We conducted multivariate analysis and constructed models for each of the diseases, but these had low predictive power and added little to the univariate findings. In particular, ethnicity did not emerge as a significant confounding covariable for blood pressure control. This suggests that ethnicity played a role in determining the level of blood pressure monitoring by general practitioners but not in the control of high blood pressure.

What is already known on this topic?

Data from the health survey for England have shown steady improvements in population blood pressure control over the decade 1994-2003.[6] However, our own results suggest that rapid improvement has occurred in the past three years and that overall population levels of blood pressure control are now likely to be better than those derived from the 2003 health survey for England. Our findings also add to the observations of others about the inverse equity hypothesis.

Limitations of this study

Several general limitations affect the interpretation of QOF as it was not designed as a research tool. The values for successful blood pressure control obtained in our study are much higher than previously reported from population surveys in the UK.[6] This may be attributable to rapid recent improvements or because the practice disease register is not validated.

Our data do not establish a causative relation because the observed association between lower blood pressure screening rates in practices located in areas with higher black populations as there were no direct ethnicity data for the people attending general practices. Others have noted that blood pressure control is poorer in hypertensive patients from an ethnic minority.[7 8]

Data verification methods differ between localities, some practices will be the subject of far greater scrutiny than others, and new audit procedures have been introduced since the early years of QOF. Prevalence data available in the QOF database are not standardised for age and sex. Individual patient data are not available through QOF, and many patients appear on more than one of the disease registers at the same time.

General practice postcodes were used as a proxy for the postcodes of registered patients at each practice. Although this makes our data susceptible to the ecological fallacy, others have found that deprivation scores linked to practice postcodes did provide a valid proxy for patient-level deprivation measures.[9]

Only the final blood pressure reading recorded during the QOF year is included in the QOF dataset, and blood pressure recording can be analysed only according to the timeframes of QOF itself. Finally blood pressure monitoring can be analysed only to determine if one or more blood pressure readings have been recorded over the preceding five years, and detailed analysis of more recent data is not possible.

We thank Stevo Durbaba for help in preparing the database and Roger Jones for helpful advice on the manuscript.

Contributors: See bmj.com

Funding: JM was supported by a grant from the King's College Centre for Caribbean Health.

Competing interests: None declared.

Ethical approval: Guy's Research Ethics Committee approved the study.

Provenance and peer review: Not commissioned; externally peer reviewed.

1 UK Department of Health. *Investing in general practice: the new GMS contract.* London: DoH, 2003.
2 Victora CG, Vaughan P, Barros F, Silva A, Tomasi E. Explaining trends in inequities: evidence from Brazilian child health studies. *Lancet* 2000;356:1093-8.
3 The NHS Information Centre for health and social care. http://www.ic.nhs.uk/statistics-and-data-collections/audits-and-performance/the-quality-and-outcomes-framework
4 Doran T, Fullwood C, Gravelle H, Reeves D, Kontopantelis E, Hiroeh U, et al. Pay-for-performance programs in family practices in the United Kingdom. *N Engl J Med* 2006;355:375-84.
5 National Statistics. Beginners' guide to UK geography: Super Output Areas (SOAs). www.statistics.gov.uk/geography/soa.asp
6 Primatesta P, Poulter NR. Improvement in hypertension management in England: results from the health survey for England 2003. *J Hypertens* 2006;24:1187-92.
7 Cappuccio FP, Cook DG, Atkinson RW, Strazzullo P. Prevalence, detection, and management of cardiovascular risk factors in different ethnic groups in south London. *Heart* 1997;78:555-63.
8 Connell P, McKevitt C, Wolfe C. Strategies to manage hypertension: a qualitative study with black Caribbean patients. *Br J Gen Pract* 2005;55:357-61.
9 Strong M, Maheswaran R, Radford J. Socioeconomic deprivation, coronary heart disease prevalence and quality of care: a practice-level analysis in Rotherham using data from the new UK general practitioner quality and outcomes framework. *J Pub Health* 2006;28:39-42.

Accepted: 31 August 2008

Patient safety indicators for England from hospital administrative data: case-control analysis and comparison with US data

Veena S Raleigh,[1,2] Jeremy Cooper,[1] Stephen A Bremner,[1] Sarah Scobie[3]

[1]Healthcare Commission, London EC1Y 8TG
[2]University of Surrey, Guildford, Surrey GU2 7WG
[3]National Patient Safety Agency, London W1T 5HD
Correspondence to: V S Raleigh veena.raleigh@healthcarecommission.org.uk

Cite this as: *BMJ* 2008;337:a1702
doi:10.1136/bmj.a1702

ABSTRACT

Objective To assess the feasibility of deriving patient safety indicators for England from routine hospital data and whether they can indicate adverse outcomes for patients.

Design Nine patient safety indicators developed by the United States Agency for Healthcare Research and Quality (AHRQ) were derived using hospital episode statistics for England for 2003-4, 2004-5, and 2005-6. A case-control analysis was undertaken to compare length of stay and mortality between cases (patients experiencing the particular safety event measured by an indicator) and controls matched for age, sex, health resource group (standard groupings of clinically similar treatments that use similar levels of healthcare resource), main specialty, and trust. Comparisons were undertaken with US data.

Setting All NHS trusts in England.

Participants Inpatients in NHS trusts.

Results There was fair consistency in national rates for the nine indicators across three years. For all but one indicator, hospital stays were longer in cases than in matched controls (range 0.2-17.1 days, P<0.001). Mortality in cases was also higher than in controls (5.7-27.1%, P<0.001), except for the obstetric trauma indicators. Excess length of stay and mortality in cases was greatest for postoperative hip fracture and sepsis. England's rates were lower than US rates for these indicators. Increased length of stay in cases was generally greater in England than in the US. Excess mortality was also higher in England than in the US, except for the obstetric trauma indicators where there were few deaths in both countries. Differences between England and the US in excess length of stay and mortality were most marked for postoperative hip fracture.

Conclusions Hospital administrative data provide a potentially useful low burden, low cost source of information on safety events. Indicators can be derived with English data and show that cases have poorer outcomes than matched controls. These data therefore have potential for monitoring safety events. Further validation, for example of individual cases, is needed and levels of event recording need to improve. Differences between England and the US might reflect differences in the depth of event coding and in health systems and patterns of healthcare provision.

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2008;337:a1702

INTRODUCTION

Safety of patients is an international problem: reviews of case notes have established that 4-16% of patients admitted to hospital experience an adverse event.[1-3] Measuring the scale and impact of safety incidents, however, is a major challenge, and estimates of deaths caused by such incidents vary widely.[4]

We examined the feasibility of deriving patient safety indicators from hospital episode data for England, whether the indicators point to adverse outcomes for patients, and how the results compare with data from the United States. We used a set of patient safety indicators designed to screen administrative data for events that indicate a potentially preventable problem of patient safety developed by the US Agency for Healthcare Research and Quality (AHRQ).[5]

METHODS

Selection of indicators

Of the 29 AHRQ patient safety indicators, we selected nine for analysis in this first phase. The choice of the nine indicators was informed by the following considerations: relative feasibility/complexity of coding conversion, potential reliability of coding in hospital episode statistics, and safety priorities for the Healthcare Commission (for example, maternity, infection control) (see bmj.com for full details).

Data used for analysis

We used hospital episode statistics for the financial years 2003-4, 2004-5, and 2005-6 for the analysis. The statistics comprise an administrative dataset of all NHS inpatients in England, covering about 13 million episodes of care annually. They contain demographic, administrative, and clinical (primary/secondary diagnoses, primary/secondary procedures, outcomes) details for every inpatient receiving NHS care.

The specifications of the US patient safety indicators use ICD-9 (international classification of diseases, ninth revision): each indicator is defined by specific numerator and denominator codes. The hospital episode statistics, however, are based on ICD-10 codes for diagnoses and Office of Population Censuses and Surveys (OPCS) codes for procedures. We translated the ICD-9 code specifications into ICD-10 and OPCS codes using semi-automated text word searches and manual coding, with the aim of obtaining the "best fit." Health resource groups are standard groupings of clinically similar treatments that use

similar levels of healthcare resource. We used health resource groups v3.5 in the analysis.

Statistical analysis

We calculated event rates at national level for each indicator and compared them across the three years with a view to testing the underlying suitability of hospital episode statistics for such analyses. We also analysed length of stay and mortality in cases (patients experiencing the particular safety event measured by an indicator) and matched controls (where such an event did not occur) for each indicator except death in low mortality health resource groups to establish whether or not the results indicated that an adverse event had occurred among cases. Each case was matched with up to four controls for age (within five years either side of case), sex, health resource groups (a derived measure of use of healthcare resources commonly used to adjust for case mix), main specialty, and trust. The case-control analysis was undertaken on only one year's data (2005-6) because we assumed that this would be adequate to test the hypothesis and because of the enormous scale of computation required to run a matched analysis on several million records.

Comparisons with US

We compared the event rates for England in 2005-6 with rates for the US in 2000.[6] We also compared the results for England for excess length of stay and mortality with US data from the same publication.

RESULTS

We analysed some 40 million episodes of inpatient care (details of the numbers of cases, numbers at risk, and rates per 1000 for the nine indicators across three years are on bmj.com). In 2005-6, the rates for the nine indicators ranged from 0.08 (postoperative hip fracture) to 60.34 (obstetric trauma—vaginal delivery with instrumentation) per 1000 discharges. The rates were fairly consistent over time for most indicators, showing little evidence of large or random variation between years.

Longer lengths of stay and higher mortality in cases compared with matched controls indicate that the measures are discriminatory and indicate the likely occurrence of a safety event. Excess length of stay and mortality is not applicable to the indicator on death in low mortality health resource groups because cases will have died during admission to hospital. The match rate for the remaining indicators was over 75%, except for postoperative sepsis (61%) and postoperative hip fracture (55%) (table 1). For all indicators except one (obstetric trauma—caesarean delivery), cases had significantly longer hospital stays than controls. Similarly, mortality in cases was significantly higher than in controls for most indicators; the exceptions (again as expected) were the obstetric trauma indicators, where there were no deaths in the matched set for two indicators, and one death for the third indicator.

For all indicators, the rates for England were lower than for the US, in most cases by a considerable margin (table 2).

For most indicators, the increased length of stay associated with cases was greater in England than in the US. Similar patterns were apparent for excess mortality, where levels in England were generally higher than in the US (see bmj.com). In both countries, the obstetric trauma indicators showed little or no differences in length of stay and mortality between cases and controls.

DISCUSSION

Our results suggest that the indicators have potential for monitoring patient safety events in the UK but require more in-depth validation of individual cases and better coding of events.

Limitations

Firstly, although widely used for analysing quality of care and clinical outcomes,[7] hospital episode statistics are primarily for administrative purposes, hence the

Table 1 | Indicators of patient safety: excess length of stay (days) and mortality in cases compared with matched controls, England (2005-6) and US (2000)

			Excess length of stay (days)				Excess mortality (percent)			
	Match rate %		England		US[6]		England		US[6]	
Indicator	England	US	Excess (SE)	P value	Excess (SE)	P value	Excess (SE)	P value	Excess (SE)	P value
Decubitus ulcer	77.2	56.0	9.14 (0.23)	<0.001	3.98 (0.10)	<0.001	13.42 (0.51)	<0.001	7.23 (0.23)	<0.001
Iatrogenic pneumothorax	80.5	66.0	4.35 (0.44)	<0.001	4.38 (0.24)	<0.001	10.59 (1.29)	<0.001	6.99 (0.73)	<0.001
Infections due to medical care	75.3	63.0	11.43 (0.37)	<0.001	9.58 (0.23)	<0.001	5.66 (0.64)	<0.001	4.31 (0.35)	<0.001
Postoperative hip fracture	55.1	51.0	17.09 (1.98)	<0.001	5.24 (0.69)	<0.001	18.20 (3.69)	<0.001	4.52 (1.34)	<0.001
Postoperative sepsis	60.5	33.0	15.90 (0.91)	<0.001	10.89 (0.90)	<0.001	27.07 (2.07)	<0.001	21.92 (1.47)	<0.001
Obstetric trauma:										
Vaginal with instrument	99.6	95.0	0.56 (0.04)	<0.001	0.07 (0.02)	<0.001	—*	NA	0.00	0.32
Vaginal without instrument	99.8	99.0	0.48 (0.02)	<0.001	0.05 (0.01)	<0.001	0.01 (0.01)	0.32	0.00	>0.99
Caesarean	99.2	99.0	0.20 (0.30)	0.49	0.43 (0.14)	0.003	—*	NA	−0.02 (0.02)	0.32

SE=standard error, NA=not applicable.
*No deaths in matched set.

depth of coding can be variable. While coding of procedures and primary diagnoses in the hospital episode statistics is fairly complete, coding of secondary diagnoses (used for several AHRQ indicators) is less complete, hence the adverse outcome rates are likely to be underestimated. In some cases, such as for postoperative sepsis, the number of events is lower than might be expected from clinical experience. This might indicate incomplete coding of events or the existence of alternative systems of recording certain events (for example, dedicated infection control systems) within hospital trusts. A new system of payment (payment by results) has been introduced in England, whereby tariffs are assigned on the basis of treatment and severity. On the basis of experience in other countries, this is likely to improve secondary coding and hence the potential utility of these indicators in the future.[8 9]

Secondly, the translation of ICD-9 diagnoses and procedure codes to ICD-10 diagnoses and OPCS procedure codes could have introduced inconsistencies with the original AHRQ specifications. We have, in consultation with others, refined the translation to capture the key coding requirements, but as thousands of detailed codes for each indicator need cross matching this remains work in progress. Furthermore, international initiatives are underway, including by the Organisation for Economic Cooperation and Development (OECD), to translate the ICD-9 codes for the AHRQ indicators to ICD-10 codes.[10] We are following these developments and will refine our coding accordingly (although the OPCS procedure codes used in hospital episode statistics are unique to England).

Thirdly, we did not attempt a cross validation of the hospital episode statistics results against patients' records or other sources of data. Although such comparisons would inevitably be costly, resource intensive, and limited in scale, such validation would be desirable to assess whether the cases identified by the indicators are confirmed patient safety events and will help to refine indicator definitions and support more appropriate use of the indicators on an on-going basis.[11]

Finally, there are caveats to the comparisons between England and the US, notably because of differences in healthcare systems and patterns of healthcare provision.

Table 2 | Indicators of patient safety: rates per 1000 events, England (2005-6) and US (2000)

Indicator	England			US[6] rate
	No of events	Population at risk	Rate	
Death in low mortality HRGs	2559	6 211 409	0.41	—
Decubitus ulcer	13 469	1 878 208	7.17	21.51
Iatrogenic pneumothorax	1313	10 662 750	0.12	0.67
Infections due to medical care	3895	3 666 667	1.06	1.99
Postoperative hip fracture	291	3 593 089	0.08	0.77
Postoperative sepsis	885	332 680	2.66	11.25
Obstetric trauma:				
Vaginal with instrument	3966	65 726	60.34	224.21
Vaginal without instrument	11 083	377 070	29.39	86.61
Caesarean	399	139 304	2.86	6.97

HRG=healthcare resource group.

WHAT IS ALREADY KNOWN ON THIS TOPIC

Hospital administrative data in the US have been used to derive the AHRQ patient safety indicators

They are used for monitoring national progress in improving the quality of health care, local quality improvement initiatives, and benchmarking of organisations

Several countries are now developing the application of these indicators to support safety monitoring

WHAT THIS STUDY ADDS

The evidence based AHRQ patient safety indicators can be reproduced for England by using routine hospital administrative data

Case-control analyses, showing longer lengths of stay and higher mortality in cases than matched controls, suggest that the indicators have potential for monitoring patient safety events, as in the US, though levels of event recording need to improve

Strengths of analysis

These caveats notwithstanding, our findings are important. Firstly, although different coding systems are used in England and the US, we were able to adapt a subset of the AHRQ indicators for use with hospital episode statistics, demonstrating their technical feasibility. Secondly, we established that although the indicators might underestimate event rates because of incomplete coding, they have potential as measures of patient safety events because cases are shown to have longer lengths of stay and higher mortality than controls for most indicators.

Thirdly, although our analysis identifies challenges for international comparisons based on these indicators, the AHRQ indicators are increasingly being developed internationally (see bmj.com). Finally, patient safety is a priority for NHS policy makers, commissioners, providers, and regulators. Although we have noted some caveats, we have, like others, also shown that routine hospital administrative data provide a pragmatic and cost effective alternative to bespoke reporting systems and case note reviews, which also don't allow for benchmarking across providers. Such data offer a resource that could potentially be used, alongside other local and national data sources, for improving completeness and quality of coding and monitoring trends in patient safety and local quality improvement initiatives.

There are challenges in developing and using safety indicators, especially in a policy environment

promoting publication of performance and quality measures, patients' choice, competition between providers, etc.[12] Use of these indicators as "performance measures" could, as with other safety measures, deter coding and reporting and impact negatively on the practical application and utility of these measures. If these risks can be managed by judicious use of the indicators, however, the potential to use routinely collected data for quality improvement will be enhanced for the benefit of patients.

We thank Richard Thomson, Adrian Cook, Jessica Chamberlain, Emma Hawe, and Ann Petruckevitch for their contribution to developing the analyses in the early stages; Frances Murphy for her expertise in translating ICD-9 to ICD-10 and OPCS procedure codes; and the Dr Foster Unit at Imperial College for collaborating on the translation of the indicators, which they did in parallel to us. There have been subsequent modifications to the codes, by them and us, which might account for some differences in rates.

Contributors: See bmj.com.

Funding: The Healthcare Commission received a small grant from the Health and Social Care Information Centre to support the initial recoding work.

Competing interests: None declared.

Ethical approval: Not required.

Provenance and peer review: Not commissioned; externally peer reviewed.

1 Brennan TA, Leape LL, Laird NM, Hebert L, Localio AR, Lawthers AG, et al. Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard medical practice study I. *N Engl J Med* 1991;324:370-6.

2 Wilson RM, Runciman WB, Gibberd RQ. The quality in Australian health care study. *Med J Aust* 1995;163:458-71.

3 Vincent C, Neale G, Woloshynowych M. Adverse events in British hospitals: preliminary retrospective record review. *BMJ* 2001;322:517-9.

4 National Audit Office *A safer place for patients: learning to improve patient safety*. London: NAO, 2005. www.nao.org.uk/publications/nao_reports/05-06/0506456.pdf

5 McDonald K, Romano P, Geppert J, Davies SM, Duncan BW, Shojania KG, et al. *Evidence report for measure of patient safety based on hospital administrative data*. Rockville, MD: Agency for Healthcare Research and Quality, 2002. (AHRQ Publication No 02-0038.) http://www.ahrq.gov/downloads/pub/evidence/pdf/psi/psi.pdf

6 Zhan C, Miller MR. Excess length of stay, charges and mortality attributable to medical injuries during hospitalisation. *JAMA* 2003;290:1868-74.

7 Bristol Royal Infirmary Inquiry. *Learning from Bristol: the report of the public inquiry into children's heart surgery at the Bristol Royal Infirmary 1984-1995. Chapter 21*.2001. (CM 5207.) www.bristol-inquiry.org.uk/final_report/report/sec2chap21_3.htm

8 Department of Health. *Payment by results guidance 2008/09*. London: DH, 2007. www.dh.gov.uk/en/Managingyourorganisation/Financeandplanning/NHSFinancialReforms/DH_081238

9 Serden L, Lindqvist R, Rosen M. Have DRG-based prospective payment systems influenced the number of secondary diagnoses in health care administrative data? *Health Policy* 2003;65:101-7.

10 Millar J, Mattke S. *Selecting indicators for patient safety at the health systems level in OECD countries*. Paris: OECD, 2004. (OECD Health Technical Papers No18.)

11 Polancich S, Restrepo E, Prosser J. Cautious use of administrative data for decubitus ulcer outcome reporting. *Am J Med Qual* 2006;21:262-8.

12 Department of Health. *High quality care for all: NHS next stage review final report*. London: DH, 2008. www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_085825

Accepted: 2 August 2008

A statistical outlier

The man who strode in lightly could not have been older than 65. He sat down smartly. Of slim build, he had a full head of hair and wore no glasses. I checked his birth date on the chart: 1923. He was actually in his 80s. His presenting complaint was relatively minor, and we agreed no investigation or treatment was required. That freed up time to discuss his youthful looks.

"I take it you don't smoke," I announced.

"No, as a matter of fact, I do. Pretty heavy all my life."

I nodded; the tar staining on his fingers was now obvious.

"Tell me about your family—many living long and healthy?"

"So so. Quite a lot of heart disease and stroke, actually."

This was going in entirely the wrong direction. I had one last try: "So what kind of person do you consider yourself? Content, no worries?"

"Maybe, but I tell you I've been a nervous sort all my life. Get upset over all silly things. My wife tells me to relax, and I just can't seem to."

I checked his age again as he left the clinic. No mistake. I sat back, puzzled and bemused, having just witnessed a statistical outlier. A man with a full house of risk factors prognosticating poor health, yet plainly defying them.

Evidence based medicine is the modern physician's mantra. The weakness of this approach is recognised in its inability to comment on an individual's prognosis from knowledge of the average. But there is another hazard in considering evidence based medicine as the only worthwhile route to clinical knowledge. We lose the opportunity to study those who bend the rules—the occasional "outliers" who never make it to the summary statistics. Yet the value of studying these cases can be exceptional. For example, it took only one instance (Scoville and Milner's famous patient HM[1]) to prove that bilateral removal of the hippocampus is disastrous. Epilepsy surgery of the temporal lobe since that day only ever contemplates a unilateral procedure; few would insist that a randomised trial would be required to prove the point.

Most physicians will agree that outstanding individual situations are just as worthy of intensive study as large scale trials with statistical outcomes. Could this single patient potentially offer enormous insight into ageing and atherogenesis? Almost certainly. Would this single patient be worth studying intensively? Absolutely. A practical problem remains though: how does one obtain a research grant that does not mention a P value somewhere in the Methods section?

Giridhar P Kalamangalam assistant professor, Department of Neurology, University of Texas Health Science Centre, Houston, TX, USA
giridhar.p.kalamangalam@uth.tmc.edu

Patient consent not required (patient anonymised, dead, or hypothetical).

1 Scoville WB, Milner B. Loss of recent memory after bilateral hippocampal lesions. *J Neurol Neurosurg Psychiatry* 1957;20:11-21.

Cite this as: *BMJ* 2008;337:a2305