



OPEN ACCESS

A randomised trial of granulocyte-macrophage colony-stimulating factor for neonatal sepsis: outcomes at 2 years

Neil Marlow,¹ Timothy Morris,² Peter Brocklehurst,³ Robert Carr,⁴ Frances M Cowan,⁵ Nishma Patel,⁶ Stavros Petrou,⁷ Maggie E Redshaw,⁸ Neena Modi,⁹ Caroline Dore²

¹Institute for Womens Health, University College London, London, UK

²MRC Clinical Trials unit, London, UK

³Institute for Women's Health, University College London, London, UK

⁴Department of Haematology, Kings College London, London, UK

⁵Department of Paediatrics, Imperial College (Hammersmith Hospital), London, UK

⁶National Perinatal Epidemiology Unit, University of Oxford, Oxford, UK

⁷CTU, University of Warwick, Coventry, UK

⁸Policy Research Unit in Maternal Health & Care, National Perinatal Epidemiology Unit, University of Oxford, Oxford, UK

⁹Department of Neonatal Medicine, Imperial College London, London, UK

Correspondence to

Neil Marlow, Institute for Womens Health, 74 Huntley Street, University College London WC1E 6AU, UK; n.marlow@ucl.ac.uk

Received 30 November 2011

Accepted 22 February 2012

Published Online First

27 April 2012

ABSTRACT

Objective The authors performed a randomised trial in very preterm small-for-gestational age (SGA) babies to determine if prophylaxis with granulocyte-macrophage colony-stimulating factor (GM-CSF) improves outcomes (the PROGRAMS trial). Despite increased neutrophil counts following GM-CSF, the authors reported no significant difference in neonatal sepsis-free survival.

Patients and methods 280 babies born <31 weeks of gestation and SGA were entered into the trial. Outcome was determined at 2 years to determine neurodevelopmental and general health outcomes, including economic costs.

Results The authors found no significant differences in health outcomes or health and social care costs between the trial groups. In the GM-CSF arm, 87 of 134 (65%) babies survived to 2 years without severe disability compared with 87 of 131 (66%) controls (RR: 1.0, 95% CI 0.8 to 1.2). Marginally, more children receiving GM-CSF were reported to have cough (RR 1.7, 95% CI 1.1 to 2.6) and had signs of chronic respiratory disease (Harrison's sulcus; RR 2.0, 95% CI 1.0 to 3.9) though this was not reflected in bronchodilator use or need for hospitalisation for respiratory disease. Overall, the rate of neurologic abnormality (7%–9%) was similar but mean overall developmental scores were lower than expected for gestational age.

Conclusions The administration of GM-CSF to very preterm SGA babies is not associated with improved or more adverse outcomes at 2 years of age. The apparent excess of developmental impairment in the entire PROGRAMS cohort, without corresponding increase in neurological abnormality, may represent diffuse brain injury attributable to intrauterine growth restriction.

INTRODUCTION

Systemic infection remains a major cause of mortality and morbidity in the newborn period. Of particular concern is the association between the inflammatory response and later risk of developmental delay and neurocognitive impairment, possibly mediated by damage to the periventricular white matter in the perinatal and neonatal periods.¹ Interventions aimed at reducing the impact of neonatal sepsis may therefore have longer term benefits in terms of developmental progress, a reduction in disability and consequent economic benefits.

Granulocyte-macrophage colony stimulating factor (GM-CSF) has been shown to be an effective treatment in neutropenia-related infections in patients with cancer after chemotherapy.^{2–3} Neutropenia is

What is known on this subject:

- ▶ Neonatal sepsis confers high mortality and morbidity after very preterm birth
- ▶ Very preterm small-for-gestational age (SGA) babies are at high risk of postnatal neutropenia and sepsis
- ▶ Postnatal GM-CSF administration increases neutrophil counts but does not reduce neonatal sepsis

What this study adds:

- ▶ At 2 years, prophylactic administration of GM-CSF does not improve neurodevelopmental or health outcomes
- ▶ This cohort of SGA babies had lower developmental scores than expected for gestation
- ▶ A detailed analysis of health and social care costs through 2 years of age

common in preterm growth-restricted infants who are at high risk of acquired infection after birth but a Cochrane review has suggested there was inadequate evidence for adoption of GM-CSF in neonatal practice.⁴ In order to resolve this matter, we undertook PROGRAMS, a single blind, multicentre, randomised trial of GM-CSF in very preterm small-for-gestational age (SGA) babies, to determine whether treatment resulted in reduced infections, mortality and morbidity in the neonatal period: 280 newborn SGA infants of 31 weeks gestational age or less were randomly allocated to GM-CSF or routine treatment within 72 h of birth. Although neutrophil counts were higher in GM-CSF-treated babies, there was no significant difference in sepsis or sepsis-free survival between the two treatment arms,⁵ in accordance with the findings of the systematic review.⁴

As part of the original design of the trial, we hypothesised that there might be more subtle benefits over and above short-term outcomes and designed an outcome evaluation at 2 years of age (adjusted for prematurity) to determine whether the administration of GM-CSF in the neonatal period produced differences in survival free of severe disability. Simultaneously, we undertook an

To cite: Marlow N, Morris T, Brocklehurst P, et al. *Arch Dis Child Fetal Neonatal Ed* 2013; **98**: F46–F53.

economic study to determine the cost benefit of the use of GM-CSF.

METHODS

Full details of the PROGRAMS trial have been reported earlier.⁶ Briefly, participants were infants born at a gestational age of ≤ 31 completed weeks of gestation with birthweight < 10 th centile (UK 1990 growth reference). An infant was not eligible if there was an immediately life-threatening congenital abnormality, or a strong likelihood of early-onset sepsis, indicated by maternal pyrexia exceeding 38°C on two occasions during labour. The study intervention, GM-CSF, in a dose of $10\ \mu\text{g}/\text{kg}$ was given subcutaneously daily for five consecutive days. No placebo injections were administered to the standard treatment arm of the study. Two commercial preparations of recombinant human GM-CSF were used during the study, molgramostim (Leucomax, Novartis, UK) and sargramostim (Leukine, Berlex, California, USA) which have equivalent biological potency for stimulating granulocyte production and function, both *in vitro* and *in vivo*.

TWO-YEAR OUTCOME EVALUATION

Contact was maintained with the families of the surviving children following their discharge from hospital. Children were traced and families contacted by the study coordinator. Paediatricians trained and validated in the outcome evaluation methods using a bespoke training course and video rating of developmental assessment techniques, evaluated each child in a clinical setting, usually in a hospital clinic room, but occasionally, at home. The assessment was carried out as close as possible to 2 years of age adjusted for prematurity (target $24\ \text{months} \pm 2\ \text{months}$) and comprised measures of growth, a formal clinical and neurological examination, developed and validated for use in infants born preterm⁶ and the Bayley Scales of Infant Development, second edition (BSID-II).⁷ In addition, parents completed a questionnaire with health and socioeconomic details and the Parent Report of Child Abilities (revised; PARCA-r).^{8,9} Data were recorded on standardised forms and collated centrally. Detailed feedback from the examination was posted to the parents to give to their general practitioner and their paediatrician (if their child was still under the care of a paediatrician).

Disability was classified as defined in 'Disability and Perinatal Care'¹⁰ but excluding malformation (a trial exclusion) and growth domains. Domains assessed were motor, vision, hearing, development and respiratory, renal and gastrointestinal function. Outcomes were classified as either severe disability (any one or more of BSID-II score $> 3\ \text{SD}$ below the mean or severe motor outcome (with inability to sit independently, use hands for feeding or control head, or severe hearing or visual impairment, or no meaningful words or signs; other disabilities (defined as BSID-2 scores -2 to $-3\ \text{SD}$, ambulant cerebral palsy (CP), and lesser degrees of sensory loss) or no disability (defined as none of the above and including children with BSID scores of $1-2\ \text{SD}$ below the mean)¹¹ Bayley standardised scores were classified in SD bands below $-1\ \text{SD}$ according to normative data provided by Bayley.⁷ Growth was measured using standardised equipment, including the Leicester height measure, standard weighing scales (Salter, model 918) and a Lass-o tape measure for head and arm circumference, and referred to UK Child Growth Foundation Standards¹² or WHO standards for mid upper arm circumference.¹³

An economic evaluation, conducted from a NHS and personal social services perspective,¹⁴ reported economic outcomes up to 2 years of age. Trial data collection forms, combined with

economic questionnaires completed by parents at 6-month intervals, provided a profile of all hospital inpatient and outpatient service use, surgeries performed, investigative tests, medications and community health and social care resource use. Unit costs (£, 2007–2008 prices) collected from primary and secondary sources in accordance with guidelines for costing healthcare services as part of economic evaluation¹⁵ were attached to each item of resource use. Cost effectiveness was expressed in two forms: incremental cost per additional severe disability-free survivor and incremental cost per additional disability-free survivor. The non-parametric bootstrap method was used to construct cost-effectiveness acceptability curves at alternative willingness to pay thresholds held by decision makers for the outcomes of interest.¹⁶

STATISTICAL METHODS

The sample size for the initial study was based on the short-term primary outcome, survival without sepsis for 14 days from trial entry.⁵

A CONSORT diagram was constructed, showing the flow of participants through the study.¹⁷ Variables were summarised as number (per cent) or median (25th centile–75th centile) for categorical or continuous/ranked data, respectively (none of the continuous variables approached approximate normality). For analysis of outcomes, RRs were used to quantify the effect of treatment on categorical variables. Median differences between treatment groups were calculated for continuous and ranked data; 95% CIs were calculated to quantify uncertainty about RRs and median differences.¹⁸

All analyses were done on an intention-to-treat basis, that is, participants were never excluded from analyses on the basis of the treatment received. The participants lost to follow-up or withdrawn were few in number and not included in analyses.

STUDY MANAGEMENT

The study was approved by the South Thames Multicentre Research Ethics Committee and by the site-specific review boards of the 26 recruiting centres. Written informed consent was obtained from parents. Trial conduct was overseen by a steering committee and independent Data-Monitoring Committee. This study is registered as an International Standard Randomised Controlled Trial, number ISRCTN42553489.

RESULTS

Of 280 babies enrolled into the study, 279 completed the study intervention, 64 babies died and 203 children (94% survivors) were evaluated at 2 years of age adjusted for prematurity (figure 1). Children were assessed at a median age of 27 months (range: 23–40 months); 64% of each group were seen within the target time window and the remainder were older at their assessment.

There were no systematic differences in study-entry characteristics or outcomes between the 13 infants (five received GM-CSF) who were not followed-up compared with those who were (web table S1). Children who received GM-CSF and controls who were followed-up were well balanced in respect of study-entry criteria, neonatal characteristics and short-term outcomes (table 1) and had similar socioeconomic profiles to the control group (web table S2).

Of the 134 children with known outcomes who received GM-CSF, 84 survived without severe disability compared with 87 of 131 control children (RR 1.0; 95% CI 0.8 to 1.2). The proportion of surviving children without severe disability in the GM-CSF group was 87 of 101 and 87 of 102 in the control

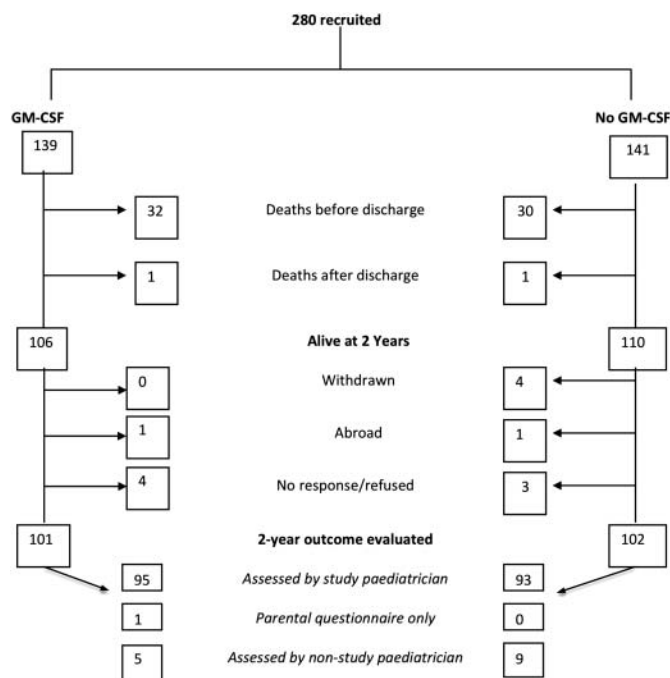


Figure 1 Flow diagram of children recruited to PROGRAMS.

group (RR 1.0; 95% CI 0.9 to 1.1); 14 of the GM-CSF group and 13 of the controls were assessed to have severe disability (RR 1.2; 0.5 to 2.7) and 24 and 19 respectively with other disabilities (RR 1.4; 0.7 to 2.7; table 2). Seven children (three non ambulant) in the GM-CSF group had a motor neurologic abnormality compared with nine controls (six non ambulant). Similar proportions had bilateral spastic CP (three each), dyskinesia (one each), unilateral spastic CP (one GM-CSF); two children who received GM-CSF had unclassifiable neurologic abnormalities compared to six controls (one child in the control arm had both bilateral spastic CP and dyskinesia). Sensory disability was infrequent; only one child had severe visual impairment (control) and none had severe hearing impairment, though three GM-CSF and six control children had no speech or signing at follow-up. There were no significant differences between the two groups in any outcome categories.

The median BSID-II mental development index (MDI) was 84 (IQR: 72–98) in the GM-CSF group and 87 (72–96) in controls and median psychomotor development indices (PDI) were 85 (73–96) and 88 (77–100), respectively. These differences were not statistically significant. There were similar proportions in each of the SD bands (table 3). Of those with severe or other disability, 5 of 12 from the GM-CSF group and 2 of 10 controls were classified as such solely on the basis of their developmental scores.

Parent report of child development using the PARCA-r questionnaire also showed no differences between groups. Growth in height and weight was similarly impaired in both groups (mean weight and height standard deviation scores (SDS) ranging from –1.6 to –1.8). Mean head circumference SDS was –2.1 for both groups (table 3).

Respiratory outcomes at 2 years are shown in table 4. Children who had received GM-CSF were more likely than controls to be reported to have cough (36% vs 22%), wheeze without infection (28% vs 20%) and chest deformity (22% vs 11%), predominantly Harrison's Sulcus indicating chronic respiratory distress though these differences were of marginal or no statistical

significance. Similar frequencies of hospital re-admissions after discharge home were reported in the two groups.

There were no significant differences in health and social care costs between the two trial arms (table 5). Mean costs during the 2-year follow-up period were £62 187 in the GM-CSF group and £66 260 in the control group, generating a mean cost difference of £4073 that was not statistically significant ($p=0.43$). GM-CSF is thus, on average, associated with marginal cost saving, but also marginally poorer outcomes in terms of survival without severe or other disability (negative incremental health effects). When the cost and outcomes data are combined within an incremental cost-effectiveness ratio (defined as incremental cost divided by incremental health effect), the mean incremental cost-effectiveness ratio appears in the southwest quadrant of the cost-effectiveness plane (web appendix figure A1). The mean incremental cost-effectiveness ratios within the PROGRAMS trial were £273 955 per additional severe disability-free survivor and £83 239 per additional disability-free survivor. In accordance with methodological guidance (NICE 2008), we also estimated the net monetary benefit of GM-CSF. The actual health benefits produced by GM-CSF were multiplied by alternative willingness to pay values for these benefits, and the net costs were subtracted. This produced a linear scale where a negative is unambiguously bad (the costs outweigh the value placed on the health benefits), and larger benefits are unambiguously better. Using this 'net-benefit' approach, we were able to assess the probability that GM-CSF is cost effective across alternative willingness to pay thresholds for the health outcomes of interest. The probability that GM-CSF is cost effective at 2 years never exceeded 0.79, which is below the level normally considered by decision makers as convincing evidence for cost effectiveness.¹⁶

DISCUSSION

We have carried out the first outcome evaluation at 2 years of age corrected for preterm birth for SGA babies randomly allocated to receive GM-CSF to prevent neonatal sepsis. In the treatment group, GM-CSF led to a significant increase in neutrophil count, but as previously reported,⁵ this did not result in reduced sepsis or increased survival. For the planned 2-year outcome evaluation, reported here, no developmental advantage was evident for the prophylaxis group and rates of disability, developmental score profiles and economic outcomes were very similar between the two groups, leading to the conclusion that GM-CSF is ineffective in improving outcomes to discharge or to 2 years of age.

We achieved a 94% follow-up rate with no systematic bias among dropouts in terms of original study variables. The high follow-up rate and well-validated testing environment mean that this study is unlikely to have missed a significant large beneficial effect despite the compelling reasons for undertaking the trial. The trial was halted for a period while the formulation of the GM-CSF was changed following a strategic decision to cease production by the concerned pharmaceutical firm, but there was no evidence of a differential effect between either of the two preparations of the compound. The outcome measures are clearly defined and well used in major epidemiological studies with good face validity.^{19 20} The Bayley (BSID-II) assessment is considered a robust measure of developmental progress and has been widely used in studies of preterm development and the results of parental report concur.

The PROGRAMS cohort is unique in representing extreme prematurity as well as SGA status. Reports of effects of fetal growth restriction on cognitive outcome are inconsistent²¹ and

Table 1 PROGRAMS: infant and maternal characteristics at trial entry and short term outcomes for all children followed-up at 2 years age corrected for prematurity

	GM-CSF (n=101)		Control (n=102)		Difference or RR	95% CI for difference or RR
	n or median	% or IQR	n or median	% or IQR		
Characteristics at trial entry						
Male, n (%)	46	46%	50	49%	0.9	0.7 to 1.2
Age at recruitment, h, median (IQR)	43	24–60	46	28–63	–4	–10 to 1
Gestational age, weeks, median (IQR)	29	27–30	29	27–30	0.1	–0.3 to 0.7
Birthweight, g, median (IQR)	839	668 to 932	788	630–928	19	–34 to 70
Neutrophil count at trial entry, $\times 10^9/l$, median (IQR)	2.7	1.3–4.1	2.6	1.3–4.3	0	–0.6 to 0.5
Neutropaenia ($<1.1 \times 10^9/l$) at trial entry, n (%)	16/97	16%	18/97	19%	0.9	0.5 to 1.6
Multiple pregnancies, n (%)	27	27%	22	22%	1.2	0.8 to 2.0
Maternal antenatal steroids administered 4 h or more before delivery, n (%)	80/87	92%	83/88	94%	1.0	0.9 to 1.1
Infants given surfactant within 4 h of birth, n (%)	70/79	89%	76/82	93%	1.0	0.9 to 1.1
Cranial ultrasound findings at trial entry						
Normal	76/100	76%	82/97	85%		
Minor abnormality	21/100	21%	12/97	12%		
Severe abnormality	3/100	3%	3/97	3%		
Short-term outcomes						
Necrotising enterocolitis (confirmed at surgery or post mortem)	2	2%	4	4%	0.5	0.1 to 2.7
Sepsis: culture positive to day 28 from trial entry	26	26%	22	22%	1.2	0.7 to 2.0
Sepsis: culture positive and probable to day 28	29	29%	34	33%	0.9	0.6 to 1.3
Cranial ultrasound findings						
Normal	62/88	70%	69/85	81%		
Minor abnormality	19/84	22%	9/85	11%		
Severe abnormality	7/84	8%	7/85	8%		
Oxygen dependency at day 28, n (%)	68	67%	57	56%	1.2	1.0 to 1.5
Oxygen dependency at 36 weeks PMA, n (%)	53	52%	44	43%	1.2	0.9 to 1.6
At discharge:						
Child exclusively breastfed	11/100	11%	12/100	12%	1.2*	0.5 to 2.8
Child partially breastfed	16/100	16%	19/100	19%	1.3*	0.6 to 2.6

For categorical variables summary measures are number, %, RR and 95% CI.

For continuous variables summary measures are median (IQR), difference (95% CI).

The denominator is the entire population, unless otherwise stated.

*RR ratio estimated by multinomial regression.

GM-CSF, granulocyte-macrophage colony-stimulating factor.

Table 2 Classification of outcome in PROGRAMS at 2 years of age corrected for prematurity

Outcome	GM-CSF (n=139)		Control (n=141)		GM-CSF – control	
	n	%	n	%	RR	95% CI
Death	33/134	24%	31/141	22%	1.1	0.7 to 1.7
Survival without severe disability†	87/134	65%	87/131	66%	1.0	0.8 to 1.2
Survival without disability‡	63/134	47%	68/131	52%	0.9	0.7 to 1.2
Overall function in survivor†						
Severe disability‡	14/101	14%	13/100	13%	1.2§	0.5 to 2.7
Other disability	24/101	24%	19/100	19%	1.4§	0.7 to 2.7
No disability	63/101	62%	68/100	68%	–	–
Neurological abnormality	7/101	7%	9/100	9%	0.8	0.3 to 2.0

* Unable to trace or abroad.

†One not done, two assessed by non-study paediatrician but insufficient information to classify disability.

‡Using criteria outlined in references 11 and 12. Malformation and growth domains have been excluded from the disability criteria.

§RR ratio estimated by multinomial logistic regression.

GM-CSF, granulocyte-macrophage colony-stimulating factor.

Table 3 Developmental and growth outcomes in children recruited to PROGRAMS at 2 years of age corrected for prematurity. The denominator indicates the number of children for whom parents returned questionnaires or who completed the Bayley Assessment

Outcome	GM-CSF		Control		GM-CSF – control		
	Med (IQR) or n%		Med (IQR) or n%		Median difference (95% CI)		
Bayley scales of infant development							
Mental development index							
	n=90		n=89				
Median (IQR) scores	84	(72–98)	87	(72–96)	0	(–4 to 5)	
Score <55	12	13%	10	11%			
Score 55–69	7	8%	6	7%			
Score 70–84	28	31%	27	30%			
Score >84	43	48%	46	52%			
Psychomotor development index							
	n=88		n=90				
Median (IQR) scores	85	(73–96)	88	(77–100)	–3	(–8 to 1)	
Score <55	9	10%	11	12%			
Score 55–69	11	13%	4	4%			
Score 70–84	24	27%	19	21%			
Score >84	44	50%	56	62%			
PARCA-r (parent report of child development)							
	(n=75)		(n=76)				
Development	26	22–29	26	23–29	0	–1 to 2	
Vocabulary	46	25–68	38	16–68	7	–3 to 17	
Sentence complexity	0	0–3	0	0–2	0	0 to 0	
Parent report composite	73	50–100	65	41–102	7	–5 to 19	
Parent report composite <49‡	17	24%	25	34%	0.7	0.4 to 1.2	
Growth							
		Mean	(SD)	Mean	(SD)	Difference in means (95% CI)	
Height	Cm	85	(5)	84	(4)		
	SDS*	–1.6	(1.4)	–1.7	(1.1)	0.1	(–0.3 to 0.5)
Weight	Kg	11	(1.8)	11	(1.3)		
	SDS	–1.7	(1.4)	–1.8	(1.2)	0.0	(–0.3 to 0.4)
Head circumference	Cm	48	(1.9)	47	(1.8)		
	SDS	–2.1	(1.5)	–2.1	(1.4)	0.1	(–0.3 to 0.5)
Midupper arm circumference	Cm	15	(1.5)	15	(1.2)		
	SDS†	–0.6	(0.8)	–0.8	(0.7)	0.1	(–0.1 to 0.4)
Abdominal circumference	Cm	47	(5.3)	46	(3.2)	0.85	(–0.5 to 2.2)

*Height, weight, head circumference standard deviation scores (SDS) scores from child growth foundation standards.

†Midupper arm circumference standard deviation scores from WHO Expert committee.

‡Parental questionnaire composite of non-verbal development, sentence complexity and vocabulary. Forty nine is the cut off for cognitive delay equivalent to bayley mental development index ≤ 70 .

GM-CSF, granulocyte-macrophage colony-stimulating factor.

there is no reference group for comparison. The proportion of children with disability and the pattern of clinical features in this particularly vulnerable group are thus of wider interest. In a population of extremely preterm babies, one would expect some depression of cognitive scores^{22–24} and this is reflected in our finding of an approximately 1 SD disadvantage in MDI and PDI scores over the normative population for the children in the PROGRAMS trial. The median gestational age of the PROGRAMS cohort was 29–30 weeks. However, their mean MDI and PDI scores (84–87 and 85–88, respectively) were of an equivalent level to those of an epidemiological cohort of babies born at a much greater degree of immaturity, 25 weeks or less, (the EPICure cohort; 84 and 87, respectively)¹⁹ and considerably lower than those of a contemporary population of babies born at 31 weeks or less gestational age and enrolled into a randomised trial of parental intervention (the PIP Study: 91–93 and 92–95, respectively).²⁵ In contrast, the rate of neurologic abnormality was, as expected, lower in the PROGRAMS group (7%–9%) compared with the EPICure cohort (24%).¹⁹ The same team trained assessors for the PROGRAMS, EPICure

and PIP studies and developmental assessment techniques were similar. We therefore speculate that the apparent excess in development impairment seen in the entire PROGRAMS cohort may reflect the vulnerability to global brain injury and developmental delay attributable to fetal growth restriction.

We might have anticipated improvement in neurodevelopment on several counts. First that neonatal sepsis is itself associated with worsened neurodevelopmental outcomes.^{26–27} Modulation of inflammatory mechanisms by GM-CSF might have improved outcomes, not evident during the neonatal period, through subtle alteration of the responses to infection. There is evidence that GM-CSF may be neuroprotective. GM-CSF crosses the blood–brain barrier and counteracts programmed cell death, and in experimental models of stroke was found to decrease damage.²⁸

The only significant difference between the trial groups was the marginally worse respiratory outcomes for the GM-CSF-treated group in terms of recurrent cough and chest deformity. This contrasts with worse short-term markers of neonatal lung disease in the control group. Twenty-three children in

Table 4 PROGRAMS: respiratory and hospital readmission outcomes from discharge from inpatient neonatal care through to 2 years of age corrected for prematurity. The denominator indicates the number of children for whom questionnaires were returned

Outcome	GM-CSF (n=100)	Control (n=101)	RR 95% CI
	n/total (%)	n/total (%)	
Chest symptoms			
Cough			
>once a week	10 (10)	16 (16)	
Weekly-monthly	12 (12)	0 (0)	
Once a month or less	30 (30)	30 (30)	
Cough without infection	36 (36)	22 (22)	1.7 (1.1 to 2.6)
Wheeze			
>once a week	5 (5)	6 (6)	
Weekly-monthly	5 (5)	5 (5)	
Once a month or less	22 (22)	19 (19)	
Wheeze without infection	28 (28)	20 (20)	1.4 (0.9 to 2.3)
Chest medicines			
Bronchodilators during last 12 months	22 (22)	26 (26)	0.9 (0.5 to 1.4)
Inhaled steroids during last 12 months	12 (12)	10 (10)	1.2 (0.5 to 2.8)
Supplemental oxygen at discharge	16/98 (16)	23 (23)	0.7 (0.4 to 1.3)
Supplemental oxygen at 2 years	1 (1)	2 (2)	0.5 (0.0 to 5.5)
Respiratory examination			
Tracheostomy	0/95 (0)	1/98 (1)	–
Chest deformity	20/92 (22)	10/90 (11)	2.0 (1.0 to 3.9)
Harrison's sulcus	19/92 (21)	10/90 (11)	
Hyperinflation	9/92 (10)	1/90 (1)	
Hospital admissions			
Respiratory admission (n)	35 (35)	37 (37)	1.0 (0.7 to 1.4)
Median no. (upper quartile; max)*	2 (3; 8)	1 (2; 4)	
Surgical admission (n)	18 (19)	22 (22)	0.8 (0.5 to 1.5)
Median no. (upper quartile; max)*	1 (1; 2)	1 (2; 4)	
ICU admission and ventilated (n)	3 (3)	5 (5)	0.6 (0.2 to 2.5)
Median no. (upper quartile; max)*	1 (1; 1)	1 (1.5; 2)	
Admission other reason (n)	10/55 (18)	8/59 (14)	1.3 (0.6 to 3.1)
Median no. (upper quartile; max)*	1 (1; 2)	1 (2; 4)	

*Median number of admissions among those who had an admission.
GM-CSF, granulocyte-macrophage colony-stimulating factor; ICU, intensive care unit.

the control arm were discharged home with supplemental oxygen compared with 16 in the GM-CSF arm. Babies who are born following fetal growth restriction appear to be at high risk for chronic respiratory disease or bronchopulmonary dysplasia.^{29–31} This is supported by our finding that 40% of the trial population were receiving supplemental oxygen at 36 weeks

postmenstrual age.⁶ Cough and wheeze were marginally more common in the GM-CSF group and more children had chest deformity. However, the implication that ongoing respiratory symptoms are more frequent in the GM-CSF arm of the trial was not reflected in the use of bronchodilators or in any excess need for hospital readmission for respiratory disease.

Table 5 Mean health and social care costs over the first 2 years of life and mean cost differences between trial groups by cost category (UK £ sterling, 2007–2008 prices)

Cost category	GM-CSF (n=134)		Control (n=131)		Mean difference	(95% CI)*
	Mean	(SE)	Mean	(SE)		
Neonatal care	50 464	(2837)	56339	(3345)	–5875	(–14 369 to 2620)
Follow-up care (to 6 months)	5321	(690)	3771	(447)	1550	(–43 to 3143)
Follow-up care (6–12 months)	2698	(478)	2349	(398)	349	(–863 to 1561)
Follow-up care (12–18 months)	1948	(312)	1837	(321)	112	(–772 to 995)
Follow-up care (18–24 months)	1753	(264)	1963	(372)	–209	(–1104 to 686)
Total health and social care	62 186	(3517)	66 260	(3775)	–4073	(–14 239 to 6093)

*Bootstrap estimation using 10 000 replications, bias corrected.
GM-CSF, granulocyte-macrophage colony-stimulating factor.

Respiratory signs were not prespecified trial outcomes and the differences may have arisen by chance.

On theoretical grounds, GM-CSF might influence respiratory function in two opposing ways. Acute activation of neutrophils and monocytes in the lungs at the time of administration might lead to lung damage, though this has not been observed in extensive adult clinical practice. Conversely, GM-CSF has been postulated to be protective, though when this effect was tested in adults with acute lung injury through a randomised trial, GM-CSF treatment led to no difference in ventilator days or mortality.³² In the PROGRAMS cohort, an increase in immediate or long-term oxygen requirement was not seen in babies receiving GM-CSF in this or our previous study.^{4–5} The second is the influence of GM-CSF on the balance between TH1 and TH2 immune responses. The newborn immune system is TH1 deficient, hence less effective at mounting antibacterial responses.³³ GM-CSF is a TH1 agonist and might be expected to accelerate the switch away from TH2 dominance that in turn might plausibly translate into a reduction in later asthma.³⁴ To our best knowledge, this has never been investigated in a clinical study. This will be of great interest as we continue monitoring respiratory outcomes at 5-year follow-up.

CONCLUSIONS

The administration of GM-CSF to very preterm SGA babies is not associated with improved or more adverse developmental or health outcomes at 2 years of age. The apparent excess of developmental impairment in the PROGRAMS cohort, without corresponding increase in neurological abnormality, may represent diffuse brain injury attributable to intrauterine growth restriction. The stability of diagnoses between 2-year assessments and childhood measures of cognitive performance has been questioned recently.^{35–36} Developmental scores may vary widely in individuals even if the proportions within each group remain roughly the same in a preterm population. It is therefore important to evaluate more subtle and predictive outcomes at school age before the hypothesis that GM-CSF has no benefit in this very high-risk population can be rejected.

Acknowledgements The authors acknowledge gratefully the contribution of the PROGRAMS administrative team, in particular, the invaluable assistance of Anne Smith, PROGRAMS coordinator responsible for contact and tracing of the participants and for data entry and collation, Liz Schroeder for assistance with the economic analysis and Ursula Bowler, Sarah Ayers and other staff of the National Perinatal Epidemiology Unit which undertook trial administration. Bliss, the premature and sick newborn charity, assisted with trial materials, and the authors thank all parent and infant participants. The trial was sponsored by Imperial College London.

Contributors Investigators NMo, RC, PB and CJD formulated the hypothesis and design of the primary PROGRAMS trial; NMo, RC, PB, CJD and NMa formulated the hypothesis and design of the follow-up study. TM analysed the data with supervision by CJD. SP and NP were responsible for the economic analysis and interpretation. NMa wrote the first draft of this paper and coordinated manuscript revisions and submission. All investigators with FC and MR were involved in management of the 2-year follow-up study, interpretation of data, and writing the report. All the authors have seen and approved the final version. Two-year outcome study oversight committees: Data Monitoring Committee: D Field, J Hutton, S Saigal, C Bennett; Follow-up Trial Steering Committee Independent Members: M Chiswick, M Deans, S Riddell, M Weindling. Trial hospitals and investigators: Brighton, Royal Sussex County (H Rabe, F Weir); Bristol, Southmead (D Evans), Bristol, St Michaels (P Cairns, A Jain); Cambridge, the Rosie (A Ogilvy-Stuart); Cardiff, University of Wales (S Barr, P Cartledge); Dundee, Ninewells (P Fowlie); Gillingham, Medway Maritime (B Jani); Glasgow, Princess Royal (A Powls), Ipswich (K O'Neill); Liverpool Women's (C Morgan, M Weindling), London, Chelsea and Westminster (D Acolet, E Ogundipe); London, Guy's & St Thomas' (M Campbell, G Fox, J Rissik); London, Hammersmith (N Murray); London, Hillingdon (M Cruwys); London, Homerton (E Maalouf); London, Newham General (V Gopinathan); London, Northwick Park

(R Nicholl); London, University College (J Hawdon); Manchester, St Mary's (A Emmerson); Norwich, Norfolk & Norwich University (P Clarke, J Eason); Nottingham, Queens (S Wardle); Plymouth, Derriford (J Lilley); Portsmouth, St Mary's (R Thwaites); Slough, Wexham Park (R Jones); Swansea, Singleton (J Mathers); Woolwich, Queen Elizabeth (B Al-Rubeyi). Two-year Outcome Study Assessors: Amanda Cundy, Angela Huertas, Huw Jones, Louise Watson, Sajjad Rahman, Andrew Cheew, Ann Humphreys, Leigh Dyet, Angela D'Amore, Georgie Siggers, Lesley McDonald, Nick Wood, Sam Johnson.

Funding Action Medical Research and the Wellcome Trust; Neil Marlow receives part funding from the Department of Health's NIHR Biomedical Research Centres funding scheme at UCLH/UCL. Robert Carr received part funding from the NIH Biomedical Research Centre at King's College London.

Competing interests None.

Ethics approval South Thames Multicentre Research Ethics Committee.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Data from this study are still under analysis as a 5-year outcome study is now under way. The database is available for researchers following application to the chief investigator (N Modi) subject to review by the investigators group.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>

REFERENCES

- Volpe JJ. Brain injury in premature infants: a complex amalgam of destructive and developmental disturbances. *Lancet Neurol* 2009;8:110–24.
- Sung L, Nathan PC, Lange B, et al. Prophylactic granulocyte colony-stimulating factor and granulocyte-macrophage colony-stimulating factor decrease febrile neutropenia after chemotherapy in children with cancer: a meta-analysis of randomized controlled trials. *J Clin Oncol* 2004;22:3350–6.
- Smith TJ, Khatcheressian J, Lyman GH, et al. 2006 update of recommendations for the use of white blood cell growth factors: an evidence-based clinical practice guideline. *J Clin Oncol* 2006;24:3187–205.
- Carr R, Modi N, Dore C. G-CSF and GM-CSF for treating or preventing neonatal infections. *Cochrane Database Syst Rev* 2003;(3):CD003066.
- Carr R, Brocklehurst P, Doré CJ, et al. Granulocyte-macrophage colony stimulating factor administered as prophylaxis for reduction of sepsis in extremely preterm, small for gestational age neonates (the PROGRAMS trial): a single-blind, multicentre, randomised controlled trial. *Lancet* 2009;373:226–33.
- Frisone MF, Mercuri E, Laroche S, et al. Prognostic value of the neurologic optimality score at 9 and 18 months in preterm infants born before 31 weeks' gestation. *J Pediatr* 2002;140:57–60.
- Bayley N. *Manual for the Bayley Scales of Infant Development*. Second edition. San Antonio, Texas: The Psychological Corporation 1993.
- Johnson S, Marlow N, Wolke D, et al. Validation of a parent report measure of cognitive development in very preterm infants. *Dev Med Child Neurol* 2004;46:389–97.
- Johnson S, Wolke D, Marlow N. Developmental assessment of preterm infants at 2 years: validity of parent reports. *Dev Med Child Neurol* 2008;50: 58–62.
- Report of two working groups. Disability and perinatal care: Measurement of health status at two years: National Perinatal Epidemiology Unit and Oxford Health Authority, 1995.
- Jones HP, Guillea ZE, Stewart JH, et al. The Health Status Questionnaire: achieving concordance with published disability criteria. *Arch Dis Child* 2002;86:15–20.
- The Child Growth Foundation. *British 1990 growth reference for height, weight, BMI and h/c*. London: The Child Growth Foundation, 1996.
- de Onis M, Yip R, Mei Z. The development of MUAC-for-age reference data recommended by a WHO Expert Committee. *Bull World Health Organ* 1997;75:11–18.
- National Institute for Health and Clinical Excellence (NICE). *Guide to the method of technology appraisal*. London: NICE, 2008.
- Drummond MF, Schulpher MJ, Torrance GW, et al. *Methods for the economic evaluation of health care programmes*. Third edition. Oxford: Oxford University Press 2005.
- Petrou S, Gray A. Economic evaluation using decision analytical modelling: design, conduct, analysis, and reporting. *BMJ* 2011;342:d1766.
- Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;340:c332.
- Altman DG, Machin D, Bryant TN, et al. *Statistics With Confidence*. Second edition. Oxford: Wiley-Blackwell 2000.

- 19 Wood NS, Marlow N, Costeloe K, *et al.* Neurologic and developmental disability after extremely preterm birth. EPICure Study Group. *N Engl J Med* 2000;343:378–84.
- 20 Johnson A. Disability and perinatal care. *Pediatrics* 1995;95:272–4.
- 21 Walker DM, Marlow N. Neurocognitive outcome following fetal growth restriction. *Arch Dis Child Fetal Neonatal Ed* 2008;93:F322–5.
- 22 Bhutta AT, Cleves MA, Casey PH, *et al.* Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis. *JAMA* 2002;288:728–37.
- 23 Johnson S, Marlow N. Developmental screen or developmental testing? *Early Hum Dev* 2006;82:173–83.
- 24 Kenyon S, Pike K, Jones DR, *et al.* Childhood outcomes after prescription of antibiotics to pregnant women with spontaneous preterm labour: 7-year follow-up of the ORACLE II trial. *Lancet* 2008;372:1319–27.
- 25 Johnson S, Whitelaw A, Glazebrook C, *et al.* Randomized trial of a parenting intervention for very preterm infants: outcome at 2 years. *J Pediatr* 2009;155:488–94.
- 26 Bassler D, Stoll BJ, Schmidt B, *et al.* Using a count of neonatal morbidities to predict poor outcome in extremely low birth weight infants: added role of neonatal infection. *Pediatrics* 2009;123:313–18.
- 27 Stoll BJ, Hansen NI, Adams-Chapman I, *et al.* Neurodevelopmental and growth impairment among extremely low-birth-weight infants with neonatal infection. *JAMA* 2004;292:2357–65.
- 28 Schäbitz WR, Krüger C, Pitzer C, *et al.* A neuroprotective function for the hematopoietic protein granulocyte-macrophage colony stimulating factor (GM-CSF). *J Cereb Blood Flow Metab* 2008;28:29–43.
- 29 Stocks J, Coates A, Bush A. Lung function in infants and young children with chronic lung disease of infancy: the next steps? *Pediatr Pulmonol* 2007;42:3–9.
- 30 Morsing E, Gustafsson P, Brodzski J. Lung function in children born after foetal growth restriction and very preterm birth. *Acta Paediatr* 2012;101:48–54.
- 31 Bose C, Van Marter LJ, Laughon M, *et al.* Fetal growth restriction and chronic lung disease among infants born before the 28th week of gestation. *Pediatrics* 2009;124:e450–8.
- 32 Paine R 3rd, Standiford TJ, Dechert RE, *et al.* A randomized trial of recombinant human granulocyte-macrophage colony stimulating factor for patients with acute lung injury. *Crit Care Med* 2012;40:90–7.
- 33 Levy O. Innate immunity of the newborn: basic mechanisms and clinical correlates. *Nat Rev Immunol* 2007;7:379–90.
- 34 Renaud JC. New insights into the role of cytokines in asthma. *J Clin Pathol* 2001;54:577–89.
- 35 Marlow N, Wolke D, Bracewell MA, *et al.* Neurologic and developmental disability at six years of age after extremely preterm birth. *N Engl J Med* 2005;352:9–19.
- 36 Hack M, Taylor HG, Drotar D, *et al.* Poor predictive validity of the Bayley Scales of Infant Development for cognitive function of extremely low birth weight children at school age. *Pediatrics* 2005;116:333–41.