BMJ Open Ophthalmology

Evaluation of the Retinopathy of Prematurity Activity Scale (ROP-ActS) in a randomised controlled trial aiming for prevention of severe ROP: a substudy of the Mega Donna Mega trial

Aldina Pivodic [ID],[1] Helena Johansson,[2,3] Lois EH Smith,[4] Chatarina Löfqvist,[1,5] Kerstin Albertsson-Wikland,[6] Staffan Nilsson,[7,8] Ann Hellström [ID] [1]

For numbered affiliations see end of article.

Correspondence to
Aldina Pivodic; aldina.pivodic@gu.se

ABSTRACT

Objective The current grading of retinopathy of prematurity (ROP) does not sufficiently discriminate disease severity for evaluation of trial interventions. The published ROP Activity Scales (original: ROP-ActS and modified: mROP-ActS), describing increasing severity of ROP, versus the categorical variables severe ROP, stage, zone and plus disease were evaluated as discriminators of the effect of an ROP preventive treatment.

Methods and analysis The Mega Donna Mega trial investigated ROP in infants born <28-week gestational age (GA), randomised to arachidonic acid (AA) and docosahexaenoic acid (DHA) supplementation or no supplementation. Of 207 infants, 86% with finalised ROP screening were included in this substudy. ROP-ActS versus standard variables were evaluated using Fisher's non-parametric permutation test, multivariable logistic and linear regression and marginal fractional response models.

Results The AA:DHA group (n=84) and the control group (n=93) were well balanced. The maximum ROP-ActS measurement was numerically but not significantly lower in the AA:DHA group (mean: 4.0 (95% CI 2.9 to 5.0)) versus the control group (mean: 5.3 (95% CI 4.1 to 6.4)), p=0.11. In infants with any ROP, the corresponding scale measurements were 6.8 (95% CI 5.4 to 8.2) and 8.7 (95% CI 7.5 to 10.0), p=0.039. Longitudinal profiles of the scale were visually distinguished for the categories of sex and GA for the intervention versus control.

Conclusions The preventive effect of AA:DHA supplementation versus no supplementation was better discriminated by the trial's primary outcome, severe ROP, than by ROP-ActS. The sensitivity and the linear qualities of ROP-ActS require further validations on large data sets and perhaps modifications.

Trial registration number NCT03201588.

Key messages

What is already known on this topic?

► A retinopathy of prematurity (ROP) Activity Scale was developed based solely on clinical experience aimed for use in clinical trials evaluating preventive therapies and treatments. The scale includes combinations of several aspects of the disease and could be a more sensitive grading measure of ROP than stage, zone and plus disease alone.

What this study adds?

► In a published randomised controlled trial examining reduction of severe ROP with enteral fatty acids supplementation, the treatment effect was better explained by standard ROP measures than the ROP Activity Scale.

How this study might affect research, practice or policy?

► The scale might help distinguish patterns of ROP profiles longitudinally in different subgroups. However, the sensitivity and the linearity of the ROP Activity Scale require further validation and potential modifications before use in clinical trials.

INTRODUCTION

Retinopathy of prematurity (ROP) is a sight-threatening disease diagnosed and monitored through repeated eye examinations.[1] Currently, in Sweden, all infants born before 30 weeks of gestation, and those weighing <1500 g at birth in the absence of a reliable gestational age (GA), are routinely screened for ROP, as well as infants who have serious illness that increase their risk of ROP.[2 3] Before 2020 (when criteria were modified), all infants born <31 weeks of gestation were routinely screened.[4] ROP strongly relates to the infants' prematurity (reflecting the degree of avascular retina after birth) and concurrent morbidities (reflecting potential inhibition of retinal vascular growth).[5 6] More extremely prematurely born babies now survive the critical neonatal period owing to continuous improvements in neonatal intensive care.[7] Therefore, there is a need to develop prevention and treatment therapies

for ROP and other morbidities affecting the increasing number of extremely preterm infants.[8]

Today, the severity of ROP is predominantly described through the categorical variables ROP stage, zone and plus disease in clinical trials.[9] However, considering each of these variables separately is thought to insufficiently reflect disease severity. Thus, in 2019, the International Neonatal Consortium was asked to develop and publish an ROP Activity Scale (ROP-ActS) based on several aspects of the disease, aimed to be used as a numerical variable.[10] Only the scale's ordering was taken into account at the development stage. The scale ranges from 0 (no ROP) to 22 (most severe ROP). Based on clinical experience, scorings 1–18 reflect the ranking of each combination of the three variables: ROP stage (1=least severe to 3=most severe), zone (I=most central to III=most peripheral) and plus disease (yes/no). Scores 19–22 are assigned to aggressive posterior ROP, and stages 4a, 4b and 5 constituting different levels of severity of retinal detachment. ROP-ActS was developed for potential use in clinical trials as a scale describing the severity of ROP with greater sensitivity than the previously used categorical variables.[10 11] Such a scale if linear with respect to relevant outcomes could help reduce the sample size in clinical trials and thereby accelerate the investigation of a therapy.

To date, ROP-ActS has been evaluated in one ROP screened cohort where a small modification of the scale was proposed (mROP-ActS).[12] In the Mega Donna Mega (MDM) trial, designed by our group to study preventive effects of arachidonic acid (AA) and docosahexaenoic acid (DHA) supplementation on severe ROP (stage 3 and type 1), we demonstrated that the incidence of severe ROP was 50% lower in the AA:DHA group compared with the control group.[13] In the current study, we evaluated the mean levels of the maximum ROP-ActS scores in the two randomised groups from the MDM cohort. We compared the scale's ability to discriminate the preventive effect of AA:DHA versus using the current standard regarding severe ROP, stage, zone and plus disease. Lastly, we studied the longitudinal patterns of ROP-ActS by treatment group, overall, and in selected subgroups.

MATERIALS AND METHODS
Written informed consent was obtained from the parents/guardians of all included infants. Patients or the public were not involved in the design, or conduct, or reporting, or dissemination plans of our research.

A detailed description of the design, treatment, primary and secondary analyses and safety evaluations is available in the MDM's original publication.[13]

Study population
The MDM trial included infants born before 28-week GA from three (Gothenburg, Lund and Stockholm) neonatal intensive care units in Sweden from December 2016 to August 2019. Out of 207 infants, randomised to receive either supplementation or non-supplementation

of AA:DHA (1:1), 177 (85.5%) were included in the current substudy having a final evaluation from ROP screening examinations, 84 (82.4%) in the AA:DHA and 93 (88.6%) infants in the control group, comprising the full analysis set, see figure 1. The randomisation was stratified by centre and GA categories (<25 weeks, 25–26 weeks and 27 weeks).

Study treatment
The interventional arm received AA:DHA supplementation (Formulaid, 100:50 mg/kg/day) starting within 3 days of postnatal age (PNA) up to 40 weeks of postmenstrual age (PMA). Both randomisation groups received conventional treatment according to national and local guidelines.

Study procedures
ROP screening and management was performed following current national guidelines.[4] SD scores (SDS) for weight, length and head circumference at birth were standardised for GA and sex using Swedish reference based on 800 000 healthy singletons born 1990–1999.[14] The diagnosis of necrotising enterocolitis (NEC) was based on the criteria by Walsh and Kliegman with stage 2A or greater considered as disease.[15] Patent ductus arteriosus (PDA) was diagnosed based on the need for surgical or medical treatment. For diagnosis of intraventricular haemorrhage (IVH, grades 0–4), cranial ultrasound was performed on postnatal days 3 and 7. Bronchopulmonary dysplasia (BPD) was diagnosed based on the need for supplemental oxygen at 36 months PMA. Sepsis was defined by clinical symptoms, C reactive protein >20 mg/dL or interleukin 6 >1000 pg/mL either confirmed or not by blood culture.

Study outcomes
The main study outcomes were ROP-ActS and mROP-ActS (online supplemental table 1). The outcomes severe ROP (yes/no), defined as ROP stage 3 or type 1 ROP, ROP stage (none, 1, 2 or 3), zone (I, II or III, none; grouping zone I and II in one category in the statistical tests due to low frequency of zone I) and plus disease (yes/no), were compared to ROP-ActS. ROP treatment was defined following early treatment for ROP criteria.[16] ROP stages were defined using the International Classification of ROP.[9] The most severe outcome of the two eyes was analysed on patient level and described longitudinally using PNA as time scale.

Statistical analysis
Statistical methods and analyses were prespecified in a statistical analysis plan.

ROP-ActS has not been studied in a prevention trial before, and it has not yet defined a minimal clinically relevant difference. SD of 5.2 was observed in the retrospective cohort where first validation of ROP-ActS was performed, selecting infants with GA <28 weeks. Assuming a difference of 2 and 2.5 scores in ROP-ActS, SD of 5.2, α of 0.05, with a two-sided Fisher's non-parametric
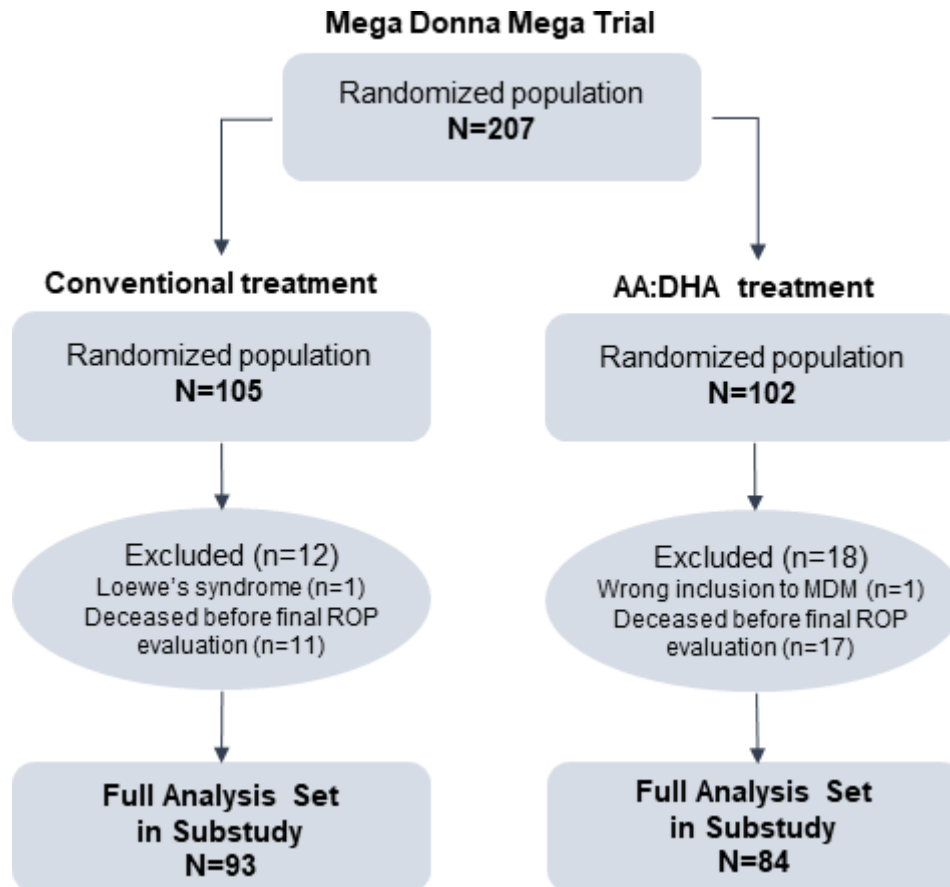
Mega Donna Mega Trial

Figure 1 Study flow chart. AA, arachidonic acid; DHA, docosahexaenoic acid; ROP, retinopathy of prematurity; MDM, Mega Donna Mega.

permutation test, a cohort including 84 infants in the AA:DHA and 93 infants in the control group would have a power of 72% and 88% to detect a difference between the two treatment groups.

Continuous variables were described by mean, SD, median and range, or median and IQR, as applicable. Categorical variables were described by number and percentage. For test of differences between the two treatment groups with respect to continuous variables, Fisher's non-parametric permutation test was used. For ordered categorical variables, Mantel-Haenszel χ^2 trend test was used, and for dichotomous variables, Fisher's exact test was used.

The primary analysis was performed using Fisher's non-parametric permutation test and described by the mean difference with 95% CIs of ROP-ActS between the two treatment groups. Adjusted analysis was performed using multivariable linear regression with ROP-ActS, stage and zone as separately evaluated dependent variables, treatment group as main effect variable, adjusting for GA and centre. Distribution of residuals was reviewed in the diagnostic plots and found satisfactory. In a similar way, logistic regression was used for adjusted analyses of severe ROP, and plus disease. Post-hoc analysis was performed by reproducing the primary and adjusted analyses on a subgroup of infants with any ROP.

In order to visually compare the linearity characteristic of ROP-ActS and mROP-ActS in the current cohort, as was previously evaluated on a retrospective cohort, the two scale variables were studied against ROP treatment.[12] Percentage of infants with ROP treatment among those having reported a certain score was presented in a bar chart.

Longitudinal ROP-ActS values were analysed using marginal fractional response models with binomial distribution and logit link function. The dependent variable was a fractional variable, ROP-ActS divided by 22 (the maximum score). The time was modelled with natural cubic splines and within infant correlation of values over time using spatial covariance structure. These analyses were performed on two data sets: (1) all infants and values at risk, representing collected data from ROP screening examinations over time, and (2) all infants and collected values from ROP screenings including imputed no ROP during the follow-up, that is, for those infants that stop their screening before PNA week 30 or PMA week 50, whichever occurs first, zeros were imputed every 2nd week. Analyses were performed overall for intervention versus control and separately by sex and GA strata (22–24 weeks, 25–26 weeks and 27 weeks), investigating treatment versus sex and treatment versus GA strata interactions.

All tests were two-sided. Fix sequential testing was aimed to be applied for the control of type 1 error for the confirmatory analyses, the primary analysis being the first one in the order. The primary analysis could not be confirmed. Hence, no other analyses in the prespecified sequential order were of relevance for evaluation. Statistical models examining longitudinal data were considered exploratory and, therefore, not confirmative. Analyses were performed using SAS software V.9.4.

RESULTS
Study population
There were no statistically significant differences between the two treatment groups with respect to infants' and mothers' characteristics at study start (table 1). In the AA:DHA and the control group, mean GA was 25.6 weeks (SD: 1.5 weeks) and 25.6 weeks (SD: 1.4 weeks), 40.5% and 46.2% were girls and mean birth weight was 818 g (SD: 205 g) and 795 g (SD: 196 g), respectively. Overall, the incidence of NEC was 7.9%, PDA 53.1%, any IVH grade 39.5%, BPD 53.7% and sepsis 48.0%. The percentage of mothers with any medical history differed numerically between the AA:DHA, 50.0%, and the control group, 36.6%.

The infants were followed up for median 15.4 weeks (IQR: 13.8–17.0 weeks) in the AA:DHA group and median 15.4 weeks (IQR: 13.7–16.9 weeks) in the control group.

AA:DHA treatment effect on ROP-ActS, mROP-ActS and traditional ROP severity variables
The evaluation of different ROP severity variables comparing the AA:DHA with the control group is presented in table 2. Graphical presentation is available in online supplemental figure 1A,B, and the distribution of infants per each maximum score in the online supplemental table 1.

The mean value of maximum ROP-ActS during the study was numerically, but not significantly, lower in the AA:DHA group compared with the control group (mean: 4.0 (95% CI 2.9 to 5.0) versus 5.3 (95% CI 4.1 to 6.4), mean difference: −1.29 (95% CI −2.87 to 0.26), unadjusted p=0.11 and adjusted for GA and centre p=0.057). Hence, the primary analysis could not be confirmed. Somewhat attenuated estimates were obtained for mROP-ActS. Studied on the full analysis set in the MDM cohort, there were significantly fewer infants with severe ROP (the MDM's originally defined primary variable, stage 3 or type 1) in the AA:DHA group than in the control group (mean: 17 (20.2%) vs 34 (36.6%), mean difference: −16.3 (95% CI −30.5 to −2.2), unadjusted p=0.025 and adjusted for GA and centre p=0.0072). No significant differences between the treatment groups were observed for maximum ROP stage, most central zone or plus disease. Post-hoc analyses, performed on the full analysis set, including infants with any ROP, identified the most discriminating variables for the effect of treatment in the

following order: ROP stage, severe ROP, ROP-ActS and mROP-ActS (table 2).

ROP-ActS versus mROP-ActS and their relation to ROP treatment
The percentage of infants with ROP treatment among ever reported ROP-ActS/mROP-ActS scores was obtained and presented in online supplemental figure 2. In total, 35 infants (19.8%) were treated for ROP, 12 (14.3%) in the AA:DHA supplemented group and 23 (24.7%) in the control group. Among infants with ever reported ROP-ActS score 3 (corresponds to mROP-ActS score 5), 12 (24.0%) were treated for ROP. Among infants with observed ROP-ActS score 5 (corresponds to mROP-ActS score 3), 2 (22.2%) were treated for ROP. No infants had reported ROP-ActS scores 4 (zone III stage 1+), 6 (zone III stage 2+), 11 (zone II stage 1+), 12 (zone I stage 2), 15 (zone I stage 1+), 16 (zone I stage 3) and 17 (zone I stage 2+).

AA:DHA treatment effect on longitudinal values of ROP-ActS
Mean curves of ROP-ActS for all infants at risk analysed continuously over time are presented in figure 2A by intervention versus the control group. The analysed data set, including imputed no ROP for infants that stop their screening early, is represented in online supplemental figure 3A, where the estimates for later time points were pressed toward 0. Using the same longitudinal method for all infants at risk, we observed that there were significant interactions between the treatment group and sex (p=0.0090) and between the treatment group and GA strata (p=0.0048) (figure 2B–F). The same analyses were performed on the data set, including the imputed values (online supplemental figure 3B–F).

DISCUSSION
Using data from the MDM trial, we could, for the first time, evaluate the advantages and disadvantages of the ROP activity scale on a prospective, controlled, randomised study, where this scale was intended for use. Our study showed that ROP-ActS did not discriminate better for the MDM study's primary variable than the originally defined severe ROP (stage 3 and type 1 ROP), although the AA:DHA treatment group had numerically lower mean levels of ROP-ActS than the control group.

The choice of a study's primary endpoint is crucial when we are interested in demonstrating an effect of a therapy. Besides the chosen statistical methodology, the variable type and the expected effect size, affecting the study's sample size, must fulfil the requirements of being the most clinically relevant candidate, a priori and clearly defined, validated, reliable and evidence based in order to follow recommendations from the authorities.[17] The MDM trial's primary variable, severe ROP, was an a priori and clearly defined variable. The primary variable was validated and reliable since the diagnosis protocol for ROP stages, and type 1 ROP was followed by all investigating ophthalmologists at the study centres

Table 1 Infants' and mothers' characteristics at study start by treatment group (full analysis set)

	Randomised group		P value
	AA:DHA (n=84)	Control (n=93)	
GA (weeks)	25.6 (1.5) 25.6 (22.6; 27.9)	25.6 (1.4) 25.9 (22.9; 27.9)	0.79
Female	34 (40.5%)	43 (46.2%)	0.54
Birth weight (g)	818 (205) 796 (455; 1345)	795 (196) 770 (425; 1330)	0.45
Birth weight SDS 14	−0.66 (1.25) −0.36 (−5.06; 1.74)	−0.78 (1.16) −0.50 (−4.06; 1.22)	0.53
Birth length (cm)	33.0 (2.7) 33 (27; 39) n=74	33.0 (2.8) 33 (28.5; 41) n=76	0.99
Birth length SDS 14	−1.24 (1.61) −1.02 (−6.76; 1.42) n=74	−1.28 (1.55) −1.08 (−4.7; 2.89) n=76	0.86
Birth head circumference (cm)	23.4 (1.9) 23.4 (18.7; 27.5) n=83	23.3 (1.8) 23.3 (20.2; 27.5) n=91	0.76
Birth head circumference SDS 14	−0.61 (0.79) −0.51 (−2.86; 0.84) n=83	−0.66 (0.78) −0.63 (−2.50; 1.20) n=91	0.62
Plurality			
Single	69 (82.1%)	75 (80.6%)	
Twin	12 (14.3%)	17 (18.3%)	
Triplet	3 (3.6%)	1 (1.1%)	0.89
Centre			
Gothenburg	32 (38.1%)	33 (35.5%)	
Stockholm	25 (29.8%)	27 (29.0%)	
Lund	27 (32.1%)	33 (35.5%)	0.89
Morbidities			
NEC	5 (6.0%)	9 (9.7%)	0.53
PDA	49 (58.3%)	45 (48.4%)	0.24
Cerebral IVH			
Grade 0	50 (59.5%)	57 (61.3%)	
Grade 1	14 (16.7%)	10 (10.8%)	
Grade 2	13 (15.5%)	12 (12.9%)	
Grade 3	4 (4.8%)	5 (5.4%)	
Grade 4	3 (3.6%)	9 (9.7%)	0.42
BPD	48 (57.1%)	47 (50.5%)	0.47
Sepsis	37 (44.0%)	48 (51.6%)	0.39
Mother's age (years)	32.2 (5.3) 32.4 (16.2; 41.2) n=82	31.8 (4.8) 31.1 (22.4; 44) n=92	0.61
Parity (number of child in order)	1.68 (0.88) 1 (1; 5) n=80	1.63 (0.90) 1 (1; 5) n=92	0.80
Caesarean section	52 (61.9%)	55 (59.1%)	0.83
Mothers with any medical history	42 (50.0%)	34 (36.6%)	0.098
Mothers with diabetes	1 (1.2%)	0 (0.0%)	0.95
Mothers with hypertension	6 (7.1%)	13 (14.0%)	0.22

For categorical variables, n (%) is presented.
For continuous variables, mean (SD)/median (minimum; maximum)/n=(unless no. missing) is presented.
AA, arachidonic acid; BPD, bronchopulmonary dysplasia; DHA, docosahexaenoic acid; GA, gestational age; IVH, intraventricular haemorrhage; NEC, necrotising enterocolitis; PDA, patent ductus arteriosus; SDS, SD scores.

Table 2 Primary, secondary, adjusted and post-hoc analyses

	Randomised group: AA:DHA (n=84)	Randomised group: control (n=93)	P value	Unadjusted difference between groups: mean (95% CI)	Adjusted P value*
Primary analysis: full analysis set					
Maximum ROP-ActS	4.0 (4.9) 2 (0; 14) (2.9; 5.0)	5.3 (5.6) 3 (0; 18) (4.1; 6.4)	0.11	−1.29 (−2.87 to 0.26)	0.057
Maximum mROP-ActS	4.2 (4.9) 2 (0; 14) (3.1; 5.3)	5.4 (5.5) 5 (0; 18) (4.2; 6.5)	0.14	−1.16 (−2.72 to 0.37)	0.087
Secondary analyses: full analysis set					
Severe ROP (stage 3 and type 1)					
No	67 (79.8%)	59 (63.4%)		16.3 (2.2 to 30.5)	
Yes	17 (20.2%)	34 (36.6%)	0.025	−16.3 (−30.5 to −2.2)	0.0072
Maximum ROP stage					
No ROP	35 (41.7%)	37 (39.8%)		1.9 (−13.8 to 17.5)	
Stage 1	15 (17.9%)	7 (7.5%)		10.3 (−0.6 to 21.3)	
Stage 2	17 (20.2%)	16 (17.2%)		3.0 (−9.6 to 15.7)	
Stage 3	17 (20.2%)	33 (35.5%)	0.12	−15.3 (−29.4 to −1.1)	0.057
Most central zone (grouped)					
No ROP	35 (41.7%)	37 (39.8%)		1.9 (−13.8 to 17.5)	
III	11 (13.1%)	7 (7.5%)		5.6 (−4.6 to 15.7)	
I or II	38 (45.2%)	49 (52.7%)	0.51	−7.5 (−23.3 to 8.4)	0.48
Plus disease					
No	72 (85.7%)	72 (77.4%)		8.3 (−4.2 to 20.8)	
Yes	12 (14.3%)	21 (22.6%)	0.22	−8.3 (-20.8 to 4.2)	0.10
Post-hoc analyses: infants with any ROP in the full analysis set					
Maximum ROP-ActS	6.8 (4.7) 7 (1; 14) (5.4; 8.2)	8.7 (4.6) 8 (1; 18) (7.5; 10.0)	0.039	−1.94 (−3.75 to 0.12)	0.024
Maximum mROP-ActS	7.2 (4.5) 7 (1; 14) (5.9; 8.5)	8.9 (4.4) 8 (1; 18) (7.7; 10.1)	0.055	−1.71 (−−3.43 to 0.03)	0.038
Severe ROP					
No	32 (65.3%)	22 (39.3%)		26.0 (5.6 to 46.4)	
Yes	17 (34.7%)	34 (60.7%)	0.013	−26.0 (-46.4 to −5.6)	0.0057
Maximum ROP stage					
Stage 1	15 (30.6%)	7 (12.5%)		18.1 (0.7 to 35.6)	
Stage 2	17 (34.7%)	16 (28.6%)		6.1 (−13.6 to 25.9)	
Stage 3	17 (34.7%)	33 (58.9%)	0.0060	−24.2 (−44.7 to −3.8)	0.0031
Most central zone (grouped)					
III	11 (22.4%)	7 (12.5%)	0.28	9.9 (−6.5 to 26.4)	0.19
I or II	38 (77.6%)	49 (87.5%)		−9.9 (−26.4 to 6.5)	
Plus disease					
No	37 (75.5%)	35 (62.5%)		13.0 (−6.4 to 32.4)	
Yes	12 (24.5%)	21 (37.5%)	0.22	−13.0 (−32.4 to 6.4)	0.11

For continuous variables, mean (SD)/median (minimum; maximum)/(95% CI for mean using the inversion of Fisher's non-parametric permutation test)/n=is presented.

*Adjusting for GA (weeks) and centre using multivariable linear regression handling AA:DHA intervenion group as the main effect variable, and ROP-ActS, mROP-ActS, ROP stage and zone as dependent variables in separate analyses. For severe ROP and plus disease as outcome (dependent variable), multivariable logistic regression was used instead.

AA, arachidonic acid; DHA, docosahexaenoic acid; GA, gestational age; mROP-ActS, modified Retinopathy of Prematurity Activity Scale; ROP, retinopathy of prematurity; ROP-ActS, Retinopathy of Prematurity Activity Scale.
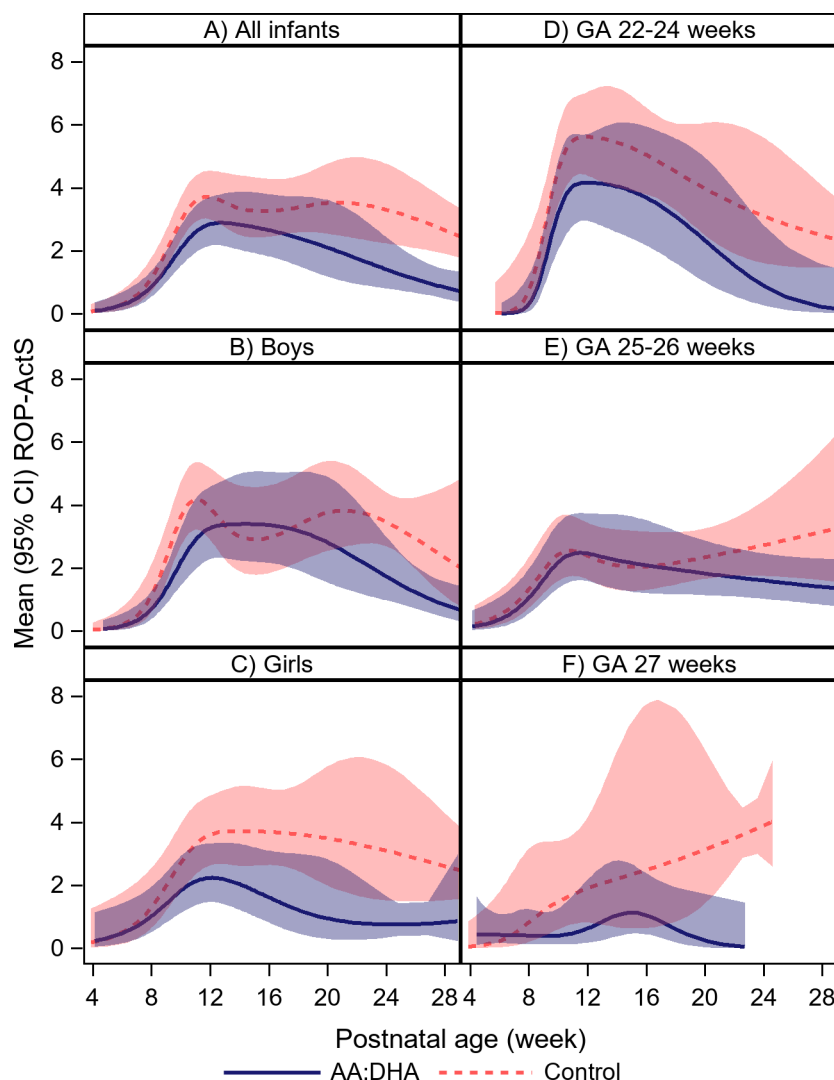
Figure 2 Longitudinal ROP-ActS values by treatment groups on at risk data set, including all data points from the existing ROP examinations without imputation for (A) all infants, (B) boys, (C) girls, (D) GA 22–24 weeks, (E) GA 25–26 weeks and (F) GA 27 weeks. The estimated mean values are not showing expected mean for the whole prematurely born population in this study, but only for infants that require ROP screening at different postnatal time points. Not all infants have regularly reported values over time, since the ROP screening examinations are individually defined for all infants. AA, arachidonic acid; DHA, docosahexaenoic acid; GA, gestational age; ROP, retinopathy of prematurity; ROP-ActS, Retinopathy of Prematurity Activity Scale.

as recommended.[9 16] It also fulfilled the requirement of clinical relevance as it identifies infants with the highest risk for end-stage ROP. Furthermore, both preclinical and clinical studies have demonstrated that higher levels of AA and DHA fatty acids have protective effects against pathological neovascularisation and severe ROP. Hence, the study's primary variable was also evidence based.[18–21]

Given the information above, we ask ourselves: What advantage would be gained by introducing an ROP-ActS in such a trial?

First, a well-ordered continuous scale would be a more sensitive measure promoting even smaller confirmatory clinical trials to have a successful outcome, provided that the treatment of interest benefits a cohort uniformly along the scale. In the retrospective validation study of ROP-ActS, including 535 infants and 3324 ROP screening

examinations, we concluded that the original ROP-ActS was relatively well ordered against the short-term outcome, treated ROP. It was proposed to switch scores 3 (zone II stage 1) and 5 (zone III stage 3).[12] In the current study, the percentage of infants with score 3 that finally led to ROP treatment was 24%, and for score 5 was 22%, compared with 33% and 14% in the retrospective study. The decision for switching the scores 3 and 5 would not be as substantiated in this study as in the previous one. Therefore, the scale's well-ordering characteristic against various outcomes needs further evaluation.

Second, an approximately linear continuous scale would imply better prediction ability and facilitate the statistical evaluation. Studying a continuous scale as a linear variable in a regression analysis would mean that we only require one parameter to be estimated, implying

a simpler statistical analysis. Besides increasing the power, this would also contribute to the well and clearly defined study variable. To achieve a completely linear scale, with regards to different populations, outcome and treatments, would be a monumental, if not impossible exercise. However, it is important to continue to try to find an approximately linear scale versus most commonly studied outcomes. It should be emphasised that in assessing the linearity of a scale, we must evaluate its linearity in predicting a certain outcome, and linearity on the components of the studied treatment. For example, in the current study, we reported similar percentages between the treatment and a non-treatment group for infants with no ROP, but noted treatment differences for more severe ROP. A discrete linear scale does not need to have the same difference between each of the two consecutive scores. Neither does it need to have one score per combination of the three variables. Some combinations might have the same risk for the outcome of interest, and some combinations might not be clinically justifiable. In the retrospective and the current study, no infants with scores 4 (zone III stage 1+), 6 (zone III stage 2+), 11 (zone II stage 1+), 15 (zone I stage 1+) and 17 (zone I stage 2+) were observed. More extensive future studies will reveal whether those scores are non-existing or are just rare. Given the rare observation of certain combinations of the three variables currently comprising the scale (stage, zone and plus disease), the methodical development of an ROP-ActS scale and evaluation of its linearity and well-ordering characteristic in relation to various ROP outcomes would require large data sets, including longitudinal follow-up. Following the recently published third edition of the International Classification of ROP, initiated due to subjectivity issues in diagnostics, regression and re-activation of the disease due to new therapies and new features identified through innovations in imaging, the classification of ROP has been further refined with new ROP-related metrics.[22] Therefore, ROP experts should help decide whether any of these new metrics are suitable for incorporation into the scale.

Third, early postnatal values of a well-ordered and an approximately linear scale might potentially detect the treatment effect already in the early follow-up that could imply shorter duration of clinical trials.

Last but not least, longitudinal postnatal values of a well-ordered and an approximately linear scale could be used for identifying different patterns of ROP development for different subgroups, important for understanding physiological processes. In an exploratory study, we visually distinguished different profiles for the two sexes and GA strata for the investigational product and the control group. Those interactions were significant. Larger data sets could be used to identify mean levels and ranges for specific GAs or other relevant subgroups.

The prevention of ROP by AA:DHA supplementation was shown to be significant on the most severe parts of the disease scale, severe ROP 20% versus 37%, but not overall, any ROP 58% vs 60%, in the AA:DHA versus the control group. Therefore, on one hand, such a continuous severity scale would not be a better primary variable candidate for therapies that do not uniformly reduce the incidence along the entire severity scale. On the other hand, therapies that are active uniformly on the complete severity profile of the disease would benefit from choosing a continuous severity scale over a dichotomous variable.

The strength of this study is its prospective design of a randomised clinical trial with well-defined and controlled study variables. The analyses were predefined. For the context of validating the scale, the relatively small sample size in relation to the rare studied outcomes and certain combinations of the scale is a limitation. However, the power calculation was performed before initiation of this substudy's analyses, although it was found to overestimate the observed treatment effect based on the current version of the scale. In the current and the retrospective study, only scores 1–18 were examined due to unavailable data for longer term outcome. Additionally, post-treatment validation is still lacking for the scale. However, having ROP-ActS planned as one of the study's secondary variables, the post-treatment validation will hopefully be feasible in the FIREFLEYE trial that aims to compare the effect of Eylea versus laser therapy.[11]

In this study, we conclude that the MDM study's primary variable, severe ROP, was a better candidate for discrimination of the preventive effect of the AA:DHA supplementation versus no supplementation than the evaluated ROP-ActS, ROP stage, zone and plus disease. The ROP-ActS requires further validation on large data sets with longitudinal follow-up concerning its well ordering and linearity against relevant short-term and long-term outcomes and ROP treatments, and potential update of the scale.

Author affiliations
[1]Department of Clinical Neuroscience, Institute of Neuroscience and Physiology, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden
[2]Mary MacKillop Institute for Health Research, Australian Catholic University, Melbourne, Victoria, Australia
[3]Sahlgrenska Osteoporosis Centre, Institute of Medicine, University of Gothenburg, Gothenburg, Sweden
[4]Department of Ophthalmology, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts, USA
[5]Institute of Health Care Science, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden
[6]Department of Physiology/Endocrinology, Institute of Neuroscience and Physiology, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden
[7]Department of Mathematical Sciences, Chalmers University of Technology, Gothenburg, Sweden
[8]Department of Laboratory Medicine, Institute of Biomedicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

administrative, technical or material support: AH. Drafting of the manuscript: AP. Statistical analyses: AP, HJ, SN and AH.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Consent obtained from parent(s)/guardian(s)

Ethics approval This study involves human participants. This study is a substudy of the randomised controlled Mega Donna Mega trial approved by the regional ethical board at the University of Gothenburg (Dnr 303-11, T570-15). Participants' parents/guardians gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available.

ORCID iDs
Aldina Pivodic http://orcid.org/0000-0002-2224-3236
Ann Hellström http://orcid.org/0000-0002-9259-1244

REFERENCES

1 Mora JS, Waite C, Gilbert CE, *et al*. A worldwide survey of retinopathy of prematurity screening. *Br J Ophthalmol* 2018;102:9–13.
2 Holmström G, Hellström A, Gränse L, *et al*. New modifications of Swedish ROP guidelines based on 10-year data from the SWEDROP register. *Br J Ophthalmol* 2020;104:943–9.
3 Swedish national guidelines for screening and management of ROP 2021, 2022. Available: https://www.medscinet.com/ROP/uploads/ Nationella%20guidelines%20ROP-revidering%2016.12.2021.pdf [Accessed 23 Feb 2022].
4 Swedish national guidelines for screening and management of ROP 2012, 2022. Available: https://neo.barnlakarforeningen.se/wp-content/uploads/sites/14/2014/03/ROP_2012.pdf [Accessed 28 Jan 2022].
5 Gunn TR, Easdown J, Outerbridge EW, *et al*. Risk factors in retrolental fibroplasia. *Pediatrics* 1980;65:1096–100.
6 Hellström ASL, Dammann O. Retinopathy of prematurity. *Lancet* 1991;337:83–4.
7 Norman M, Hallberg B, Abrahamsson T, *et al*. Association between year of birth and 1-year survival among extremely preterm infants in Sweden during 2004-2007 and 2014-2016. *JAMA* 2019;321:1188–99.
8 Hellström A, Hård A-L. Screening and novel therapies for retinopathy of prematurity - A review. *Early Hum Dev* 2019;138:104846.
9 International Committee for the Classification of Retinopathy of Prematurity. The International classification of retinopathy of prematurity revisited. *Arch Ophthalmol* 2005;123:991–9.
10 Smith LEH, Hellström A, Stahl A, *et al*. Development of a retinopathy of prematurity activity scale and clinical outcome measures for use in clinical trials. *JAMA Ophthalmol* 2019;137:305–11.
11 The FIREFLEYE trial. Available: https://clinicaltrials.gov/ct2/show/ NCT04004208?term=NCT04004208&draw=2&rank=1 [Accessed 21 May 2021].
12 Pivodic A, Nilsson S, Stahl A, *et al*. Validation of the retinopathy of prematurity activity scale (ROP-ActS) using retrospective clinical data. *Acta Ophthalmol* 2021;99:201-206.
13 Hellström A, Nilsson AK, Wackernagel D, *et al*. Effect of enteral lipid supplement on severe retinopathy of prematurity: a randomized clinical trial. *JAMA Pediatr* 2021;175:359–67.
14 Niklasson A, Albertsson-Wikland K. Continuous growth reference from 24th week of gestation to 24 months by gender. *BMC Pediatr* 2008;8:8.
15 Walsh MC, Kliegman RM. Necrotizing enterocolitis: treatment based on staging criteria. *Pediatr Clin North Am* 1986;33:179–201.
16 Early Treatment For Retinopathy Of Prematurity Cooperative Group. Revised indications for the treatment of retinopathy of prematurity: results of the early treatment for retinopathy of prematurity randomized trial. *Arch Ophthalmol* 2003;121:1684–94.
17 Ich topic E 9 statistical principles for clinical trials. Available: https:// www.ema.europa.eu/en/documents/scientific-guideline/ich-e-9-statistical-principles-clinical-trials-step-5_en.pdf [Accessed 21 May 2021].
18 Connor KM, SanGiovanni JP, Lofqvist C, *et al*. Increased dietary intake of omega-3-polyunsaturated fatty acids reduces pathological retinal angiogenesis. *Nat Med* 2007;13:868–73.
19 Fu Z, Lofqvist CA, Shao Z, *et al*. Dietary ω-3 polyunsaturated fatty acids decrease retinal neovascularization by adipose-endoplasmic reticulum stress reduction to increase adiponectin. *Am J Clin Nutr* 2015;101:879–88.
20 Löfqvist CA, Najm S, Hellgren G, *et al*. Association of retinopathy of prematurity with low levels of arachidonic acid: a secondary analysis of a randomized clinical trial. *JAMA Ophthalmol* 2018;136:271–7.
21 Bernabe-García M, Villegas-Silva R, Villavicencio-Torres A, *et al*. Enteral docosahexaenoic acid and retinopathy of prematurity: a randomized clinical trial. *JPEN J Parenter Enteral Nutr* 2019;43:874–82.
22 Chiang MF, Quinn GE, Fielder AR, *et al*. International classification of retinopathy of prematurity, third edition. *Ophthalmology* 2021;128:e51–68.