A COHORT STUDY OF TRANSCUTANEOUS OXYGEN TENSION AND THE INCIDENCE AND SEVERITY OF RETINOPATHY OF PREMATURITY*

BY *John T. Flynn,* MD, *Eduardo Bancalari,* MD (BY INVITATION), *Ellen Sim Snyder,* MS (BY INVITATION), *Ronald N. Goldberg,* MD (BY INVITATION), *William Feuer,* MS (BY INVITATION), *Janet Cassady,* MS (BY INVITATION), *Joyce Schiffman,* MS (BY INVITATION), *Harold I. Feldman,* MD (BY INVITATION), *Brian Bachynski,* MD (BY INVITATION), *Edward Buckley,* MD (BY INVITATION), *Jacqueline Roberts,* RN (BY INVITATION), AND (BY INVITATION) *Dennis Gillings,* PhD

INTRODUCTION

ALTHOUGH IT IS LIKELY THAT THE ETIOLOGY OF RETINOPATHY OF PREMATURITY (ROP) is multifactorial,[1] several previous studies have suggested a relationship between ROP and the duration of oxygen exposure.[2,3] These studies related the duration of supplemental inspired oxygen to ROP.[2-4] No specific threshold level of inspired oxygen, however, could be identified. The Second National Cooperative Study of ROP, a multicenter case-control study undertaken in 1969 for the purpose of identifying specific levels of arterial oxygen tension (Pa_{O_2}) that may increase the risk of ROP,

*From the Department of Ophthalmology, Bascom Palmer Eye Institute; the Division of Neonatology, Department of Pediatrics, University of Miami School of Medicine, Miami; the Clinical Epidemiology Unit, University of Pennsylvania School of Medicine, Philadelphia; and the Department of Biostatistics, School of Public Health, University of North Carolina, Chapel Hill. Supported in part by Public Health Service research grants R01-EY03513 and 3PO EY 02180 from the National Eye Institute, Bethesda, Maryland; the Alex J. Weinstein Foundation, New York; Fight for Sight Children's Diagnostic Eye Clinic of the Bascom Palmer Eye Institute through grants from Fight for Sight, Inc, New York, and the Department of Health and Rehabilitative Services, State of Florida, Children's Medical Services and University of Miami Project Newborn.

failed to identify such a relationship, possibly because it utilized only intermittent measurements of Pa_{O_2}.[5]

The present study was part of a larger clinical trial undertaken to evaluate whether the incidence and severity of ROP could be reduced by continuous transcutaneous oxygen monitoring in neonates who required oxygen therapy.[6-8] This study demonstrated that continuous transcutaneous oxygen monitoring may prevent ROP in infants weighing between 1100 and 1300 g at birth, but not in those weighing less than 1100 g. The specific role of hyperoxia in ROP, however, was not clarified by that study. The purpose of the present cohort study was to evaluate the relationship between hyperoxic oxygen exposure assessed by transcutaneous oxygen measurements and the incidence and severity of ROP in the subset of continuously monitored subjects from the previously mentioned clinical trial.

MATERIALS AND METHODS

STUDY POPULATION

All infants born at the University of Miami-Jackson Memorial Hospital Medical Center from November 1, 1982, to May 31, 1984, were eligible for the clinical trial (and therefore this study) if their birth weights were between 500 and 1300 g and they required oxygen in excess of room air at some time during the first 7 days after birth. Infants were excluded from this study if parental consent was withheld, if major congenital anomalies were present, or if the infants were considered nonviable by the attending neonatologist upon admission to the intensive care unit. Infants who became eligible when no transcutaneous monitors were available were also excluded.

A total of 438 infants with birth weights between 500 and 1300 g were born during the study period of the clinical trial. As a result of differences in the mortality observed in historical controls,[9] infants were stratified into two birth-weight groups (500 to 899 g and 900 to 1300 g) and were randomized to either a transcutaneously monitored group or a standard care group. Two hundred ninety-six eligible infants were enrolled: 148 were randomized to the transcutaneously monitored group and 148 to a standard care group. Details of the clinical trial have been published previously.[6-8] The subjects included in this report are the 101 infants in the continuously monitored group who were discharged alive.

MONITORING

Each infant in the study was connected to a transcutaneous oxygen monitor that recorded $tcPo_2$ levels as long as supplemental oxygen was required to maintain Pao_2 above 50 mm Hg. During monitoring, a microprocessor attached to the transcutaneous oxygen monitor recorded minutes in each 10 mm Hg of $tcPo_2$ range (eg, 0 to 9 mm Hg, 10 to 19 mm Hg) for each infant. The number of minutes in each range was then tabulated by week after birth for each infant. The primary measure of hyperoxic exposure for each infant was obtained by tabulating the number of hours $tcPo_2$ was $\geq$ 80 mm Hg during the first 4 weeks of life. The number of hours $tcPo_2$ was $<$ 50 mm Hg during the first 4 weeks of life was tabulated to be used as a measure of hypoxia. Since $tcPo_2$ was not recorded when the infants were not receiving supplemental oxygen, $tcPo_2$ was assumed to be $\geq$ 50 and $<$ 80 mm Hg during the unmonitored periods for purposes of analysis. Arterial Pao_2 was measured intermittently from blood obtained through umbilical or peripheral arterial lines approximately every 4 hours while infants were on mechanical ventilation and less frequently once they were weaned from assisted ventilation.

A research assistant was present in the intensive care unit 24 hours a day and was responsible for calibrating and repositioning the $tcPo_2$ electrode every 2 to 4 hours, and adjusting the Fio_2 or the ventilator in an attempt to maintain the $tcPo_2$ between 50 and 70 mm Hg. The $tcPo_2$ electrode temperature was maintained at 43.5° to 44°C, and the electrode was positioned on the right upper chest during the first day of life in all infants and later for those in whom extrapulmonary right-to-left shunting was suspected. Subsequently, the $tcPo_2$ electrode was placed on the trunk, abdomen, or thighs. Infants with cardiovascular collapse and those whose $tcPo_2$ and Pao_2 values differed by more than 10 mm Hg, however, had their Fio_2 adjusted according to arterial blood gas values rather than $tcPo_2$ values to maintain the Pao_2 between 50 and 70 mm Hg. The $tcPo_2$ electrode was left in place to follow trends, but these values were not used in the analyses.

DIAGNOSIS OF ROP

Ophthalmologic examinations using indirect ophthalmoscopy were performed in all infants by a pediatric ophthalmologist experienced in the examination of premature infants. The examiner was unaware of the infant's identity, results of previous examinations, and treatment group assignment. These examinations were begun when infants reached 32 weeks postconceptional age and were judged to be in stable clinical condition. Examinations were repeated every 2 to 4 weeks until babies were

TABLE I: SEVERITY INDEX*

VARIABLE		DESCRIPTION	WEIGHT (g)
Location	A.	Posterior pole: macula region	18
	B.	Macula to equator	8
	C.	Anterior pole to equator	2
Extent	A.	Full circumference (12 clock hours)	12
	B.	Between 6 and 12 clock hours	8
	C.	Less than 6 clock hours	2
Vascular tortuosity	A.	Extending to disc	3
	B.	Primary branches of arcade	2
	C.	Region of shunt	1
Neovascular membranes	A.	Large: region of shunt, vitreous, over retina	3
	B.	> 1 disc diameter	2
	C.	< 1 disc diameter	1

*A score of < 20 indicated mild disease; a score of 21 to 36 indicated moderate to severe disease.

discharged from the hospital. ROP was diagnosed when one examination demonstrated the minimal diagnostic criteria, including the presence of a structure perpendicular to and located at the tips of developing retinal blood vessels, which divided the vascular from avascular retina.[10] A minimum of two examinations was performed for each infant, one taking place during the 35 to 45 weeks' conceptional age interval.

SEVERITY OF ROP

A severity index outlined in Table I was devised to encompass the most important characteristics of the infant's retinal disease,[11-13] including the location of the disease in the retina; its extent in clock hours; the extent of vascular tortuosity; and the degree of neovascular membrane formation. A numerical score was assigned to each of these components. The sum of these 4 scores constituted the severity index for that examination, with values ranging from 6 to 36. An application of this severity index to the infants in the clinical trial is detailed elsewhere.[8] This index was developed prior to publication of the International Classification of Retinopathy of Prematurity, which is based on the same clinical characteristics.[14,15] The highest severity score attained on any examination was used to categorize infants with ROP into one of two levels of disease: (1) a severity index of ≥ 20 defined moderate to severe ROP, and (2) an index of ≥ 6 and < 20 defined mild disease. An ordinal scale of ROP with three levels

(none, mild, and moderate to severe) served as the primary outcome variable of this study.

CONFOUNDING FACTORS

Infants with long exposures to $tcPo_2$ may have had other clinical characteristics predisposing to ROP, such as a more severe respiratory illness, low Apgar scores, and low birth weight. These characteristics were taken into account in the analysis to control for their potentially confounding influence on the relationship between $tcPo_2$ and ROP. The total duration of supplemental oxygen administered in concentrations exceeding an Fio_2 of 0.4 during each infant's entire hospital stay was used as a measure of the severity of neonatal respiratory illness, not as a measure of oxygen exposure. The 5-minute Apgar score was categorized into two levels (≤ 7 and > 7) and birth weight was measured in grams.

ANALYTIC METHODS

The mean, standard deviation, median, minimum, and maximum were used to describe the hours of $tcPo_2$ exposure ≥ 80 mm Hg or < 50 mm Hg during the first 4 weeks of life and the potentially confounding variables, including birth weight and the hours of Fio_2 required in concentrations ≥ 0.4 during the infant's entire hospital stay. Other indicators of respiratory health (intermittent positive pressure ventilation and number of arterial blood gases) were also described for the study infants, stratified by birth weight. Frequencies were tabulated for categorical variables such as ROP outcome and 5-minute Apgar score. Contingency tables were used to display the unadjusted association between ROP outcome and the $tcPo_2$ exposure variables, which were categorized for this purpose only.

Ordinal logistic regression using the proportional odds model[16-19] was the primary method of analysis that was used to examine the relationship between $tcPo_2$ and ROP incidence and severity. Dichotomous logistic regression[17,19] was used to examine the relationship between ROP incidence (of any severity) and the $tcPo_2$ exposure variables for the subset of infants weighing ≥ 900 g at birth. These methods of analysis allowed the calculation of odds ratio and 95% confidence intervals (CI) from model coefficients and standard errors while adjusting for the potential influence of confounding factors. Unadjusted and adjusted odds ratios were calculated for selected increments in the duration of exposure to specific $tcPo_2$ levels, and results for 12-hour increments are presented.[20] The arbitrary cut-off of < 12 versus ≥ 12 is used in some tables for descriptive purposes only.

PROTOCOL MONITORING

The protocol for this study's parent randomized clinical trial was approved by the Committee for Protection of Human Subjects, University of Miami-Jackson Memorial Hospital Medical Center. An ad hoc data monitoring and safety committee approved the manual of operations and monitored the progress of the study.

RESULTS

ROP developed in 52 (51%) of the 101 infants studied, including 19 of 22 infants (86%) weighing < 900 g. Disease was mild in 10 of the 19 and moderate to severe in 9; 4 of the 9 had cicatricial disease.[21] Thirty-three of 79 infants (42%) weighing ≥ 900 g had ROP; disease was mild in 27 and moderate to severe in 6. No infant weighing ≥ 900 g had cicatricial disease. Birth weight, 5-minute Apgar score, and the duration of supplemental oxygen concentration ≥ 0.4 during the infant's entire hospital stay, stratified by birth weight category and ROP status, are shown in Table II. Infants with ROP weighed less at birth ($P < 0.001$), had lower 5-minute Apgar scores ($P < 0.001$), and had more hours of supplemental oxygen at a concentration ≥ 0.4 during the entire hospital stay than did infants without the disease ($P = 0.029$).

The percentage of infants who received oxygen and were monitored fell from 100% at the beginning of the first week of life to 49% by the beginning of the second week. Beyond the fourth week the number of infants monitored fell to less than 10%. This decline was due to the design of the clinical trial that required infants to be monitored only as long as supplemental oxygen concentration ≥ 0.4 during the infant's entire hospital stay, stratified by birth weight category and ROP status, are shown in Table II. Infants with ROP weighed less at birth ($P < 0.001$), had lower 5-minute Apgar scores ($P < 0.001$), and had more hours of supplemental oxygen at a concentration ≥ 0.4 during the entire hospital stay than did infants without the disease ($P = 0.029$).

The percentage of infants who received oxygen and were monitored fell from 100% at the beginning of the first week of life to 49% by the beginning of the second week. Beyond the fourth week the number of infants monitored fell to less than 10%. This decline was due to the design of the clinical trial that required infants to be monitored only as long as supplemental oxygen or mechanical ventilation was administered. While the transcutaneous electrode was continuously positioned on the infant during monitoring, $tcPo_2$ readings could be classified into a 10-mm Hg range 91% of the time. The failure to classify 9% of "monitored" hours was

TABLE II: DESCRIPTION OF BIRTH WEIGHT, 5-MINUTE APGAR SCORE, AND DURATION OF INSPIRED O_2 > 0.4, STRATIFIED BY ROP AND BIRTH WEIGHT CATEGORY

	BIRTH WEIGHT CATEGORY			
	< 900 g		900-1300 g	
	ROP	NO ROP	ROP	NO ROP
Birth weight (g)				
No. of infants	19	3	33	46
Mean	784	790	1038	1151
SD	73	78	82	113
Minimum-maximum	620-890	740-880	910-1240	940-1300
5-minute Apgar score				
n (%) ≤ 7	12 (63)	0 (0)	27 (82)	16 (35)
n (%) > 7	7 (37)	3 (100)	6 (18)	30 (65)
Hours Fio_2 ≥ 0.4 for entire hospital stay				
Mean	168.4	26.5	71.6	22.0
SD	245.7	21.9	117.6	32.0
Median	104	39	45	2
Minimum-maximum	1-988	1-39	0-648	1-140

SD, standard deviation.

due to the repositioning of the $tcPo_2$ electrode and other technical difficulties. When $tcPo_2$ readings could not be classified and when infants were not receiving supplemental oxygen therapy, $tcPo_2$ was assumed to be ≥ 50 mm Hg and < 80 mm Hg for analysis purposes. Respiratory therapy during the first 28 days after birth for the study infants is described in Table III.

HYPEROXIC EXPOSURE AND ROP

The hours of exposure to $tcPo_2$ ≥ 80 mm Hg during the first 4 weeks after birth, the primary measure of hyperoxic exposure, are described in Table IV for the 101 study infants. These data suggest that an increasing number of hours of exposure to this range of oxygen tension was associated with both incidence and severity of ROP. The $tcPo_2$ ≥ 80 mm Hg exposure data are also displayed in Table V, where oxygen exposure has been categorized into two levels of duration: ≥ 12 hours and < 12 hours. ROP developed in 75% of infants exposed to ≥ 12 hours of $tcPo_2$ ≥ 80 mm Hg during the first 4 weeks of life but in only 26% of infants exposed to < 12 hours of $tcPo_2$ ≥ 80 mm Hg. The association between hyperoxic exposure ($tcPo_2$ ≥ 80 mm Hg) during the first 4 weeks of life and the incidence and severity of ROP was explored using ordinal logistic regression (Table VI). The unadjusted odds ratio derived from an ordinal logistic regression model for a 12-hour increment in the hours of exposure to $tcPo_2$ ≥ 80 mm Hg during the first 4 weeks of life was 3.0 (range, 2.0 to 4.5). After

TABLE III: RESPIRATORY THERAPY DURING THE FIRST 28 DAYS OF LIFE

	BIRTH WEIGHT CATEGORY	
	< 900 g MEDIAN (RANGE)	900-1300 g MEDIAN (RANGE)
No. of infants	22	79
Hours of oxygen therapy	466 (85-659)	89 (0.8-672)
No. of infants requiring intermittent positive pressure ventilation	22	60
Hours of intermittent positive pressure ventilation	236 (59-672)	49 (0-672)
No. of arterial blood gas measurements	69 (38-206)	33 (0-123)

TABLE IV: HOURS ACCUMULATED IN THE HYPEROXIC $tcPo_2$ RANGE (≥ 80 mm Hg) DURING WEEKS 1 TO 4 (n = 101)

	ROP STATUS*		
HOURS $tcPo_2$ ≥ 80 mm Hg	MODERATE OR SEVERE (n = 15)	MILD (n = 37)	NONE (n = 49)
Mean	33.9	18.6	9.8
SD	19.4	11.3	9.9
Median	35.6	16.5	6.3
Range	2-61	1-54	1-50

*P 0.02, pairwise Bonferroni-corrected Mann-Whitney U tests.

TABLE V: FREQUENCY OF ROP SEVERITY BY $tcPo_2$ ≥ 80 mm Hg EXPOSURE DURING WEEKS 1 TO 4 (n = 101)

	SEVERITY OF ROP			
EXPOSURE TO $tcPo_2$ ≥ 80 mm Hg	NO. (%) MODERATE OR SEVERE	NO. (%) MILD	NO. (%) NONE	TOTAL
≥ 12 hr	13 (25)	26 (50)	13 (25)	52
< 12 hr	2 (4)	11 (22)	36 (73)	49

adjusting for birth weight, 5-minute Apgar score, and the duration of administered supplemental Fio_2 ≥ 0.4, this odds ratio remained significant; the adjusted odds ratio (95% CI) was 1.9 (range, 1.2 to 3.0). As dictated by the analytic technique, the statistical significance for this association remained the same no matter what increment was chosen; the odds ratios for incremental exposures bigger than 12 hours were larger

than 1.9, and odds ratio for incremental exposures smaller than 12 hours were smaller than 1.9. The unadjusted and adjusted odd ratios for the confounding factors, birth weight, 5-minute Apgar, and the hours of $Fio_2 \geq 0.4$ are also presented in Table VI. The unadjusted odds ratio for each of these factors was highly significant. Only the odds ratio for the hours of $Fio_2 \geq 0.4$ during the infant's entire hospital stay lost significance after controlling for the other variables in the model.

TABLE VI: UNADJUSTED AND ADJUSTED ODDS RATIOS (95% CI) FROM ORDINAL LOGISTIC REGRESSION MODEL (n = 101)

	ODDS RATIO (95% CI)	
	UNADJUSTED	ADJUSTED*
$tcPo_2 \geq 80$ mm Hg weeks 1-4 (12-hr increment)	3.0 (2.0-4.5)	1.9 (1.2-3.0)
Birth weight (100-g decrement)	2.4 (1.8-3.2)	2.3 (1.6-3.4)
5-Minute Apgar score (≤ 7 *vs* > 7)	5.3 (2.3-12.2)	7.2 (2.5-21)
Hours $Fio_2 \geq 0.4$ during entire hospitalization (12-hr increment)	1.4 (1.1-1.8)	1.0 (0.97-1.05)

*Adjusted for the other variables in this table.

An ancillary analysis considered the exposure to $tcPo_2 \geq 80$ mm Hg during the first week after birth separately from $tcPo_2 \geq 80$ mm Hg during weeks 2 through 4. Exposure during these two periods was treated as two separate variables because of the possibility that $tcPo_2$ values recorded during the first week of life in infants with unstable hemodynamic function may not have reflected the true Pao_2 levels as well as those recorded during weeks 2 through 4. With this approach, the unadjusted odds ratio for week 1 was 2.3 (range, 1.2 to 4.4) for a 12-hour increment in exposure to $tcPo_2 \geq 80$ mm Hg. After controlling for birth weight, 5-minute Apgar score, and total hours of $Fio_2 \geq 0.4$, $tcPo_2 \geq 80$ mm Hg during week 1 was no longer an important predictor of ROP: the adjusted odds ratio for a 12-hour increment was 1.0 (range, 0.5 to 2.3). The unadjusted odds ratio for a 12-hour increment in exposure to $tcPo_2 \geq 80$ mm Hg during weeks 2 through 4, however, was 2.3 (range, 1.2 to 4.3). This association became even stronger after controlling for the confounding factors and for the number of hours of exposure to $tcPo_2 \geq 80$ mm Hg during week 1, with the odds ratio (95% CI) rising to 3.1 (range, 1.6 to 6.1).

TABLE VII: FREQUENCY OF ROP BY HOURS OF $tcPo_2 \geq 80$ mm Hg IN WEEKS 1 TO 4 IN SUBSET OF INFANTS WEIGHING 900 TO 1300 G ($n = 79$)

EXPOSURE TO $tcPo_2 \geq 80$ mm Hg	ROP STATUS		TOTAL
	ROP NO. (%)	NO ROP NO. (%)	
≥ 12 hr	22 (67)	11 (33)	33
< 12 hr	11 (24)	35 (76)	46

An analysis of the subgroup of infants weighing ≥ 900 g was performed using dichotomous logistic regression to examine the relationship between the incidence of ROP (a diagnosis of ROP versus no ROP) and $tcPo_2 \geq 80$ mm Hg (Table VII). This analysis was not performed in the smaller infants (< 900 g), since ROP developed in almost all (19 of 22 [86%]) of these infants. Infants with ROP in the ≥ 900-g stratum accumulated more hours in the hyperoxic $tcPo_2$ range (19.6 ± 14.0) than did infants without ROP (8.7 ± 7.7) in this stratum ($P = 0.0001$). The unadjusted odds ratio for a 12-hour increment in hyperoxic exposure during the first 4 weeks after birth for ≥ 900-g infants was 3.3 (range, 1.7 to 6.4) ($P < 0.001$). However, the odds ratio was reduced to 1.6 (range, 0.7 to 3.5) after controlling for the following potentially confounding factors: birth weight, low 5-minute Apgar score (≤ 7), and total duration of supplemental oxygen therapy with $Fio_2 \geq 0.4$. Although this is not statistically significant, it does not rule out a clinically important effect.

HYPOXIC EXPOSURE AND ROP

The description of hypoxic exposure ($tcPo_2 < 50$ mm Hg) among the ROP categories is presented in Table VIII. Infants with ROP had more hours of hypoxia than infants without ROP ($P < 0.01$). The association between hypoxic $tcPo_2$ exposure during the first 4 weeks after birth and the

TABLE VIII: HOURS ACCUMULATED IN THE HYPOXIC $tcPo_2$ RANGE (< 50 mm Hg) DURING WEEKS 1 TO 4 ($n = 101$)

HOURS $tcPo_2 < 50$ mm Hg	ROP STATUS*		
	MODERATE TO SEVERE ($n = 15$)	MILD ($n = 37$)	NONE ($n = 49$)
Mean	78.0	64.8	21.7
SD	58.2	59.9	37.1
Median	78.4	38.9	8.5
Range	0-161	0-181	0-151

*$P < 0.01$, pairwise Bonferroni-corrected Mann-Whitney U tests.

incidence and severity of ROP was small but significant. The unadjusted odds ratio (95% CI) for a 12-hour increment was 1.2 (range, 1.1 to 1.3). This association did not remain significant, however, when the confounding influence of birth weight, 5-minute Apgar score, and the hours of $Fio_2 \geq 0.4$ during the infant's entire hospital stay were controlled in the analysis; the adjusted odds ratio (95% CI) was 1.0 (range, 0.9 to 1.1). The examination of exposure to $tcPo_2 < 50$ mm Hg during the first week of life separately from the second through fourth weeks of life also did not reveal any association with ROP after controlling for the three confounding factors and hours of hypoxia during the first week of life. Unlike hyperoxic exposure, hypoxic retinal exposure did not independently predict the incidence and severity of ROP.

DISCUSSION

This study presents evidence for a relationship between prolonged exposure to hyperoxic arterial blood ($tcPo_2 \geq 80$ mm Hg) and the incidence and severity of ROP in infants weighing ≤ 1300 g at birth and surviving at least 28 days. This relationship could not be attributed to the confounding influence of birth weight, Apgar score, or the duration of administered supplemental oxygen.

While these results are consistent with the literature implicating prolonged oxygen administration as one of the causes of ROP, they specifically demonstrate a relationship between duration of hyperoxia and ROP. Several studies in the 1950s have shown that premature infants treated with oxygen regimens of longer durations and/or higher concentrations of inspired oxygen had significantly more ROP than controls.[1-3] The Second National Cooperative Study of ROP,[5] which took place between 1969 and 1972, was designed to identify the relationship of specific Pao_2 levels and ROP. While that study was unable to demonstrate a relationship between Pao_2 levels and ROP, it did confirm that longer durations of oxygen therapy were associated with the development of ROP in infants weighing < 1200 g at birth. Finally, while the clinical trial portion of this study[6-8] detected no overall effect of $tcPo_2$ monitoring, analysis of the subset of infants weighing ≥ 1100 g at birth showed a significant reduction of the incidence of ROP in the monitored group. A notable survival difference between the continuously monitored and standard care groups, although not statistically significant, may have contributed in part to these results. Assuming that the difference in ROP incidence was not due to differential survival across treatment groups and that the continuously monitored group remained in the target range of $tcPo_2$, 50 to 70 mm Hg, for a

greater proportion of the time than did the unmonitored group, these data represent the most recent evidence implicating oxygen exposure in the pathogenesis of ROP. This study, however, did not clarify if the protective mechanism of $tcPo_2$ monitoring was through avoidance of hyperoxic or hypoxic exposures.

The current study, which emanates from that clinical trial, supports the role of oxygen in the pathogenesis of ROP and implicates high blood oxygen tensions in the development and severity of this disease. Blood oxygen levels were measured by intensive transcutaneous monitoring in this study, in contrast with previous studies, which employed only intermittent Pao_2 measurements using arterial blood gases. This difference in the method of blood oxygen tension ascertainment may explain why the previous studies were unable to detect such an association.

The well-described incidence of ROP in infants receiving no supplemental oxygen and the lack of disease in babies exposed for long durations demonstrates that high oxygen levels are neither necessary nor sufficient to induce retinal disease.[1] These observations implicate the role of factors other than oxygen in the pathogenesis of ROP. Associations with ROP were confirmed in this study for birth weight, 5-minute Apgar score, and severity of respiratory failure as measured by hours of $Fio_2 > 0.4$ as well as hours of Fio_2 greater than room air. Only birth weight and 5-minute Apgar score, however, displayed associations with ROP that were independent of $tcPo_2 \geq 80$ mm Hg. The number of hours of supplemental oxygen therapy may have been a less powerful predictor of ROP in this study than has been found in previous studies, if infants currently receive less supplemental oxygen therapy than was true in the 1950s and 1970s, when these other studies were conducted. Factors in addition to birth weight, Apgar score, and the duration of supplemental oxygen therapy have also been identified in the literature.[1] Although the confounding influences of one or more of these other factors may have influenced the results of this study, most of them are highly correlated with birth weight, 5-minute Apgar score, and hours of supplemental oxygen therapy and therefore were controlled by the analytic techniques used here.

The use of duration of $Fio_2 > 0.4$ as a confounding variable theoretically could have led to underestimates of the relationship between $tcPo_2 \geq 80$ mm Hg and ROP, because high levels of inspired oxygen are, in part, directly responsible for elevated $tcPo_2$. To examine this possibility, a second set of multivariate models was constructed not including the duration of the $Fio_2 \geq 0.4$ as a covariate. The results of these analyses were virtually identical to those derived from analyses incorporating Fio_2. This study, therefore, contributes to the body of literature indicat-

ing that high $tcPo_2$ levels independently promote the incidence and severity of ROP.

A precise threshold of $tcPo_2$ in the hyperoxic range that was toxic to the retina could not be established. Since $tcPo_2$ exposure was recorded in 10-mm Hg intervals, hyperoxic exposure could only be defined as, for example, $\geq$ 70 mm Hg, $\geq$ 80 mm Hg, $\geq$ 90 mm Hg. We were unable to explore the independent influence of exposure in the range $\geq$ 90 mm Hg on the incidence and severity of ROP because of the paucity of hours infants were exposed to this range. Analyses using hours $\geq$ 70 mm Hg for the definition of hyperoxic exposure, however, did demonstrate the same associations as those using $\geq$ 80 mm Hg, but were of smaller magnitude. Infants weighting $<$ 900 g had the greatest incidence of ROP and were exposed to greater durations of hyperoxia than infants weighing $\geq$ 900 g at birth. The small number of infants weighing $<$ 900 g (n = 22), however, did not permit the estimation of an adjusted odds ratio relating hyperoxia to the severity of ROP specific to this birth weight subgroup. While the size of the $\geq$ 900 g subgroup (n = 79) did permit a separate analysis, the small number of these larger infants with moderate to severe disease did not permit the analysis of ROP measured on an ordinal scale. The adjusted odds ratio relating a 12-hour increment in hyperoxic exposure to ROP of any severity in infants weighing $\geq$ 900 g was similar in magnitude to the analysis of all babies using the ordinal severity measure of ROP (1.6 *vs* 1.9, respectively), but it did not reach conventional statistical significance. This analysis of the $\geq$ 900 g subgroup using a dichotomous ROP outcome variable, however, was of substantially reduced statistical power.

This study was unable to demonstrate an association between hypoxic retinal oxygen exposure, defined by a transcutaneously measured oxygen tension $<$ 50 mm Hg, and ROP. Previously, Yu and associates[4] were also unable to implicate hypoxia as a cause of ROP in a historical cohort study of premature infants. In contrast to both of these findings, animal and human studies have suggested an important role of hypoxia in ROP. Phelps and Rosenbaum[22,23] reported that kittens exposed to hypoxic environments following hyperoxic exposures had more retinal vascular disease than kittens spared an hypoxic insult. Cantolino and associates[24] reported a relationship between hypoxemia and retinal arterial constriction, photographically documented, in newborns. Katzman and colleagues[25] and Shohat and co-workers[26] independently found that hypoxia, defined as $Pao_2 < 50$ mm Hg, was associated with the severity of ROP. Our study's inability to detect a relationship between hypoxia and ROP, therefore, must be viewed cautiously, especially in view of the probable undercounting of hypoxemia episodes during unmonitored periods.

The results of this observational study must be interpreted in light of several constraints. This study was not a randomized clinical trial of levels of hyperoxic exposure on the occurrence and severity of ROP, because such a trial would be unethical in our opinion. Several other constraints arise from the design of the larger clinical trial. No monitoring took place while infants were not receiving supplemental oxygen. The $tcPo_2$ was assumed to be $\geq$ 50 mm Hg and $<$ 80 mm Hg during these unmonitored periods. Unobserved episodes of hypoxemia may have contributed to undercounting hours at $tcPo_2 < 50$ mm Hg and, therefore, to the lack of association between hypoxic exposures and ROP. Undercounting of hyperoxic exposure ($tcPo_2 \geq 80$ mm Hg), also a theoretical concern, was much less likely, because infants received no supplemental oxygen during unmonitored periods. During monitoring, 9% of hours, on average, were not classifiable into a $tcPo_2$ range. The $tcPo_2$ during these hours was also assumed to be $\geq$ 50 mm Hg and $<$ 80 mm Hg for analysis purposes. Unlike unmonitored hours, however, it is conceivable that a proportion of these unclassified hours was actually in the hyperoxic range. This potential misclassification may have reduced the magnitude of the observed relationship between the incidence and severity of ROP and $tcPo_2 \geq 80$ mm Hg, because undercounting of these hours was more likely for infants with (more severe) ROP.

Because of the frequency of ophthalmologic examinations (every 2 weeks when technically feasible) and the difficulty of examining premature newborn eyes, neither the onset nor the timing of maximally severe ROP was possible to determine precisely. The period of risk during which oxygen may influence the development of ROP, however, was assumed to be before the fifth week of life. This definition was based on the observed natural history of ROP in the newborn period. It has been estimated that by the end of the third week of life, only 1.7% of premature newborns weighing $\leq$ 1300 g have diagnosable ROP. By the end of the fourth week, this percentage rises to approximately 3%.[8] Despite these observations, the uncertainty of the timing of ROP may have led to the inclusion of hours of exposure to oxygen after maximally severe ROP had already developed, the effect of which may have been to either inflate or diminish the apparent effect of $tcPo_2$ exposure. However, the analysis presented in this report based on incremental hours of exposure, rather than on exposures greater than threshold durations of $tcPo_2$, should have minimized the influence of this potential misclassification.

The results of this study rely on the validity of the assumptions underlying the ordinal logistic regression model, which were supported by these data. This technique enabled pooling of information on infants in both

birth weight subgroups, despite the fact that virtually all of the infants weighing < 900 g had ROP. In addition, it also permitted the control of important confounders in the analysis.

CONCLUSIONS

This study provides the first direct and quantitative link between arterial oxygen tension as measured by $tcPo_2$ monitoring and the incidence and severity of ROP. Durations of exposure to $tcPo_2$ values over 80 mm Hg appears to be a clinically important and significant variable affecting the incidence and severity of this disease.

REFERENCES

1. Lucey JF, Dangman B: A reexamination of the role of oxygen in retrolental fibroplasia. *Pediatrics* 1984; 73:82-96.
2. Kinsey VE: Retrolental fibroplasia: Cooperative study of retrolental fibroplasia and the use of oxygen. *Arch Ophthalmol* 1956; 56:481-543.
3. Lanman JT, Guy LP, Dancis J: Retrolental fibroplasia and oxygen therapy. *JAMA* 1954; 155:223-226.
4. Yu VY, Hookham DM, Nave JRM: Retrolental fibroplasia: Controlled study of 4 years' experience in a neonatal intensive care unit. *Arch Dis Child* 1982; 57:247-252.
5. Kinsey VE, Arnold HJ, Kalina RE, et al: Pao_2 levels and retrolental fibroplasia: A report of the cooperative study. *Pediatrics* 1977; 60:655-668.
6. Bancalari E, Flynn J, Goldberg RN, et al: Influence of transcutaneous oxygen monitoring on the incidence of retinopathy of prematurity. *Pediatrics* 1987; 79:663-669.
7. Flynn JT, Bancalari E, Bawol R, et al: Retinopathy of prematurity: A randomized, prospective trial of transcutaneous oxygen monitoring. *Ophthalmology* 1987; 94:630-638.
8. Flynn JT, Bancalari E, Bachynski BN, et al: Retinopathy of prematurity: Diagnosis, severity, and natural history. *Ophthalmology* 1987; 94:620-629.
9. Flynn JT: Acute proliferative retrolental fibroplasia: Multivariate risk analysis. *Trans Am Ophthalmol Soc* 1983; 81:549-591.
10. ———: Notes on a classification of acute proliferative retrolental fibroplasia (retinopathy of prematurity), in *Retinopathy of Prematurity Conference Syllabus*. Washington, DC, Ross Laboratories, 1981, Vol 1, pp 247-253.
11. Uemura Y: Current status of retrolental fibroplasia: Report of the joint committee for the study of retrolental fibroplasia in Japan. *Jpn J Ophthalmol* 1977; 21:366-378.
12. Tanabe Y, Ikema M: Retinopathy of prematurity and photocoagulation therapy. *Acta Soc Ophthalmol Jpn* 1972; 76:260-266.
13. Quinn GE, Schaffer DB, Johnson L: A revised classification of retinopathy of prematurity. *Am J Ophthalmol* 1972; 94:744-749.
14. The Committee for the Classification of Retinopathy of Prematurity: An international classification of retinopathy of prematurity. *Arch Ophthalmol* 1984; 102:1130-1134.
15. The International Committee for the Classification of the Late Stages of Retinopathy of Prematurity: An international classification of retinopathy of prematurity: II. The classification of retinal detachment. *Arch Ophthalmol* 1987; 105:906-912 [Erratum: *Arch Ophthalmol* 1987; 105:1498].
16. Walker SH, Duncan DB: Estimation of the probability of an event as a function of several independent variables. *Biometrika* 1967; 54:167-179.

17. McCullagh P, Nelder JA: *Generalized Linear Models*, 2nd ed. London, Chapman and Hall, 1989, pp 151-155.
18. Schlesselman JJ: *Case-Control Studies: Design, Conduct, Analysis*. New York, Oxford University Press, 1982, pp 227-230.
19. Agresti A: *Analysis of Ordinal Categorical Data*. New York, John Wiley & Sons, 1984, pp 104-132.
20. Kleinbaum DG, Kupper LL, Morgenstern H: *Epidemiologic Research: Principles and Quantitative Methods*. Belmont, CA, Lifetime Learning Publications, 1982, pp 419-491.
21. Reese AB, King MJ, Owens WC: A classification of retrolental fibroplasia. *Am J Ophthalmol* 1953; 36:1333-1335.
22. Phelps DL, Rosenbaum AL: Effects of marginal hypoxemia on recovery from oxygen-induced retinopathy in the kitten model. *Pediatrics* 1984; 73:1-6.
23. Phelps DL: Reduced severity of oxygen-induced retinopathy in kittens recovered in 28% oxygen. *Pediatr Res* 1988; 24:106-109.
24. Cantolino SJ, O'Grady GE, Herrera JA, et al: Ophthalmoscopic monitoring of oxygen therapy in premature infants: Fluorescein angiography in acute retrolental fibroplasia. *Am J Ophthalmol* 1971; 72:322-331.
25. Katzman G, Satish M, Krishnan V, et al: Comparative analysis of lower and higher stage retrolental fibroplasia. *Pediatr Res* 1982; 16:294A.
26. Shohat M, Reisner SH, Krikler R, et al: Retinopathy of prematurity: Incidence and risk factors. *Pediatrics* 1983; 72:159-163.

DISCUSSION

DR ARNALL PATZ. It has been well established in studies by Kinsey and co-workers in the 1950s and 1960s that the duration of exposure to oxygen was related to the development of ROP. Doctor Flynn and colleagues, utilizing transcutaneous oxygen tension monitoring, have documented for the first time the association of oxygen tension levels with the development of ROP. The authors used a well-designed, prospective nursery study with highly sophisticated and appropriate statistical analysis and interpretation. They found a significant association between incremental exposure to transcutaneous oxygen tensions $\geq$ 80 mm Hg during the first 4 weeks of life and both the incidence and severity of ROP.

Transcutaneous oxygen monitoring used by the authors provided an essentially continuous recording of the level of oxygenation. They pointed out that in the earlier studies the failure to demonstrate a relationship between arterial oxygen blood levels and ROP incidence could have resulted from the inadequate data obtained from blood gas measurements at 6-hour intervals between blood sampling.

Phelps reported in 1980 (*J Pediatr* 1991; 1:7-17) that the survival rate of premature infants with birth weights under 1000 g was approximately 8% in 1950 and that by 1980 the survival rate had increased to 35%. Estimates for 1990 project the survival rate of these very small infants to be approximately 75%. The medical community is now faced with an almost tenfold increase in survival of these very high-risk infants, and as a result we are now seeing a significant number of new cases of ROP.

The present study is therefore most timely and of considerable importance in light of the markedly improved survival rates of the smallest-birth-weight infants,

those with the most immature retinal vasculature and at the highest risk for developing ROP.

Animal studies conducted during the early 1950s resulted in the development of an experimental model of retinal neovascularization that had many similarities to the early stages of ROP. The animals studies revealed that the incompletely vascularized retina was generally susceptible to increased oxygen administration. As the retinal vessels matured with the vessels approaching the retinal periphery, the response to oxygen progressively decreased. The very-low-birth-weight premature infant's susceptibility to ROP may be related to the immaturity of retinal vascularization.

The authors are to be congratulated on their very important contribution, which documents for the first time the role of oxygen tension levels in the development of ROP. Doctor Flynn and colleagues have demonstrated that transcutaneous monitoring of blood oxygen tension levels provides an extremely valuable tool for safer oxygen administration to the premature infant. When properly applied, their findings will certainly have a significant impact on the prevention of this important cause of blindness in the premature infant.

DR ROBERT DREWS. I would also like to congratulate Doctor Flynn and his group on a superb and fundamental study in retinopathy of prematurity.

In 1951 in the *American Journal of Ophthalmology* a special article appeared by Szewczyk calling the world's attention to oxygen as a cause of retinopathy of prematurity. It was his hypothesis and his experience, however, that babies kept at high levels of oxygen could avoid, in many instances, retinopathy of prematurity if the oxygen level did not *vary*. I wonder if you could look at your data to see if variation in oxygen level was a factor as well as the total exposure to oxygen.

DR LEONARD APT. I would like Doctor Flynn's comments on several issues that primarily concern instrumentation of his study.

To my knowledge transcutaneous oxygen monitoring is now rarely used in most newborn intensive care units for several reasons. First, the procedure can cause severe skin burns, especially in prematurely born infants whose skin normally is extremely thin and delicate. This complication can result because the transcutaneous monitor requires extraneous heating to dilate dermal capillaries and to thus arterialize the oxygen reading.

Second, the PO_2 values obtained on skin oximetry may be erroneous on a number of accounts, including variation in skin thickness over different parts of the body, presence of edema, the effect of certain drugs, and the development of hypothermia or shock-like states.

Perhaps Doctor Flynn will tell us how he avoided serious skin burns on the babies he studied. Also, would he say how he dealt with the many factors that could affect the accuracy and dependability of oxygen assessment in the infants in his study?

I may point out that in recent years conjunctival oxygen monitoring has been introduced to measure tissue oxygen tension. Doctor Sherwin Isenberg at our

institute has had an interest in this technique. Monitoring the conjunctiva has a number of advantages over transcutaneous oxygen monitoring. A heated oxygen electrode is not needed, the conjunctival vessels are close to the surface and are not covered by a layer of oxygen-consuming tissue. Furthermore, the conjunctiva lacks variable thickness, is unkeratinized, is only a few cells thick, and therefore is approximately equal to peripheral tissue PO_2. Finally, conjunctival oxygen tension ($P_{cj}O_2$) has been found to correlate directly with arterial oxygen tension in uncompromised patients.

Since transconjunctival oxygen monitoring deals with vascularization of tissue belonging specifically to the organ with which ophthalmologists are concerned, it seems logical that this technique at present is the preferred one to measure tissue oxygen. Doctor Flynn, have you had any experience with transconjunctival oxygen monitoring in the study of retinopathy of prematurity?

DR JOHN T. FLYNN. I would like to thank my discussants. I thought this paper would end up with a single discussant following that interesting cataract paper that had about 20 discussants.

First of all, regarding the point which Doctor Patz raised, why week 1 *vs* week 2 through 4? The reason I think is the fact that if you look at the profile of monitored babies, we had 100% of babies in 1 week whether they were going to get ROP or not, who were all on supplemental oxygen and, therefore, on monitors. There will then be a smaller cohort (about 50% to 60%) who are sick and they have respiratory disease over long periods of time. These are the infants whose oxygen tension is undulating over time, because they literally are exposed to the hundred and hundreds of supplemental oxygen, is undulating over time. They are getting small exposures to accumulated risk of hyperoxia and ROP.

Regarding the question that Doctor Drews raised, the hypoxia issue is a critical one. We looked at hypoxia in this study but I took it out of the talk because, to understand its significance, you have to go through a very complicated analysis. These babies accumulated an enormous amount of hours of what we arbitrarily defined as hypoxia during their stay. Yet when we compared statistically whether that stood up against their confounding variables of low birth weight and low Apgar scores and long, continued supplemental, it (hypoxia) always washed out as a risk factor. To answer the question of hypoxia thoroughly, Dale Phelps, MD, a neonatologist from Rochester, New York, is organizing a randomized clinical trial that will put babies who develop retinopathy of prematurity and reach a level of severity of that disease, back in low grades of oxygen. This is based on the theory that the oxygen insult produces the disease and infants who come out of oxygen are relatively hypoxic. This hypoxia is the stimulus to the tremendous neovascularization that then develops. Nothing would delight me more than to see Doctor Szewczyk's work, these many years later, confirmed. That was his hypothesis and the oxygen exposure story sort of brushed it off the stage. In my conception, oxygen insult plus hypoxia exposure are the necessary but not sufficient causes to account for the full-blown picture of retinopathy of prematurity.

Regarding the points that Doctor Apt raised, we did have complications from Tc

monitoring. The first seven babies, were who on constant monitoring, died. We convened an emergency meeting of our Data Monitoring and Safety Committee to ask, should we stop the study? They said no, go on with it. We paid strict attention to the problems of asepsis, to the problems of the burn, which occurs in these very tiny babies with almost translucent skin, and we found that with exquisite care these complications disappeared. It is true that transcutaneous oxygen monitoring does not always accurately reflect arterial oxygen levels at the aorta or the carotids. When that was the case, we did not accept the data, but put these infants on every 2-hour arterial blood gas analysis and used that data instead. I have no experience with the transconjunctival monitoring mentioned by Doctor Apt but, seemingly, that would be a way to go if the study is ever going to be replicated.

Thank you very much.