Neuroblastoma Screening at 1 Year of Age: The Final Results of a Controlled Trial

Frank Berthold (iD), MD,[1,*,†] Claudia Spix (iD), PhD,[2,†] Rudolf Erttmann, MD,[3] Barbara Hero (iD), MD,[1] Joerg Michaelis, MD,[2] Joern Treuner, MD,[4] Angela Ernst (iD), MSc,[5,‡] Freimut H. Schilling, MD[6,‡]

[1]Department of Pediatric Oncology and Hematology, University of Cologne, Cologne, Germany; [2]Division of Childhood Cancer Epidemiology, Institute of Medical Biostatistics, Epidemiology and Informatics, University Medical Center, Mainz, Germany; [3]Clinic of Pediatric Oncology and Hematology, University of Hamburg, Hamburg, Germany; [4]Child and Adolescent Health, Pediatrics 5, Olgahospital, Stuttgart, Germany; [5]Institute of Medical Statistics and Bioinformatics, University of Cologne, Cologne, Germany; and [6]Department of Pediatric Oncology Hematology Children's Hospital, Cantonal Hospital, Luzern, Switzerland

[†]F. B. and C. S. contributed equally to this manuscript.

[‡]A. E. and F. S. contributed equally to this manuscript.

*Correspondence to: Frank Berthold, MD, Pediatric Oncology and Hematology, University of Cologne, Kerpener Straße 62, D50924 Cologne, Germany (e-mail: frank.berthold1@uk-koeln.de).

Abstract

Background: Neuroblastoma screening aims to reduce neuroblastoma-related mortality. A controlled trial showed no reduction in stage 4 disease incidence and preliminary mortality data. This article presents epidemiologic and clinical data 20 years after cessation of the screening program. **Methods:** The patients with detected disease in the screening area were compared with the clinically diagnosed patients in the control area and in the prestudy and poststudy cohorts. All statistical tests were 2-sided. **Results:** The cumulative incidence for children aged 1 to 6 years in the birth study cohorts (1994-1999) in the screening arm was 13.4 cases per 100 000 births (95% confidence interval [CI] = 12.2 to 14.6) based on 61.2% of screening participants and 38.8% of nonparticipants. Screening participants had a cumulative incidence of 15.7 (95% CI = 14.0 to 17.4) per 100 000 births. The cumulative incidence in the contemporary control cohort was 9.3 (95% CI = 8.2 to 10.3) per 100 000 births, 7.6 (95% CI = 6.8 to 8.4) in the prestudy cohort, and 8.1 (95% CI = 7.4 to 8.9) in the poststudy cohort from 2000 to 2004 (P < .001 each). The increased incidence in the screening cohort was restricted to stages 1 through 3, while stage 4 incidence was not reduced. The cumulative mortality for deaths within 10 years from diagnosis and per 100 000 births remained unchanged. Patients with stage 4 disease detected by screening had better biological characteristics and an improved outcome compared with those stage 4 cases not detected by screening. **Conclusions:** Neuroblastoma screening at 1 year of age reduced neither stage 4 incidence nor neuroblastoma mortality and was affected by overdiagnosis, leading to unnecessary treatment. A few screening-detected stage 4 cases represent a biologically interesting subgroup but do not change the recommendation to close the "catecholamine-based neuroblastoma screening book."

The aim of detecting cancer while it is in its presymptomatic phase is to prevent the development of disease dissemination. In neuroblastoma, metastatic stage 4 (International Neuroblastoma Staging System) disease (1) typically presents later in life than regional stage 3 disease, and stage 3 disease presents later than localized stage 1 and 2 disease (2), suggesting a potential stepwise progression. The 10-year overall survival (OS) for patients with stage 4 disease has improved but is still less than 50% (3); therefore, reduction in stage 4 incidence is a worthwhile aim. The existence of catecholamine

metabolites in urine as highly tumor-specific markers has made a biochemical screening of diaper specimens possible.

Japanese colleagues pioneered such a mass screening in 6-month-old infants in 1974 (4). During the late 1980s and 1990s, many similar studies followed—reviewed by Woods (5) and the PDQ Pediatric Treatment Editorial Board (6)—but only 2 studies had concurrent control cohorts and an epidemiologic approach: the Quebec study, which screened children between 3 weeks and 6 months of age (7), and the German study, which screened children between 9 and 18 months of age (8). As a result of these

studies, the neuroblastoma screening programs were stopped in Canada, the United States, Germany, and Japan (9).

The ability of neuroblastoma to spontaneously regress, resulting in more or less complete disappearance of the tumor without any interventions in infants with stage 4S disease (10) and stage 1 to 3 disease (11,12) demonstrates the complex biology of these tumors and points to the interference with screening programs, in particular during infancy (13). Comparing birth cohorts in Japan of children born during and after the neuroblastoma mass screening program, the cumulative incidence rate dropped for children up to 5 years of age from 15.6 to 33.8 per 100 000 (1993-2003) to 6.6 to 11.6 (2004-2011), while the cumulative mortality rate showed no difference between the 2 cohorts (14). No recurrences or progressions to malignant transformation or metastasis were observed for patients with screening-detected disease in a follow-up study for up to 15 years (9). In Canada, implementation of the screening program led to a transient increased awareness and incidence for 3 years after the screening but disappeared in years 4 and 5 (15). To our knowledge, long-term analyses of epidemiologically controlled neuroblastoma screening programs do not exist.

We investigated the incidence, stage-related overdiagnosis, tumor- and toxicity-related mortality, potential benefit of subcohorts, and survival more than 20 years after cessation of the German screening program. The patients in the federal states with screening (screening cohort) were compared with those in federal states that had no screening program (control cohort) and with the pre- and poststudy cohorts.

Methods

Study Design

The study design is described elsewhere (8). In short, urinary screening for elevated catecholamine metabolites was offered in 6 German states to children at the routine medical health care check-up at 1 year of age (range = 9-18 months) on a voluntary basis; 61% participated. For the study, the children born between July 1, 1994, and October 31, 1999, were eligible (8). The study area covered roughly half of the German population. The remaining 10 states served as the control area (controlled trial for incidence and mortality). The areas were selected for practical reasons and comparable regarding all disease-related indicators before the study. Diagnostics and treatment remained centralized and comparable for all patients throughout the study. No randomization procedure was performed.

The screening study analyses were restricted to the birth cohorts (1994-1999). Because of the high frequency of cases in early infancy and the study's screening time point at 1 year of age, patients diagnosed during the first year of life were excluded from the analyses. Because more than 90% of cases present within the first 6 years of life and the lead-time effect of the screening is unlikely to occur afterwards, patients aged older than 71 months were also excluded from the analyses (16).

Definitions of screening false-negatives, overdiagnosis, toxic death, incidental diagnosis, and the acquisition of clinical data are outlined in the Supplementary Methods (available online).

Statistical Analyses

The main endpoint of the screening study was mortality. The original study aim was a 50% reduction in the screened group.

The secondary endpoint was the incidence rate of stage 4 disease as a necessary prerequisite for a decrease in mortality and thus a potential early surrogate. The statistical tests (χ^2 tests) were based on the assumption of a Poisson distribution at a 2-sided statistical significance level of 5%. All comparisons were restricted to children aged 12 to 71 months.

Results

Patients

Of the 1 475 773 screened children in the aforementioned birth cohorts and age windows, 1605 had false-positive results and 149 were diagnosed with neuroblastoma after a positive test (8). Detailed clinical information was available for 146 patients; 80 screened children in the screening area who had a negative result presented until 71 months of age with neuroblastoma. Of those patients, 20 showed normal catecholamine excretion at clinical diagnosis (7 of 10 with stage 1 disease; 4 of 9 with stage 2 disease; 4 of 14 with stage 3 disease; 5 of 47 with stage 4 disease; all 20 of 80 [25.0%]). The control cohort had a comparable proportion of cases with nonelevated catecholamine excretion (10 of 26 with stage 1 disease; 5 of 18 with stage 2 disease; 13 of 29 with stage 3 disease; 12 of 102 with stage 4 disease; all stages 40 of 175 [22.9%]). Among the unscreened children in the screening area, 110 presented with neuroblastoma (the same birth cohort, the same age group).

Incidence

The cumulative incidence for children aged 1 to 6 years in the study's birth cohorts (1994-1999) in the screening area was 13.4 cases per 100 000 births (95% confidence interval [CI] = 12.2 to 14.6) based on 61.2% of participants screened and 38.8% nonparticipants. This figure is 44.1% higher compared with the control cohort (9.3 per 100 000 births, *P* < .001) (Table 1). The cumulative incidence within the group of screening participants was even higher, with a cumulative incidence of 15.7 (95% CI = 14.0 to 17.4) per 100 000 births. The cumulative incidence in the contemporary control cohort was 9.3 (95% CI = 8.2 to 10.3) per 100 000 births, 7.6 (95% CI = 6.8 to 8.4) in the prestudy cohort, and 8.1 (95% CI = 7.4 to 8.9) in the poststudy cohort from 2000 to 2004 (*P* < .001 each). (Supplementary Table 1, available online).

This increase was the result of the cases detected by screening that were mostly diagnosed in the second year of life and consisted of additional children with stage 1, 2, and 3 disease (*P* < .001) but not of stage 4 disease (Supplementary Table 2, available online). The observed proportion of localized neuroblastoma in the second year of life was 412.8% higher than expected. The increase was highest for stage 1 disease (665.2%). Stage 2 disease showed an increase of 473.1%, and stage 3 disease showed 179.2% (all *P* < .001). In contrast, there was no difference for stage 4 (−3.1%, *P* = .86). Supplementary Table 4 (available online) demonstrates the absolute numbers of patients (per birth year and per stage).

The cumulative incidences of neuroblastoma cases in patients aged 12 to 71 months did not differ relevantly between the control cohort (1994-1999) and the prestudy (1990-1993) or the poststudy birth cohorts (2000-2004, 2005-2009, 2010-2014) (Table 1). Thus, exempting the study area cohort during the screening, no cumulative incidence change has been observed.

Table 1. Cumulative incidence in the screening and control areas according to neuroblastoma stage during the study cohort and in the prestudy and poststudy birth cohorts[a]

	Cumulative incidence, No. of cases per 100 000 births (95% CI)					
	Prestudy birth cohort	Study birth cohort		Poststudy birth cohort		
Stage	1990-1993	Screening area 1994-1999	Control area 1994-1999	2000-2004	2005-2009	2010-2014[b]
Total	7.6 (6.8 to 8.4)	13.4[c] (12.2 to 14.6)	9.3 (8.2 to 10.3)	8.1 (7.4 to 8.9)	7.7 (6.9 to 8.5)	7.1 (6.3 to 7.8)
1	0.9 (0.6 to 1.1)	3.6[c] (3.0 to 4.2)	1.5 (1.0 to 1.9)	0.9 (0.7 to 1.2)	1.1 (0.6 to 1.1)	0.7 (0.5 to 0.9)
2	0.7 (0.5 to 1.0)	2.3[c] (1.8 to 2.8)	0.9 (0.5 to 1.2)	0.9 (0.6 to 1.1)	1.2 (0.9 to 1.5)	1.1 (0.8 to 1.4)
3	1.4 (1.1 to 1.7)	2.3[c] (1.8 to 2.8)	1.8 (1.3 to 2.3)	1.4 (1.0 to 1.7)	1.2 (0.9 to 1.5)	1.1 (0.8 to 1.4)
1-3	3.0 (2.5 to 3.5)	8.3[c] (7.3 to 9.2)	4.2 (3.5 to 4.9)	3.2 (2.7 to 3.6)	3.2 (2.7 to 3.7)	2.9 (2.4 to 3.4)
4S	0.0 (0.0 to 0.0)	0.0 (0.0 to 0.0)	0.0 (0.0 to 0.0)	0.0 (0.0 to 0.0)	0.1 (0.0 to 0.2)	0.0 (0.0 to 0.1)
4	4.5 (3.8 to 5.0)	5.0 (4.3 to 5.7)	5.0 (4.2 to 5.8)	4.9 (4.3 to 5.5)	4.3 (3.7 to 4.9)	4.0 (3.4 to 4.5)
Stage missing	0.1 (0.0 to 0.2)	0.2 (0.0 to 0.3)	0.0 (0.0 to 0.1)	0.1 (0.0 to 0.2)	0.1 (0.0 to 0.2)	0.1 (0.0 to 0.2)

[a]Data include cases diagnosed at 12 to 71 months (second to sixth year) of age. CI = confidence interval.
[b]Observations until the sixth year of life in 2019 still incomplete.
[c]$P < .001$ comparing screening area and control area or screening area and prestudy or screening area and poststudy cohorts (Poisson model, χ^2 tests, 2-sided).

Table 2. Cumulative mortality within 10 years after diagnosis in the screening and control areas according to neuroblastoma stage during the study's birth cohort and in the prestudy and poststudy birth cohorts[a]

	Cumulative 10-y mortality, No. of deaths per 100 000 births (95% CI)				
	Prestudy birth cohort	Study birth cohort		Poststudy birth cohort	
Stage	1990-1993	Screening area 1994-1999	Control area 1994-1999	2000-2004[b]	2005-2009[b]
Total	3.4 (2.8 to 3.9)	3.5[c,d] (2.9 to 4.1)	3.8[c] (3.1 to 4.5)	3.2 (2.7 to 3.7)	2.7 (2.2 to 3.1)
1	0.0 (0.0 to 0.1)	0.2 (0.0 to 0.3)	0.0 (0.0 to 0.0)	0.0 (0.0 to 0.0)	0.0 (0.0 to 0.1)
2	0.0 (0.0 to 0.0)	0.1 (0.0 to 0.2)	0.0 (0.0 to 0.1)	0.1 (0.0 to 0.2)	0.1 (0.0 to 0.2)
3	0.4 (0.2 to 0.5)	0.3 (0.1 to 0.5)	0.4 (0.2 to 0.6)	0.4 (0.2 to 0.5)	0.1 (0.0 to 0.2)
1-3	0.4 (0.2 to 0.6)	0.6 (0.2 to 0.9)	0.6 (0.4 to 0.9)	0.4 (0.3 to 0.6)	0.3 (0.1 to 0.4)
4S	0.0 (0.0 to 0.0)	0.0 (0.0 to 0.0)	0.0 (0.0 to 0.0)	0.0 (0.0 to 0.0)	0.0 (0.0 to 0.1)
4	2.9 (2.4 to 3.3)	2.9 (2.3 to 3.4)	3.3 (2.6 to 3.9)	2.8 (2.3 to 3.2)	2.3 (1.9 to 2.8)
Stage missing	0.1 (0.0 to 0.2)	0.0 (0.0 to 0.0)	0.0 (0.0 to 0.1)	0.0 (0.0 to 0.0)	0.0 (0.0 to 0.1)

[a]Data include cases diagnosed at 12 to 71 months (second through sixth year) of age. CI = confidence interval.
[b]Follow-up for part of the cohorts still incomplete; thus, cumulative mortality is underestimated, especially for patients born between 2005 and 2009.
[c]$P = .53$ testing the hypothesis that the birth cohort (1994-1999, study screening and control area combined) is not different from all other cohorts and areas, adjusted for age groups (Poisson model, χ^2 tests, 2-sided).
[d]$P = .56$ testing the hypothesis that the study area (1994-1999) is not different from all other birth cohorts and areas, adjusted for age groups (Poisson model, χ^2 tests, 2-sided).

Mortality

The cumulative mortality within 10 years after diagnosis was not reduced for the screening cohort compared with the control cohort ($P = .78$). Similarly, the prestudy, study, and poststudy cohorts showed no mortality differences. This observation applied to the entire cohorts and for selected stages (Table 2). Supplementary Table 3 (available online) demonstrates that the 5-year mortalities did not differ when comparing the screening cohort with the control, prestudy, and poststudy cohorts.

Event-Free and Overall Survival

The event-free survival (EFS) and OS probabilities were lower in the control cohort compared with the screening cohort (10-year EFS control cohort, 51.1% [95% CI = 43.7% to 58.5%]; screening cohort, 66.5% [95% CI = 64.5% to 71.5%]; 10-year OS control cohort, 57.7% [95% CI = 50.5% to 64.9%]; screening cohort, 72.3% [95% CI = 67.5% to 77.1%], $P < .001$) for all patients with neuroblastoma aged 12 to 71 months. This lower probability was mainly caused by the higher proportion of lower-stage disease in the screening cohort (Supplementary Figure 1, available online [all stages], and Supplementary Figure 2, available online [by stage categories]).

The subgroup of patients aged 12 to 23 months with stage 4 disease at diagnosis ("the screening age")—that is, disease detected by screening (true-positive patients)—had a better long-term EFS and OS (EFS = 88.9% [95% CI = 75.5% to 100%]; OS = 94.4% [95% CI = 84.4% to 100%]) compared with unscreened patients aged 12 to 23 months with stage 4 disease in the control cohort (EFS = 29.0% [95% CI = 16.7% to 50.3%], $P = .001$; OS = 29.0% [95% CI = 16.7% to 50.3%], $P = .002$) and to patients aged 12 to 23 months with a false-negative screening result (EFS = 16.7% [95% CI = 2.8% to 99.7%]; OS = 16.7% [95% CI = 2.8% to 99.7%]; $P < .001$ each) (Figure 1). In contrast, no outcome advantage was seen for patients with stage 3 disease if detected by screening or for patients with stage 2 or 1 disease.
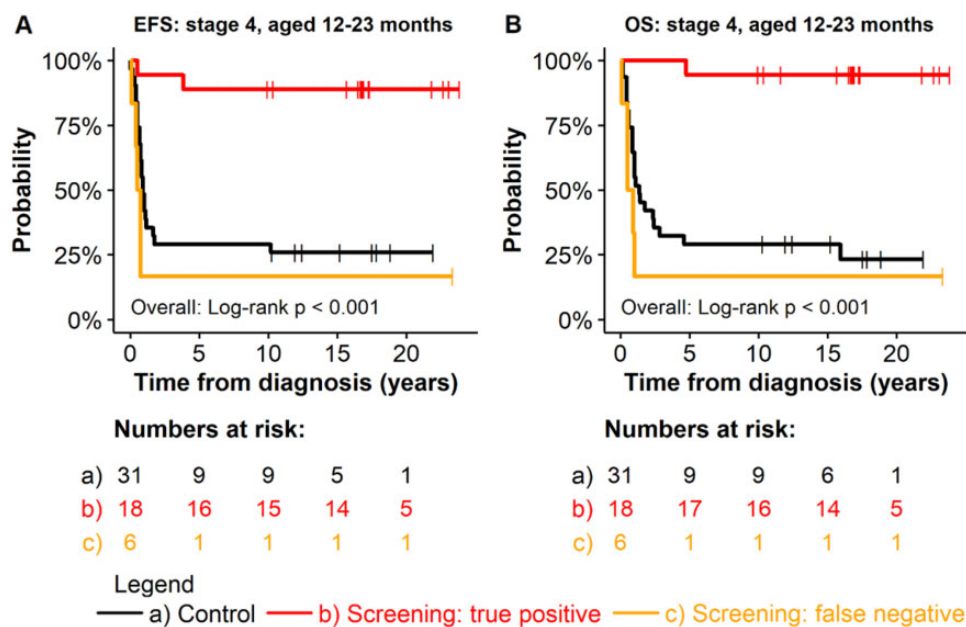
Figure 1. Outcome for patients with neuroblastoma aged 12 to 23 months at diagnosis, comparing patients of the screening cohort (true-positive and false-negative) with patients of the control cohort by stage. Screening nonparticipants in the screening cohort were excluded. All statistical tests were 2-sided. EFS = event-free survival; OS = overall survival.

No survival differences were observed in the subgroups of patients aged 24 and 71 months at diagnosis between screening false-negative patients and the control cohort (no true-positive cases beyond "the screening age") (Supplementary Figure 3, available online).

Association of Screening and Risk Factors

The stage 4 group of the screening cohort (both participants and nonparticipants), with patients aged 12 to 71 months, had 18.0% asymptomatic patients compared with 2.9% in the control cohort (Supplementary Table 5, available online). No differences were detected regarding median age ($P = .86$); MYCN proto-oncogene, bHLH transcription factor (*MYCN*) oncogene amplification ($P = .89$); or tumor load at diagnosis (estimated by the median number of involved metastatic organs; $P = .17$).

The subgroup of 18 patients aged 12 to 23 months with stage 4 disease detected by screening was younger (median = 14.3 months [95% CI = 13.4 to 16.3]) compared with screening-negative patients aged 12 to 23 months (median = 18.4 months [95% CI = 16.6 to 20.0]; $P < .001$) and control cohort patients (median = 17.6 months [95% CI = 15.3 to 19.3]; $P = .01$).

Symptomatic at diagnosis in stage 4 were 1 patient from the screening-detected group, all the screening false-negatives, and most (94.4%) patients in the control cohort ($P < .001$). Only 1 of the 18 stage 4 cases identified through screening showed an *MYCN* amplification (1 of 17 [5.9%]) compared with 3 of 5 (60.0%) patients in the screening-negative group and 17 of 25 (68.0%) patients in the control cohort ($P < .001$). Furthermore, the tumor load of the stage 4 screening-detected cases was lower. Those patients had a median of 1 organ with metastases (interquartile range [IQR] = 1-2), while those in the false-negative group had 3 such organs (IQR = 2-4) and the control cohort had 2 (IQR = 1-3) ($P = .03$).

Death From Toxicity

In addition to the burden of surgical treatment, chemotherapy, radiation therapy, and other therapies, death caused by treatment is a rare but certainly particularly harmful toll of a screening project. Of 352 patients aged 12 to 71 months with stage 1 to 3 neuroblastoma in the screening area, 4 died of toxicity (n = 1 at stage 1, n = 2 at stage 2, n = 1 at stage 3). Of those 4 patients, only 2 had cancer detected by screening. The other 2 lived in the screening area but did not participate in the screening program. In the corresponding control area, no patients died of toxicity (n = 84) (Supplementary Table 6, available online). In the pre-study and poststudy cohorts, the numbers of patients with stage 1, 2, and 3 disease who died of toxicity were comparably low (1990-1993: 5 of 101 patients; 2000-2004: 2 of 119 patients; 2004-2009: 1 of 116 patients; 2010-2014: 1 of 110). Thus, death clearly caused by toxicity did not notably increase during the screening study. Of the 219 patients with stage 4 disease in the screening area, 80 died. Seven deaths were caused by toxicity. None of the 7 malignancies were detected by screening.

Altogether, the majority of deaths were caused by tumor progression or tumor progression and/or toxicity (Supplementary Table 6, available online). This was true for all stages.

Excess and Asymptomatic Cases

Table 3 demonstrates that a statistically significant number of excess cases (overdiagnosis) was seen in the participant group. This trend did not persist after the end of the project. The percentage of asymptomatic cases is traditionally high in Germany because of recommended routine health care check-ups in infants and toddlers. Supplementary Table 6 (available online) demonstrates an increase of the percentage of cases diagnosed without symptoms for the screening cohorts and a return to the

Table 3. Estimated excess cases (overdiagnosis) in the screening cohort for children who underwent actual screening (screening participants) compared with the prestudy birth cohort and the birth cohort of the control area[a]

| | Cumulative incidence difference of participants only in the study area birth cohort (1994-1999) compared with: | |
Variable	Prestudy birth cohort (1990-1993), excess No. of cases[b] per 100 000 births	Study birth cohort control area (1994-1999), excess No. of cases[b] per 100 000 births
Cumulative	8.1	6.4
Stage 1	4.0	3.4
Stage 2	2.7	2.5
Stage 3	1.3	1.1
Stage 4	−0.1	−0.5

[a]Data include cases diagnosed at 12 to 71 months (second through sixth year) of age.

[b]Including the postscreening ages until the catch-up point at the sixth birthday. The excess is corrected for a lead-time effect and thus an estimate for the extent of overdiagnosis.

prescreening level during the first poststudy period (2000-2004). During the second poststudy period (2005-2009), again, more asymptomatic cases were detected but not in the following years (2010-2014).

Discussion

This study confirms that a screening program for neuroblastoma for children at 1 year of age does not reduce the incidence of metastatic (stage 4) neuroblastoma and has no influence on neuroblastoma mortality. It also confirms that not introducing neuroblastoma screening into routine examinations in 2001 was the correct decision (last birth cohort evaluated was 1999).

To our knowledge, this is the only long-term epidemiologic analysis on this topic. A primary strength of our study is the comparison of assigned areas (screening and control cohorts), which were demonstrated to be comparable with respect to neuroblastoma incidence, age distribution, stage distribution, and mortality before the screening study. Other strengths of our study are the comparison of the screening cohort with the control cohort, the long observation period (>20 years after the cessation of the screening study), and the detailed corresponding clinical data obtained within the framework of the national clinical trials.

Limitations include the relative low compliance rate (only 61.2%) with participation in the screening program within the screening area (8). Similarly, the results are limited to the screening age—at approximately the first birthday leading—to screening-detected cases almost exclusively within the second year of life. Although earlier screening during infancy as performed in Canada (7) and Japan (4,9,17) are in agreement with our results and support the concept of a more benign neuroblastoma type early in life and a more aggressive type later (7,8,11), it cannot be ruled out that screening later in life (18) may result in less overdiagnosis and a reduction in neuroblastoma mortality. Biological or epidemiologic data suggesting this outcome, however, have not been presented so far.

Some general aspects should be highlighted. First, 80 patients had been observed with a negative screening result but presented later with disease. Of those, 20 (25.0%) showed no elevated catecholamine excretion at clinical diagnosis, which is similar to the 22.9% of patients in the concurrent control cohort in stage-related comparison. The metabolic activity of the tumor requires some level to exceed the age-related reference limits. Furthermore, the distribution of the interval between detectable disease (by catecholamine screening or by imaging techniques) and clinical presentation—called the *sojourn time*—

is unknown. Earlier, we estimated the mean lead time to be about 15 months (16). The empirical lead time distribution was right skewed (because of the age distribution of patients with neuroblastoma). We assume that the sojourn time distribution may similarly be skewed. Based on this hypotheses, the mean sojourn time could even be shorter than the mean lead time (16).

The relatively high number of false-negatives is particularly regrettable for high-risk disease (*MYCN*-amplified tumors and stage 4 disease). The known fast growth of *MYCN*-amplified tumors leaves only a short window for early detection within a screening program. The actual number of cases detected earlier than without screening (ie, those who might in principle have benefitted from the screening) was far lower (n = 18) than hoped for. Thus, the tumor biology leads to a low sensitivity, which is a relevant issue for a neuroblastoma screening program.

An unfavorable effect of any screening program is overdiagnosis, which was particularly pronounced here. The excess cases were exclusively seen in the group with stage 1, 2 and 3 disease but not in stage 4 disease. Approximately 7 cases per 100 000 children were additionally diagnosed with neuroblastoma in the screening cohort compared with the control cohort and with the pre- and poststudy incidences. This overdiagnosed group is interpreted as consisting of patients whose tumor would have regressed spontaneously before clinical diagnosis, a well-known characteristic of neuroblastoma stage 4S disease (10) and in infants with stage 1 through 3 disease (11,12). The lack of decrease in the incidence of unfavorable advanced-stage disease in older children aged more than 18 months is in agreement with the Quebec study (19). Moreover, the authors demonstrated for screening at 3 weeks and 6 months an increase in the standardized incidence ratio to 2.85 at birth to 11 months of age, to 1.58 at 12 to 23 months of age at diagnosis, and normalization in the age groups from 24 to 71 months of age (19). Together with our overdiagnosis data (Supplementary Table 1, available online), both reports suggest that most overdiagnosis occurs in the first 2 years of life. The overdiagnosed patients carry the burden of unnecessary treatment. Luckily, our data did not indicate an increase in the toxic death rate as the worst case of toxicity.

Patients with stage 4 disease detected by screening had surprisingly high EFS and OS probabilities. Diagnosis at an earlier age, being asymptomatic at detection, having less tumor burden, and in only 1 case having *MYCN* amplification indicated a more favorable biology. For these 18 patients with stage 4 disease who may have benefitted from screening, 1.5 million children were screened. It remains open whether the better

prognosis resulted from more favorable biological characteristics or from earlier detection by screening (ie, before the acquisition of poorer features if detected later clinically). In view of the nonincreased stage 4 disease incidence rate in the screening cohort (ie, no overdiagnosis), the saved lead time could contribute to a more favorable biology. Thus, the detection of this subgroup is biologically interesting but does not change the recommendation to close the book on catecholamine-based neuroblastoma screening. Overdiagnosis and potential harm by unnecessary treatment add to that decision. Generally, it is assumed that disease detected by a screening program has benign biological properties (6). Arakawa and colleagues reported that they did not observe any recurrence of disappeared tumors or any progression of persisting neuroblastomas for up to 15 years (9).

In Canada, the incidence of neuroblastoma remained increased in the screening area of Quebec 1 to 3 years after the end of the screening program compared with the control areas (Minnesota and Ontario) but dropped to the previous levels in years 4 and 5 (15). In Germany, the stage distribution and cumulative overall incidence after the first birthday returned to the prescreening and control-area levels.

In Japan, the neuroblastoma mortality rate for children was not different before and after cessation of the screening program in 2003 (14). The Canadian study, with screening between 3 weeks and 6 months of age, drew the same conclusion from the data (7). To our knowledge, only 1 report published a decline in mortality. The relative risk decreased to 0.73 (95% CI = 0.58 to 0.90) using a qualitative catecholamine approach and to 0.53 (95% CI = 0.42 to 0.63) using a quantitative catecholamine determination approach (19). This finding, however, has been challenged for methodologic reasons (20). Our population-based comparison does not indicate a drop in neuroblastoma mortality associated with the screening program.

This long-term evaluation of the German study on neuroblastoma screening at one 1 of age confirms the unintended side effect of overdiagnosis and a failure to reduce population-related neuroblastoma mortality. The high number of false-positive screening results leading to unnecessary medical examinations as well as the number of false-negative results potentially leading to a false perception of safety represent additional burdens. A small subgroup of screening-detected stage 4 neuroblastoma cases had a surprisingly good outcome. It remains open as to whether this outcome was the result of early detection or to more favorable tumor biology independent of time.

Funding

Notes

Data Availability

For this report data are given in the article and the Supplementary Materials (available online). Additional data are available in the original report published at doi:10.1056/NEJMoa012277.

References

1. Brodeur GM, Pritchard J, Berthold F, et al. Revisions of the international criteria for neuroblastoma diagnosis, staging, and response to treatment. *J Clin Oncol.* 1993;11(8):1466–1477. doi:10.1200/J Clin Oncol.1993.11.8.1466.
2. Berthold F, Hunneman DH, Käser H, et al. Neuroblastoma screening: arguments from retrospective analysis of three German neuroblastoma trials. *Am J Pediatr Hematol Oncol.* 1991;13(1):8–13.
3. Pinto NR, Applebaum MA, Volchenboum SL, et al. Advances in risk classification and treatment strategies for neuroblastoma. *J Clin Oncol.* 2015;33(27):3008–3017. doi:10.1200/JCO.2014.59.4648.
4. Sawada T, Hirayama M, Nakata T, et al. Mass screening for neuroblastoma in infants in Japan. Interim report of a mass screening study group. *Lancet.* 1984;324(8397):271–273. doi:10.1016/s0140-6736(84)90311-8.
5. Woods WG. Screening for neuroblastoma. In: NKV Cheung, S Cohn, eds. *Neuroblastoma.* Berlin, Germany: Springer; 2005:7–19.
6. PDQ Pediatric Treatment Editorial Board. Neuroblastoma treatment (PDQ®): health professional version. In: PDQ Cancer Information Summaries. Bethesda, MD: US National Cancer Institute; 2020.
7. Woods WG, Gao RN, Shuster JJ, et al. Screening of infants and mortality due to neuroblastoma. *N Engl J Med.* 2002;346(14):1041–1046. doi:10.1056/NEJMoa012387.
8. Schilling FH, Spix C, Berthold F, et al. Neuroblastoma screening at one year of age. *N Engl J Med.* 2002;346(14):1047–1053. doi:10.1056/NEJMoa012277.
9. Arakawa A, Oguma E, Aihara T, et al. Long-term follow-up results of the observation program for neuroblastoma detected at 6-month mass screening. *J Pediatr.* 2014;165(4):855–857.e1. doi:10.1016/j.jpeds.2014.06.055.
10. D'Angio GJ, Evans AE, Koop CE. Special pattern of widespread neuroblastoma with a favourable prognosis. *Lancet.* 1971;1(7708):1046–1049. doi:10.1016/s0140-6736(71)91606-0.
11. Yamamoto K, Hanada R, Kikuchi A, et al. Spontaneous regression of localized neuroblastoma detected by mass screening. *J Clin Oncol.* 1998;16(4):1265–1269. doi:10.1200/JCO.1998.16.4.1265.
12. Hero B, Simon T, Spitz R, et al. Localized infant neuroblastomas often show spontaneous regression: results of the prospective trials NB95-S and NB97. *J Clin Oncol.* 2008;26(9):1504–1510. doi:10.1200/JCO.2007.12.3349.
13. dos Santos Silva I. Cancer prevention. In: *Cancer Epidemiology: Principles and Methods.* Lyon, France: International Agency for the Research of Cancer; 1999:355–383.
14. Shinagawa T, Kitamura T, Katanoda K, Matsuda T, Ito Y, Sobue T. The incidence and mortality rates of neuroblastoma cases before and after the cessation of the mass screening program in Japan: a descriptive study. *Int J Cancer.* 2017;140(3):618–625. doi:10.1002/ijc.30482.
15. Barrette S, Bernstein ML, Robison LL, et al. Incidence of neuroblastoma after a screening program. *J Clin Oncol.* 2007;25(31):4929–4932. doi:10.1200/JCO.2007.12.1905.
16. Spix C, Michaelis J, Berthold F, Erttmann R, Sander J, Schilling FH. Lead-time and overdiagnosis estimation in neuroblastoma screening. *Stat Med.* 2003;22(18):2877–2892. doi:10.1002/sim.1533.
17. Sawada T, Kidowaki T, Sakamoto I, et al. Neuroblastoma. Mass screening for early detection and its prognosis. *Cancer.* 1984;53(12):2731–2735. doi:10.1002/1097-0142(19840615)53:12.
18. Hiyama E, Iehara T, Sugimoto T, et al. Effectiveness of screening for neuroblastoma at 6 months of age: A retrospective population-based cohort study. *Lancet.* 2008;371(9619):1173–1180. doi:10.1016/S0140-6736(08)60523-1.
19. Woods WG, Tuchman M, Robison LL, et al. A population based study of the usefulness of screening for neuroblastoma. *Lancet.* 1996;348(9043):1682–1687. doi:10.1016/S0140-6736(96)06020-5.
20. Maris JM, Woods WG. Screening for neuroblastoma: a resurrected idea? *Lancet.* 2008;371(9619):1142–1143. doi:10.1016/S0140-6736(08)60500-0.