ORIGINAL ARTICLE

The Care of Preterm Infants With Birth Weight Below 1250 g

Risk-Adjusted Quality Benchmarking as Part of Validating a Caseload-Based Management System

Marcus Kutschmann, Sven Bungard, Joachim Kötting, Andrea Trümner, Christoph Fusch, Christof Veit

SUMMARY

Background: In Germany, controversy currently surrounds the contention that the quality of care for preterm infants weighing less than 1250 g is best assured by requiring that centers treat a minimum of 30 such cases per year.

Methods: A risk-adjusted model was developed on the basis of neonatal data from 7405 preterm infants treated in German centers, and the effect of caseload on risk-adjusted mortality was analyzed. In addition, the discriminative ability of the minimal caseload requirement for quality assessment was studied. The authors designate the quality of care in a particular center as above average if the observed mortality is lower than would have been expected from the risk profile of the preterm infants treated there.

Results: Risk-adjusted mortality was found to be significantly higher in smaller centers (those with fewer than 30 cases per year) than in larger ones (odds ratio, 1.34). Even among centers whose caseload exceeded the minimum requirement, there was still marked variability in risk-adjusted mortality (range: 3.5% to 28.6%). Of all the preterm infants treated in larger centers, 56% were treated in centers with above-average quality of care. 44% of the centers with above-average quality of care had caseloads in the range of 14 to 29 cases per year.

Conclusion: Because of the marked variability in risk-adjusted mortality, even among larger centers, a caseload of 30 or more cases per year is not a suitable indicator of the quality of care. The neonatal data of external quality assurance should be used to develop an instrument for quality-based coordination of care that takes not just morbidity and mortality, but also the treating centers' competence profiles into account.

► Cite this as:

Kutschmann M, Bungard S, Kötting J, Trümner A, Fusch C, Veit C: The care of preterm infants with birth weight below 1250 g: risk-adjusted quality benchmarking as part of validating a caseload based management system. *Dtsch Arztebl Int* 2012; 109(31–32): 519–26. DOI: 10.3238/arztebl.2012.0519

More than 9000 preterm infants weighing less than 1500 g are born every year in Germany (1). In order to guarantee high-value care for these high-risk patients, in 2009 Germany's Federal Joint Committee (G-BA, Gemeinsamer Bundesausschuss) set a minimum caseload (2) for this area, which was raised to 30 per year for infants weighing less than 1250 g in 2010. (3). However, the minimum is currently suspended (4). The G-BA bases the required caseload on a report by the Institute for Quality and Efficiency in Health Care (IQWiG, Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen) (5). The IQWiG concludes that for the care of neonates with birth weight below 1500 g mortality falls as a hospital's caseload increases, but the lower mortality rate is not necessarily caused by the higher volume of care. The idea arose of examining the basis for argumentation and the potential effects of the minimum caseload using neonatal data that have been gathered all over Germany for several years. As in other, comparable studies, mortality was selected as a surrogate outcome parameter for other aspects of quality of care. Using neonatal data, the effect of the caseload on risk-adjusted mortality was analyzed on the basis of a differentiated risk-adjustment, and subsequently the discriminative ability of a minimum annual number of 30 cases was investigated.

Methods

Analysis included data from preterm infants with birth weight below 1250 g born between 2007 and 2009, from the German federal states of Bavaria, Baden-Württemberg, Hesse, Lower Saxony, and North Rhine-Westphalia. These states account for approximately 68% of all births in Germany (6).

A systematic review of the literature identified risk factors that affect the mortality rate of preterm infants and are not affected by care. A risk-adjustment model with mortality as outcome parameter was then developed.

Calculation of mortality rates did not include preterm infants admitted to neonatal units more than 24 hours after birth, because they had probably received neonatal treatment before admission. Infants referred to other units were also excluded, because

BQS Institut für Qualität & Patientensicherheit, Düsseldorf: Dr. rer. medic. Kutschmann, Dr. med. Bungard, Kötting, Trümner, Dr. med. Veit

McMaster University, Hamilton (Canada): Prof. Dr. med. Fusch

TABLE 1

Univariate descriptive analysis of factors included in the risk-adjustment model (p-values of chi-square test)

Risk factor	Risk category	No. of preterm infants	Preterm infants who do not survive		p-value
			No.	(%)	
Multiple birth	Yes	2375	368	(15.5)	0.048
	No	5884	811	(13.8)	
Sex	Male	4124	699	(16.9)	<0.001
	Female	4135	480	(11.6)	
Birth weight (g)	≤499	484	275	(56.8)	<0.001
	500 to 749	1858	516	(27.8)	
	750 to 999	2856	277	(9.7)	
	1000 to 1249	3061	111	(3.6)	
Week of gestation	≤22	123	101	(82.1)	<0.001
	23 to 24	1124	481	(42.8)	
	25 to 26	1952	347	(17.8)	
	>26	5060	250	(4.9)	
Body length (cm)	≤35	4340	1007	(23.2)	<0.001
	>35	3915	169	(4.3)	
Malformations	Severe	110	78	(70.9)	<0.001
	Mild	552	82	(14.9)	
	None	7597	1019	(13.4)	
Greatest base deficit	≤−15	442	230	(52.0)	<0.001
	−14 to −10	842	223	(26.5)	
	≥−9	6325	602	(9.5)	
Minimum FiO ₂	0.91 to 1.00	184	111	(60.3)	<0.001
	0.61 to 0.90	100	50	(50.0)	
	0.41 to 0.60	201	84	(41.8)	
	≤0.40	7227	819	(11.3)	
Maximum FiO ₂	0.91 to 1.00	1265	588	(46.5)	<0.001
	0.81 to 0.90	128	28	(21.9)	
	0.41 to 0.80	1845	192	(10.4)	
	≤0.40	4479	261	(5.8)	

mortality rates cannot be calculated for them. Quality was therefore assessed on the basis of data from 8259 preterm infants who received all their treatment in a single unit. Risk factors were not fully recorded for 854 (10.3%) of these 8259 infants. As a result, the risk-adjustment model was based on 7405 (89.7%) preterm infants.

The risk-adjustment model was used to calculate each hospital's expected mortality, E, according to the risk profile of the treated preterm infants. E was then compared to observed mortality, O. A quotient $O \div E$ of less than 1 (or greater than 1) should be considered evidence of above-average (or below-average) quality of care.

In order to assess the effect of the caseload, the factor "caseload" was additionally included in the model, using the categories "fewer than 30 cases" and "30 cases or more." In a second version, caseload was considered as factor with 12 intervals of five up to 59 and the category "60 or more." The caseloads-recorded in 2009 were used as the basis for these categories. Also investigated was the question of whether the minimum caseload could distinguish sufficiently well between above-average and below-average quality of care.

The methods used are described in detail in the supplementary *eMethods* section.

TABLE 2

Risk-adjustment model: results of multiple logistic regression with mortality as outcome parameter

Risk factor	Category	Reg. coeff.	Standard error	p-value	Odds ratio	95% CI
Multiple birth	Yes	0.4537	0.0934	<0.001	1.5741	1.3108–1.8903
	No	Reference category				
Sex	Male	0.4554	0.0866	<0.001	1.5768	1.3305–1.8686
	Female	Reference category				
Birth weight	1 g	–0.0028	0.0004	<0.001	0.9972	0.9964–0.9980
Week of gestation	1 week	–0.1965	0.0271	<0.001	0.8216	0.7791–0.8664
Body length	1 cm	–0.0610	0.0227	0.007	0.9408	0.8999–0.9836
Malformations	Severe	2.9461	0.3149	<0.001	19.0318	10.2663–35.2816
	None/mild	Reference category				
Greatest base deficit	≤–15	1.2948	0.1378	<0.001	3.6504	2.7866–4.7819
	–14 to –10	0.5591	0.1119	<0.001	1.7491	1.4046–2.1781
	≥–9	Reference category				
Minimum FiO ₂	per 10% oxygen saturation	0.2309	0.0238	<0.001	1.2597	1.2023–1.3199
Maximum FiO ₂	per 10% oxygen saturation	0.2139	0.0152	<0.001	1.2385	1.2021–1.2760
Constant		5.0562				

Reg. coeff.: regression coefficient; 95% CI: 95% confidence interval

Results

Of the 8259 preterm infants born between 2007 and 2009, 1179 (14.3%) died. In 2009, 39 hospitals met the criterion of at least 30 cases. In these hospitals, 14.1% of preterm infants (613 of 4341) died. The mortality rate in the 127 units with fewer than 30 cases was 14.4% (566 of 3918).

At 18.2% the proportion of preterm infants referred to or from a neonatal unit in larger hospitals is somewhat higher than in hospitals with smaller caseloads (16.5%). The expected mortality rate for these excluded cases (1738 infants in all) is 14.5% for large centers and 16.7% for smaller ones.

The risk-adjustment model

The search of the literature identified the parameters multiple birth, sex, birth weight, week of gestation, head circumference on admission, body length on admission, malformations, base deficit, and minimum and maximum FiO₂ administered as potentially relevant risk factors. The parameter “head circumference” was excluded from the multiple logistic regression model, because it was not significant at a level of 0.05. *Table 1* shows the results of a univariate descriptive analysis of the factors included in the risk-adjustment model.

Data on risk factors were complete for 89.7% of the preterm infants (7405 of 8259). The incomplete datasets are distributed equally between large and small hospitals (10.3% and 10.4%). The mortality rate for these cases is 18.9%, higher than for infants with complete datasets (13.7%). It is 23.9% in the 39 large units, and 13.3% in the 127 small units.

The results yielded by the multiple logistic risk-adjustment model are shown in *Table 2*. The risk factors included show a highly significant effect on mortality. Nagelkerke’s R-square value is 0.46. The Hosmer–Lemeshow test indicates a good model fit, with a p-value of 0.65.

The effect of volume of care

If the caseload is divided into two categories (preterm infants from hospitals with fewer than 30 cases in 2009 versus preterm infants from hospitals with 30 cases or more), the adjusted risk of death for infants from small units is higher overall than for infants from large units (odds ratio 1.34, $p = 0.001$). If the caseload is divided into 13 categories (*Table 3*), there is a significantly increased risk-adjusted mortality in the categories “5 to 9,” “15 to 19,” “20 to 24,” and “35 to 39” than in the reference category. In contrast, for the four categories “30 to 34,” “40 to 44,” “50 to 54,” and “55 to 59” the risk-adjusted mortality is lower than in the reference category, although not significant. No monotonic increase in the effect of the caseload across all caseload categories is observed. Taking caseload into account in the risk-adjustment model, there are only minor changes in the results of the other risk factors, as expected.

Table 4 shows the values of O÷E and the risk-adjusted mortality rate for each interval. Risk-adjusted mortality is calculated as the product of O÷E and overall mortality ($n = 7405$, 13.7%). It varies substantially within individual caseload categories (*Figures 1 and 2*). For example, in units with more than 60 cases it ranges from 5.9% to 17.9%.

TABLE 3

Risk-adjustment model with caseload as categorical factor

Risk factor	Category/unit	Reg. coeff.	Standard error	p-value	Odds ratio	95% CI
Multiple birth	Yes	0.4832	0.0943	<0.001	1.6212	1.3475–1.9505
	No	Reference category				
Sex	Male	0.4605	0.0872	<0.001	1.5849	1.3359–1.8805
	Female	Reference category				
Birth weight	1 g	–0.0028	0.0004	<0.001	0.9972	0.9964–0.9980
Week of gestation	1 week	–0.2015	0.0273	<0.001	0.8175	0.7750–0.8624
Body length	1 cm	–0.0669	0.0228	0.003	0.9353	0.8945–0.9780
Malformations	Severe	2.9457	0.3154	<0.001	19.0242	10.2525–35.3007
	None/mild	Reference category				
Greatest base deficit	≤–15	1.3008	0.1388	<0.001	3.6721	2.7977–4.8196
	–14 to –10	0.5630	0.1134	<0.001	1.7559	1.4060–2.1928
	≥–9	Reference category				
Minimum FiO ₂	per 10% oxygen saturation	0.2266	0.0240	<0.001	1.2543	1.1967–1.3147
Maximum FiO ₂	per 10% oxygen saturation	0.2134	0.0155	<0.001	1.2379	1.2008–1.2761
Caseload*	0–4	0.1981	0.2393	0.408	1.2191	0.7627–1.9487
	5–9	0.4853	0.2335	0.038	1.6246	1.0279–2.5677
	10–14	0.1835	0.2080	0.378	1.2014	0.7991–1.8060
	15–19	0.3361	0.1592	0.035	1.3995	1.0244–1.9120
	20–24	0.7105	0.1838	<0.001	2.0349	1.4194–2.9173
	25–29	0.1162	0.1631	0.476	1.1233	0.8159–1.5464
	30–34	–0.1412	0.2387	0.554	0.8683	0.5438–1.3864
	35–39	0.3835	0.1604	0.017	1.4674	1.0715–2.0097
	40–44	–0.0204	0.1827	0.911	0.9798	0.6849–1.4019
	45–49	0.0157	0.3216	0.961	1.0158	0.5408–1.9079
	50–54	–0.3407	0.2357	0.148	0.7113	0.4482–1.1288
	55–59	–0.2713	0.2307	0.240	0.7624	0.4851–1.1982
	≥60	Reference category				
Constant		5.2278				

*Significant values shown in bold; Reg. coeff.: regression coefficient; 95% CI: 95% confidence interval

Table 4 also shows the distribution of preterm infants and hospitals among the 13 caseload categories. Only just under one-quarter of hospitals (n = 39; 24%) achieved the minimum caseload of 30 cases in 2009. 46.8% (n = 4671) of preterm infants were cared for in hospitals with small caseloads, and 53.2% (5306) in large units.

Discriminative ability of the minimum caseload

Assessment of the discriminative ability of the minimum caseload criterion included the 91 hospitals in which at least 14 preterm infants had been cared for, in line with the minimum caseload requirement for 2009 in Germany (Table 5). 52 hospitals were deemed to provide insufficient quality of care according to the current

minimum caseload criterion, because they cared for fewer than 30 cases. However, at the same time, 17 of these hospitals are among the 39 hospitals with above-average quality of care ($O+E<1$). This means that there is a false negative rate of 44% at hospital level. In contrast, only 2975 of the 5306 preterm infants cared for in hospitals with supposedly good quality of care (30 cases or more) were actually treated in hospitals with above-average quality of care. In other words, the positive predictive value of the minimum caseload criterion is 56% at patient level.

Discussion

A minimum caseload criterion is intended to substantially improve the quality of care provided to very small

TABLE 4

Risk-adjusted mortality of infants with birth weight <1250 g by caseload category*

Caseload category	Hospitals		Preterm infants		Risk-adjusted mortality	Range (adjusted mortality)	O+E	Range (O+E)
	No.	(%)	No.	(%)				
≥60	8	(5)	1582	(15.9)	12.5%	5.9–17.9%	0.91	0.43–1.30
55–59	3	(2)	562	(5.6)	10.7%	7.1–18.9%	0.78	0.52–1.37
50–54	3	(2)	501	(5.0)	10.4%	6.1–19.2%	0.76	0.45–1.39
45–49	2	(1)	228	(2.3)	12.6%	11.4–13.5%	0.92	0.83–0.98
40–44	7	(4)	904	(9.1)	12.5%	3.5–27.0%	0.91	0.25–1.97
35–39	10	(6)	1037	(10.4)	15.4%	8.2–28.6%	1.12	0.60–2.08
30–34	6	(4)	492	(4.9)	11.6%	5.2–24.7%	0.84	0.38–1.80
25–29	13	(8)	1125	(11.3)	13.4%	6.2–23.7%	0.98	0.45–1.72
20–24	11	(7)	667	(6.7)	18.5%	11.7–31.6%	1.34	0.85–2.30
15–19	23	(14)	1210	(12.1)	15.0%	8.2–31.4%	1.09	0.59–2.29
10–14	17	(10)	669	(6.7)	13.8%	0.0–36.1%	1.01	0.00–2.63
5–9	22	(13)	548	(5.5)	16.1%	0.0–67.2%	1.17	0.00–4.89
0–4	38	(23)	452	(4.5)	14.0%	0.0–183.9%	1.02	0.00–13.38
Total	163	(100)	9977	(100.0)	13.7%	0.0–183.9%	1.00	0.00–13.38

*For 10 hospitals that cared for a total of 20 preterm infants, O+E could not be calculated, because either only referred infants or infants with incomplete records were treated (O = observed mortality; E = expected mortality)

preterm infants by concentrating care in centers with higher caseloads. To assess this, neonatal data were used to examine the potential effect of a minimum annual caseload of 30 cases.

Analysis is limited to risk-adjusted mortality as a central indicator of quality of care, because it is an outcome parameter on which information can be reliably gathered. Simultaneously examining complication morbidity was beyond the scope of this article. However, it should be examined in future quality assessments.

Because records are insufficiently validated, it is often insinuated that neonatal data are incomplete. The higher mortality rate in official statistics is seen as evidence of this. Analyses by Hummer et al. (7) indicate that approximately one-third of infants who die do not show up in neonatal units' figures. A decisive consideration in this regard might be that before 2010 records were only kept on such infants if they were admitted to a pediatric unit on or before their tenth day of life. No records were kept on infants who died before the beginning of treatment or while receiving palliative care. Various calculations show that these infants account for a large part of the gap in records. Perinatal data gathered in 2008 (8) show that half of postpartum deaths occurred within six hours (56%, 539 of 954 cases). A Swedish working group estimated the proportion of infants whose treatment was discontinued at 40% of all those who died before they were one day old (9). From 2010 onwards, new procedures for gathering neonatal data have closed this gap in perinatal mortality

data in Germany. However, infants who receive palliative care play no role in the comparison of the quality of neonatal care units. Hospitals which document the data of palliatively treated infants, compared with facilities that do not, have no disadvantage because these cases are considered in the risk-adjusted mortality. This means that the available neonatal data do provide a sufficiently valid basis to address the question posed.

In addition to the unclear status of data, particularly regarding infants given palliative care, a further limitation of the study is that only data from the five German federal states named above were available. In addition, in 11.3% of the preterm infants whose data were available the factors used in the risk-adjustment model were incompletely recorded, which meant that they could not be included in risk-adjusted analysis. However, the analyzed sample is large enough to obtain meaningful results. Also, the fact that infants referred to other units were excluded may mean that the ability of particularly well-qualified units to treat high-risk infants referred from other hospitals is not taken into account. However, differences in quality should be revealed by comparing infants treated only in the unit in question.

The neonatal parameters used in the risk-adjustment model primarily form the mortality-related risk profiles of preterm infants. They are largely independent of quality of care and correspond to the factors described in the literature (10–14). The Hosmer–Lemeshow test indicates good model fit, with a p-value of 0.65.

FIGURE 1

Risk-adjusted mortality of preterm infants with birth weight below 1250 g in 2007 to 2009, plotted against volume of care in 2009

Values of 0 for volume of care in 2009 with associated values above 0% for risk-adjusted mortality refer to hospitals that still treated preterm infants in 2007 and 2008 but no longer did so in 2009. Three hospitals with risk-adjusted mortality figures of 62% and 67% (5 cases each, 2009) and 184% (no cases in 2009) are not shown.

FIGURE 2

Risk-adjusted mortality in 2007 to 2009 for individual caseload categories in 2009 with ranges of hospital results

The x-axis shows caseload categories and the numbers of hospitals and preterm infants weighing less than 1250 g in these categories. For the caseload categories "0 to 4" and "5 to 9," the maximum values of risk-adjusted mortality (67% and 184%) are not shown.

TABLE 5

Discriminative ability of the minimum caseload criterion at hospital level (H) and patient level (P), in hospitals with at least 14 cases

Quality of care	Above average (O÷E <1)		Below average (O÷E >1)		Total	
	H	P	H	P	H	P
30 cases or more	22	2975	17	2331	39	5306
	True positive		False positive			
Fewer than 30 cases	17	1072	35	2173	52	3245
	False negative		True negative			
Total	39	4047	52	4504	91	8551

O = observed mortality; E = expected mortality

Minimum and maximum FiO₂ administered are part of medical care, but like base deficit are recorded on admission. They are therefore rarely affected by care and so largely reflect preterm infants' health.

If caseload is divided into only two categories in the risk-adjustment model, risk-adjusted mortality in the group of units with annual caseloads of 30 or more is significantly lower than in the group of smaller units. Bartels et al. (15) had already reported a significantly higher adjusted mortality rate for infants whose birth weight was under 1500 g who were born in small hospitals (annual caseload fewer than 36). Trotter and Pohlandt (16) described a significant connection between caseload and mortality in infants weighing under 750 g or born before the 26th week of gestation. Phibbs et al. (17) and Chung et al. (18) showed a mostly significantly lower risk-adjusted mortality rate for preterm infants from Californian hospitals with the highest quality levels and annual caseloads of more than 100 than for infants from hospitals at levels 1 to 3c combined with various categories for caseloads of 100 or fewer. However, no monotonic increase in the effect of the caseload across all caseload categories can be determined.

This study also reveals a heterogeneous picture across caseload categories if caseload is divided into 13 categories. This means that some of the results on risk-adjusted mortality are as good or even better in smaller caseload categories than in the group of the largest hospitals. The limited relationship between caseload and mortality is particularly clear in the categories from 50 cases per year upwards, in which risk-adjusted mortality ranges from 6% to 19% (Figure 2). This high variation even in centers with large caseloads is also described by other authors (11, 19, 20) and means that the minimum caseload criterion can distinguish between low and high risk-adjusted mortality only to a limited extent, for both individual hospitals and individual cases. Thus this analysis shows a 44% false negative rate in the assessment of quality of care of individual hospitals on the basis of their caseloads. In addition, only 56% of preterm infants treated in units with caseloads of 30 or more receive above-average quality of care.

The variations in the results even of hospitals with very large caseloads indicate that caseload is only a surrogate parameter for other factors that directly affect quality of care, such as staffing structure, equipment, organization, process and internal quality management, and a culture of quality (21). Future studies should research the role actually played by these factors and how they can be optimized to achieve a high quality of care. One idea would be to use neonatal data gathered according to section 137a of the German Code of Social Law (SGB), Part V to develop tools for quality-based care management that include morbidity parameters in addition to mortality itself.

Good risk adjustment allows direct qualitative comparison of units with different patient populations and competencies. For the purpose of care planning, how-

ever, despite equally high quality of care the successful treatment of low-risk patients cannot be placed on the same footing as equally successful treatment of high-risk patients. Care policy management therefore requires a tool that combines both aspects: competence profiles and quality of care.

KEY MESSAGES

- Overall risk-adjusted mortality is significantly lower for preterm infants with birth weight below 1250 g who are cared for in units with at least 30 cases per year than in preterm infants from smaller units.
- However, risk-adjusted mortality of preterm infants varies even between hospitals with high yearly caseloads.
- Caseload is therefore not a sufficient indicator of quality of care.
- Neonatal data gathered since January 2010 as part of compulsory external quality assurance according to section 137a of the German Code of Social Law (SGB), Part V are now more complete. They therefore can and should be used to develop a quality-based management tool.
- In addition to mortality, such a management tool should include mortality parameters and hospitals' competency profiles.

Conflict of interest statement

The BQS Institute for Quality and Patient Safety (BQS Institut für Qualität & Patientensicherheit) received financial support from the German Hospital Association (DKG, Deutsche Krankenhausgesellschaft e. V.) for this analysis.

Dr. Kutschmann, Dr. Bungard, Mr. Kötting, Mrs. Trümner, and Dr. Veit are employees of the BQS Institute for Quality and Patient Safety. The authors declare that no other conflict of interest exists.

Prof. Fusch has received lecture fees from Abbott.

Manuscript received on 13 September 2011, revised version accepted on 27 April 2012.

Translated from the original German by Caroline Devitt, MA.

REFERENCES

1. AQUA-Institut für angewandte Qualitätsförderung und Forschung im Gesundheitswesen GmbH. Qualitätsreport 2010: Geburtshilfe. www.sgg.de/sqg/upload/CONTENT/Qualitaetsberichte/2010/AQUA-Qualitaetsreport-2010.pdf.
2. Gemeinsamer Bundesausschuss (G-BA): Bekanntmachung eines Beschlusses des Gemeinsamen Bundesausschusses zur Versorgung von Früh- und Neugeborenen. 20 August 2009 (BAnz. Nr. 195 (p. 4450) 24 December 2009. Köln: Bundesanzeiger 2009.
3. Gemeinsamer Bundesausschuss (G-BA). Bekanntmachung eines Beschlusses des Gemeinsamen Bundesausschusses über eine Änderung der Anlage 1 der Mindestmengenvereinbarung: Mindestmengen bei Früh- und Neugeborenen. 17 June 2010. BAnz. Nr. 123 (p. 2840) 18 August 2010. Köln: Bundesanzeiger 2010.

4. Gemeinsamer Bundesausschuss (G-BA). Bekanntmachung eines Beschlusses des Gemeinsamen Bundesausschusses über eine befristete Außervollzugsetzung einer Änderung der Mindestmengenvereinbarung: Mindestmengen für Früh- und Neugeborene Perinatalzentren Level 1. 16. Dezember 2010. BAnz. Nr. 200 (p. 4480) 31. Dezember 2010. Köln: Bundesanzeiger 2010.
5. Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG): Zusammenhang zwischen Leistungsmenge und Ergebnis bei der Versorgung von Früh- und Neugeborenen mit sehr geringem Geburtsgewicht. Köln: Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) 2008.
6. Statistisches Bundesamt Deutschland: GENESIS-Online Datenbank. www.genesis.destatis.de/genesis/online.jsessionid=02F336D5DD5CC255F2F2FF1023F5BD8A.tomcat_GO_1_1?sequenz=statistikTabellen&selectionname=12612. Wiesbaden 2011.
7. Hummler HD, Poets C: Mortalität sehr unreifer Frühgeborener – Erhebliche Diskrepanz zwischen Neonatalerhebung und amtlicher Geburten-/Sterbestatistik. *Z Geburtshilfe Neonatol* 2011; 215: 10–17.
8. Bundesgeschäftsstelle Qualitätssicherung GmbH: BQS-Bundesauswertung 2008 Geburtshilfe. www.bqs-outcome.de/2008/ergebnisse/leistungsbereiche/geburtshilfe/buaw/bu_Gesamt_16n1-GBH_2008.pdf. Düsseldorf: 2009.
9. Fellman V, Hellstrom-Westas L, Norman M, et al.: One-year survival of extremely preterm infants after active perinatal care in Sweden. *JAMA* 2009; 301: 2225–33.
10. Costeloe K, Hennessy E, Gibson AT, Marlow N, Wilkinson AR: The EPICure study: outcomes to discharge from hospital for infants born at the threshold of viability. *Pediatrics* 2000; 106: 659–71.
11. Horbar JD, Badger GJ, Lewit EM, Rogowski J, Shiono PH: Hospital and patient characteristics associated with variation in 28-day mortality rates for very low birth weight infants. Vermont Oxford Network. *Pediatrics* 1997; 99: 149–56.
12. Maier RF, Rey M, Metzke BC, Obladen M: Comparison of mortality risk: a score for very low birthweight infants. *Arch Dis Child Fetal Neonatal Ed* 1997; 76: F146–F151.
13. Moro M, Figueras-Aloy J, Fernandez C, et al.: Mortality for newborns of birthweight less than 1500 g in Spanish neonatal units (2002–2005). *Am J Perinatol* 2007; 24: 593–601.
14. The International Neonatal Network (INN): The CRIB (clinical risk index for babies) score: a tool for assessing initial neonatal risk and comparing performance of neonatal intensive care units. *Lancet* 1993; 342: 193–8.
15. Bartels DB, Wypij D, Wenzlaff P, Dammann O, Poets CF: Hospital volume and neonatal mortality among very low birth weight infants. *Pediatrics* 2006; 117: 2206–14.
16. Trotter A, Pohlandt F: Aktuelle Ergebnisqualität der Versorgung von Frühgeborenen < 1500g Geburtsgewicht als Grundlage für eine Regionalisierung der Risikogeburten. *Z Geburtshilfe Neonatol* 2010; 214: 55–61.
17. Phibbs CS, Baker LC, Caughey AB, Danielsen B, Schmitt SK, Phibbs RH: Level and volume of neonatal intensive care and mortality in very-low-birth-weight infants. *N Engl J Med* 2007; 356: 2165–75.
18. Chung JH, Phibbs CS, Boscardin WJ, Kominski GF, Ortega AN, Needleman J: The effect of neonatal intensive care level and hospital volume on mortality of very low birth weight infants. *Med Care* 2010; 48: 635–44.
19. Heller G: Ermittlung der klinikspezifischen Ergebnisqualität der Behandlung von Früh- und Neugeborenen mit sehr niedrigem Geburtsgewicht (VLBW) auf der Basis von Routinedaten. In: Klauber J, Geraedts M, Friedrich J, Wasem J (eds.): Krankenhaus-Report 2011 – Schwerpunkt: Qualität durch Wettbewerb. Stuttgart: Schattauer 2010; 39–47.
20. Rogowski JA, Horbar JD, Staiger DO, Kenny M, Carpenter J, Geppert J: Indirect vs direct hospital quality indicators for very low-birth-weight infants. *JAMA* 2004; 291: 202–9.
21. Curry LA, Spatz E, Cherlin E, et al.: What distinguishes top-performing hospitals in acute myocardial infarction mortality rates? *Ann Intern Med* 2011; 154: 384–90.

Corresponding author

Dr. rer. medic. Marcus Kutschmann
BQS Institut für Qualität & Patientensicherheit
Kanzlerstr. 4
40472 Düsseldorf, Germany
m.kutschmann@bqs-institut.de

eMethods:
www.aerzteblatt-international.de/12m519

ORIGINAL ARTICLE—eMETHODS

The Care of Preterm Infants With Birth Weight Below 1250 g

Risk-Adjusted Quality Benchmarking as Part of Validating a Caseload-Based Management System

Marcus Kutschmann, Sven Bungard, Joachim Kötting, Andrea Trümner, Christoph Fusch, Christof Veit

Detailed description of methods

Data

With the consent of the steering committees of the quality assurance working groups in German federal states, anonymized neonatal data for 2007 to 2009 from the states of Bavaria, Baden-Württemberg, Hesse, Lower Saxony, and North Rhine–Westphalia were analyzed. These states account for approximately 68% of all births in Germany (5). The data refer to 9997 preterm infants with birth weight below 1250 g from 173 centers.

Analysis included preterm infants for whom the reason for discharge was recorded as “Sent home” or “Died.” Infants with the reason for discharge “Referred back to maternity ward” or “Referred to another hospital/unit” were not included, because the outcome parameter “mortality” cannot be calculated for them. Preterm infants who were admitted to neonatal units more than 24 hours after birth were excluded. It is very probable that these infants had received neonatal treatment before admission, so the outcome of treatment cannot be unambiguously attributed to a single unit. As a result, a total of 1738 cases were excluded, so analysis was based on data from 8259 preterm infants. It was analyzed whether the exclusion of the 1738 referred infants leads to a bias regarding the question the research aims to address: Investigation was carried out into whether the excluded cases were distributed evenly between hospitals with small and large caseloads.

Risk-adjustment model

In order to develop a robust risk-adjustment model, a systematic review of the literature was performed initially. Risk factors that affect the outcome parameter mortality were identified. We considered only risk factors that are not principally affected by care but play a main role in determining preterm infants’ risk profiles. Searches were performed in PubMed and the Cochrane Library using specific keywords and free search terms (e.g. “very low birth weight,” “mortality”). The risk-adjustment model was then developed using multiple logistic regression with mortality as the outcome parameter. Variables were selected (with a 5% significance level) using the SPSS procedures “forward

selection” and “backward selection,” as well as manually. The p-value of the Hosmer–Lemeshow test was used as a measure of the fit of the model. All analyses were performed using PASW Statistics, version 18.

The logistic regression model provides the probability of death for each child, on the basis of his/her individual risk profile. Calculating the mean of these probabilities for each hospital yields the mortality rate that would have been expected according to the risk profiles of the preterm infants treated in the hospital. If this expected mortality (E) is compared with the actually observed mortality (O), the quotient $O \div E$ provides a measure for risk-adjusted mortality. If $O \div E$ is less than 1, which means that the observed mortality is less than the mortality expected according to the preterm infants’ risk profiles, this can be considered evidence of above-average quality of care. If, in contrast, observed mortality is higher than expected mortality ($O \div E$ greater than 1), this indicates below-average quality of care. $O \div E$ can be calculated not only for individual hospitals but also for groups of hospitals based on particular criteria, such as caseload.

Only datasets for which all risk factors have been fully recorded can be used to calculate $O \div E$. This is the case for 7405 (89.7%) of the 8259 preterm infants of the total sample population. In order to examine whether this led to a bias, mortality in infants whose data were not included because their records were incomplete ($n = 854$, 10.3%) was compared with the mortality of infants whose data were fully recorded.

The effect of volume of care

After the risk-adjustment model had been developed, the variable “caseload” was incorporated into it. This made it possible to assess the effect of volume of care on risk-adjusted mortality in the preterm infants. Caseloads recorded in 2009 were used to categorize hospitals by caseload.

The variable “caseload” was processed in two different ways. On the one hand, it was divided into two categories, “2009 caseload fewer than 30 preterm infants” and “2009 caseload 30 or more preterm infants,” in line with the minimum caseload criterion. This made it possible to assess whether adjusted mortality really was

lower overall in hospitals with a volume of care of at least 30 cases than in hospitals with smaller volumes of care. On the other hand, caseload was treated as a discrete variable with 13 categories, “fewer than 5,” “5 to 9,” and so on up to “55 to 59” and “60 or more.” Using these categories made it possible to examine the caseload categories that were particularly relevant to the minimum caseload criterion (14 cases for the previous minimum caseload criterion, and 30 for the current criterion) more closely. It also allowed these categories to be compared directly with the reference category “ ≥ 60 ,” in which hospitals supposedly provide particularly high levels of quality. The value of 0.05 was chosen as the significance level in both cases.

The discriminative ability of the minimum caseload criterion

The discriminative ability of the minimum caseload criterion was analyzed in order to assess its suitability as a care management tool. Model-fit measures were calculated in a similar fashion as for diagnostic tests. The false negative rate describes the number of patients classified as healthy (negative test result) as a percent-

tage of the total number of patients who are actually ill. The positive predictive value is the number of patients who are actually ill as a percentage of the total number of patients classified as ill (positive test result). The false negative rate and the positive predictive value are measures of fit that are also relevant here.

In terms of the structure of care, the percentage of hospitals in the group of hospitals with above-average quality of care ($O \div E < 1$) that would no longer provide care according to the minimum caseload criterion (because they treat fewer than 30 cases) is of interest. Therefore, at hospital level the false negative rate was calculated. Parents of preterm infants, however, are interested in knowing which hospital provides the highest chance of their child receiving above-average quality of care ($O \div E < 1$). According to the minimum caseload criterion, parents must assume that it is highest in large hospitals (30 cases or more). Calculating the positive predictive value at patient level therefore estimates the probability that infants are actually treated in large hospitals with above-average quality of care.