medRxiv preprint doi: https://doi.org/10.1101/2024.11.22.24317721; this version posted November 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

Prevention of Transmissions by Effective Colonisation Tracking in Neonates (PROTECT-Neo)

Timmy Nguyen[1*], Fabian Bürkin[1*], Stefany Ayala Montaño[1*], Daniel Jonas[1], Martin Kuntz[2], Tjibbe Donker[1], Sandra Reuter[1#], Tim Götting[1#], Philipp Henneke[1,3#], for the PROTECT-Neo-Study-Group

University Medical Freiburg, Germany:

[1]*Institute for Infection Prevention and Control,*

[2]*Center for Pediatrics and Adolescent Medicine,*

[3]*Institute for Immunodeficiency (CCI)*

Correspondence to:
Philipp Henneke; philipp.henneke@uniklinik-freiburg.de
Tim Götting; tim.goetting@pkd.de

Abstract

Importance: Patients in neonatal intensive care units (NICU) are at risk of transmission events (TE) by bacteria with multidrug-resistance or with epidemic potential, such as *Enterobacterales*, *P. aeruginosa, S. marcescens*, *A. baumannii*, MSSA and MRSA, which may precede invasive infections ("multidrug-resistant organisms plus", MDRO+). Resolution of MDRO+ transmission clusters (TC) is important to invoke timely and thus effective infection prevention control (IPC) measures. **Objectives**: Exploration of the potential of timely whole genome sequencing (WGS) in resolving putative MDRO+ transmission chains on a NICU, based on screening-isolates. Exploration of patient- and environmental-specific risk factors of becoming part of a TC. **Design, Setting, Participants**: Prospective monocentric cohort study at a level III NICU of the University Medical Center Freiburg, Germany. Inclusion of 434 of 551 patients (preterm and term) that were on the NICU for at least 48h and screened at least once in the time between February 15 2019 and November 16 2020. **Exposures**: Integration of (1) routine culture-based screening for MDRO+ (at admission, then once a week; in accordance with national guidelines), (2) genetic typing with amplified fragment length polymorphism (AFLP) and WGS and (3) granular clinical and staffing data. Statistical analysis of time-dependent risk factors based on multivariate model analysis, using logistic regression and moving averages (MA). **Main Outcomes**: Primary outcome: Identification of bacterial transmission events (TE) with pathogens of the same species in different patients, which were indistinguishable by AFLP or WGS. Secondary outcomes: MDRO+ colonization rates; identification of TE influencing factors; blood stream infection (BSI) rates.
Results: Among 434 NICU patients (27.9% [95% CI, 23.9%-32.3%] with birth weight < 1,500 g), 51.8 % (95% CI, 47.1%-56.5%) were colonized with at least one MDRO+ species; 32.5% (95%CI, 28.3%-37.0%) were colonized as part of a TE as revealed by WGS. With 38 unique TC, *E. coli* was the most common cluster-forming MDRO+, whereas *K. oxytoca* formed the largest cluster involving up to 19 patients. Out of ten bloodstream infections, four originated from TE. Multivariate model analysis revealed three key factors for the risk of becoming part of a TC: Increased nurse staffing levels and antibiotic administration lowered the risk of being part of a bacterial transmission cluster, while vascular catheter usage increased it. **Conclusions and Relevance**: Prospective WGS of routine screening isolates from newborn infants in intensive care is a powerful tool for resolving MDRO+ transmission chains, exceeding AFLP in precision. Despite the associated costs, systematic strain identification by WGS seems justified in high-risk neonates. Both delayed TE identification and "false" TE/TC, which inevitably occur in conventional microbiological screening, have grave organizational consequences. Within model boundaries, we observe factors influencing the risk of becoming part of a bacterial transmission cluster.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.11.22.24317721; this version posted November 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

Introduction

In healthy infants, the exit from the sterile womb and the establishment of metabolic independence are accompanied by swift acquisition of maternal microorganisms from birth canal, skin and breast milk. This drives a dynamic microbiome development in a postnatal relationship between mother and infant best termed 'separate, but intertwined'. However in hospitalized and in particular in preterm infants, the microbiome receives major input from hospital-adapted microorganisms residing in other patients and on inanimate surfaces. Conceptually, these microorganisms are qualitatively different because they have been shaped e.g. by hosts that carry diseases, by the selection pressure of antibiotics and by the hospital environment. In contrast to older children and adults, the microbiome of infants, in particular those born prematurely, is less stable, i.e. less resilient to the incorporation of new strains. The largely increased risk for serious infections in newborn infants acutely links microbiome composition to life-threatening infections. This susceptibility is held responsible for the dreaded infection outbreaks on Neonatal Intensive Care Units (NICUs). The difficulty of preventing transmission events (TE) in infants even with sophisticated hygiene measures was recently confirmed in a randomized probiotic study which we conducted: 48.8% (95% CI, 38.4%-59.3) of the infants in the placebo group, who had no direct contact with the probiotics, carried the probiotic *Bifidobacterium* strain [1]. Overall, the gradual colonization of preterm infants on NICUs includes bacteria with multidrug resistance (MDR) or with increased potential for patient-to-patient transmission [2–5]. Nosocomial bacterial colonization, broad spectrum antibiotics and invasive indwelling catheters coincide to increase the risk of nosocomial infections (NI) [6] with increased mortality rates [7]. Nosocomial bacterial colonization and potential infection therefrom is furthermore linked to long term adverse effects such as sepsis induced immunosuppression [8] and microbiome disturbance [9]. Factors influencing bacterial colonization include prematurity, mode of delivery [10,11], antibiotic therapy [12,13] and the hospital environment [14]. However, factors driving transmission events on NICUs remain largely elusive. The spatial and timely clustering of bacteria of the same species strongly suggests an outbreak [15] and should be investigated with appropriate typing methods to confirm or refute putative TE [16]. According to multiple reports of outbreaks on NICUs, *Enterobacterales* prevail as the most common causative agents. Despite extensive infection prevention control (IPC) measures, including environmental investigations, there is frequent inability to find a point source [17,18], indicating that transmissions are primarily driven by indirect contact transmissions, e.g. due to improper hand hygiene rather than environmental point sources. In Germany, weekly microbiological surveillance based on cultural screening techniques, to systematically identify, track and possibly prevent outbreaks, has been recommended by the Commission for Hospital Hygiene and Infection Prevention (KRINKO) at the RKI (Robert Koch-Institute) since 2012 [19]. The culture-based screening identifies the presence of potentially problematic bacterial species as defined by the KRINKO, i.e. species belonging to the order of *Enterobacterales* and non-fermenters declared as multi resistant gram negative (MRGN) bacteria, and methicillin resistant *Staphylococcus aureus* (MRSA) [20], together with their non-MRGN counterparts, which are collectively referred to as "multidrug resistant organisms plus" (MDRO+) in this study. MDRO are relatively rare on NICUs in Germany (and in many other European countries), i.e. parallel detection of MDRO of similar species in two or more patients is highly indicative of a TE. Yet, this is not the case for parallel detection of the more frequent non-MDR *Enterobacterales*, since they often represent distinct strains of the same species, yet with similar resistance pattern. Accordingly, the extent of the recommended and potentially disruptive organizational IPC measures, ranging from patient

medRxiv preprint doi: https://doi.org/10.1101/2024.11.22.24317721; this version posted November 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

and staff cohorting to environmental sampling and ward closures, is often in conflict with an inaccurate characterization of MDRO+ organisms and their genomic relationship. Moreover, a stepwise complex typing of isolates usually takes place with considerable delay, so that the necessity of the hygiene measures taken can only be assessed retrospectively.
Here we asked in which way the timely and accurate detection of TE by amplified fragment length polymorphism (AFLP) and whole genome sequencing (WGS) could affect the recognition and management of TC and how, at the proposed granular level, treatment and management strategies influence the emergence and development of MDRO+ clusters in the NICU.

Methods and Analysis

Design, Setting, Participants

The premise of this single-center prospective cohort study was the 17-bed Level III NICU at the University Medical Center Freiburg, Germany, which is one of the largest tertiary care hospitals in Germany with around 2200 beds. The NICU treats both preterm and term newborn infants, covering the entire spectrum of diseases and treatment of congenital malformations. 434 newborns (preterm and term) were enrolled from February 15 2019, to November 16 2020. Patients were included in the study if admitted to the NICU for ≥ 48 hours and if screened at least once. Patients were excluded from the study with a length of stay < 48 hours and/or if no screening was performed. The study outline and key metrics are shown in *Figure 1*. The initial study protocol was published beforehand [21].

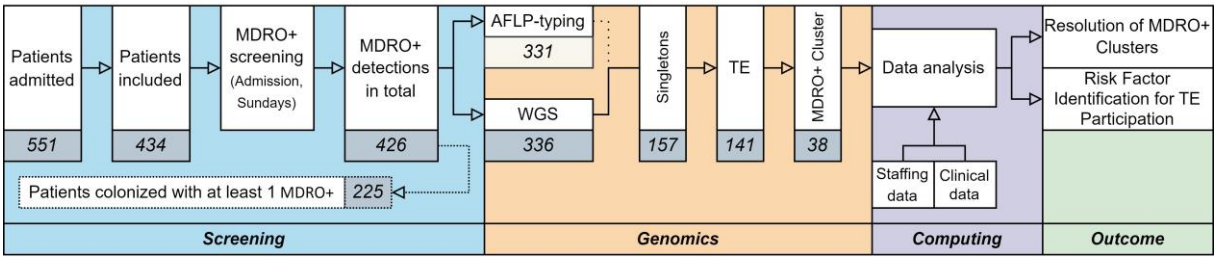

Figure 1: *Study Outline Flow Chart*
Bacterial isolates (MDRO+) from weekly screenings were used for genomic cluster determination via AFLP-typing and WGS. The genomic information was combined with clinical and staffing data. The genomic analysis resolved putative MDRO+ clusters and the statistical analysis identified associated influential factors of being part of a transmission cluster within its model's boundaries.
*7940 patient days, **95% CI, 24.8-32.3, ***95% CI, 48.7-59.0, ****95% CI, 15.0-20.9

Outcomes

Primary outcome: Identification of bacterial transmission events with pathogens of the same species in different patients, which were indistinguishable after molecular or genomic typing.
Secondary outcomes: Rate of patients colonized with at least one MDRO+ per 1,000 patient days, rate of MDRO+ colonization per 1,000 patient days, identification of influential factors of transmission events, blood stream infection rates per 1,000 patient days.

Bacterial Transmission – Definitions

Transmission cluster (TC): Group of two or more bacterial isolates obtained from different patients that are genetically indistinguishable based on molecular typing methods (WGS), including the index patient.

medRxiv preprint doi: https://doi.org/10.1101/2024.11.22.24317721; this version posted November 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

Transmission event (TE): Instance where a bacterial isolate is transmitted from one patient to another patient, as revealed by molecular typing methods (WGS), excluding index patients.

Pre-Existing (Study-Independent) Screening Regimen of the NICU

In adherence to the KRINKO-recommendations [22], culture-based bacterial screenings at admission or birth as well as every subsequent Sunday were performed. The swab sites included nasopharyngeal, anal, rectal as well as other risk sites, e.g. stomata if applicable.

Routine Cultural Diagnostics

Screening swabs (eSwab, Copan, Brescia, Italy) and clinical samples were inoculated on both selective and non-selective media including following agar plates: Columbia Blood, MacConkey, Hematin Chocolate Agar, chromID® (Biomérieux, Nuertingen, Germany) Extended spectrum betalactamase (ESBL)/MRSA. Media were incubated under aerobic conditions for 48 hours at 36°C and with 5% CO_2. If growth on plates was detected, matrix-assisted-laser-desorption-ionization-time-of-flight (MALDI-TOF) mass spectrometry (Bruker Daltonics, Bremen, Germany) was employed to identify bacterial species. Susceptibility testing was performed using VITEK2® (Biomérieux, Nuertingen, Germany) or MIC test strips (Liofilchem, Piane Romano, Italy), respectively, and the results were interpreted according to EUCAST (European Committee on Antimicrobial Susceptibility Testing) clinical breakpoints. Presence of resistance genes was confirmed by nucleic acid amplification tests for carbapenem-resistant gram-negative bacteria and MRSA.

Collection and Preparation of Isolates

If MDRO+ species were found, bacterial isolates were transferred to the typing laboratory. The study included the following bacteria in our genetic and statistical analysis and are referred to as “MDRO+”, specifically: *A. baumanii, C. freundii, E. coli, E. cloacae, K. oxytoca, K. pneumonia, P. mirabilis, S. marcescens, P. aeruginosa* and *S. aureus* (methicillin sensitive *Staphylococcus aureus* [MSSA], MRSA) on the basis of the initially published study protocol [21].

Collection of Healthcare-Worker-Specific Procedural Data

Procedural data contained anonymous daily and shift-specific metrics on bed occupancies, nurse staffing levels and application of minimum staffing criteria. The patients’ birth weight, maturity level and the need for intensive care or intensive monitoring guided definitions of ‘overstaffing’ and ‘understaffing’, which are based on algorithms predefined by the G-BA (German: “Gemeinsamer Bundesausschuss”) [23] at shift level.

Genetic Typing of Bacterial Strains

AFLP-typing for gram-negative bacteria, *spa*-typing for *S. aureus* (MSSA, MRSA) and WGS were performed for typing of bacterial species.

AFLP

AFLP typing from pure bacterial cultures was performed with the Genetic Analyzer abi3500. The comparison of the resulting fragment patterns was software-assisted (abi3500 analysis software „Genemapper“ or BioNumerics). The assignation of fragment differences to genotypes was performed as described [24,25].

medRxiv preprint doi: https://doi.org/10.1101/2024.11.22.24317721; this version posted November 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

WGS

Deoxyribonucleic acid (DNA) extraction was performed with a High Pure polymerase chain reaction (PCR) Template (Roche) following standard protocol. Sequencing was carried out using the Illumina MiSeq Nextera DNA Flex Library pre-preparation and V2 300 cycle PE kit according to manufacturer's instructions. Bioinformatics analysis was carried out using smalt [26], samtools v0.1.19 [27] and GATK (mapping) [28], SPAdes v3.13.1 [29] with kmer sizes 21, 33, 55, 77, 99, 109, and 123 with filtering to only include contigs with a minimum of 500bp (assembly), kraken v1.1.1 [30] (species identification), mlst v2.10 [31] (multilocus sequence typing [MLST] identification). Quality control parameters were a minimum of 30X coverage, appropriate length of the sequence, number of contigs <500, N50 <100,000bp, correct species and MLST identification. Phylogenetic reconstruction was based on core genome alignment, single nucleotide polymorphism (SNP) identification (snp-sites v2.5.1, snp-dists v0.6), followed by estimation using RAxML v8.2.12. iTOL was used for phylogenetic visualization. Sequencing data has been deposited in the ENA project PRJEB81699 and individual accession identifiers can be found in the supplemental material.
If bacterial isolates are indistinguishable from one another via WGS, based on respective cut-off values, we interpret them as resulting from transmission events. The first chronological appearance of an isolate in a cluster was regarded as the index case. The minimum number of indistinguishable bacterial isolates required to form a cluster is two.

Data Management

Data was collected with Microsoft Access, further analyzed with Microsoft Excel as well as the statistical Software R (Version 4.4.1). Patient data pseudonymization was performed electronically using hash-functions. We collected clinical data from patient charts or electronic systems (laboratory information system, clinical information system) including patient- and ward-specific data, potential risk factors and confounders as well as procedural and operational data. Information on the case-report-form is provided (*eTable 1 in Supplement*). Outliers and influential values were not corrected or excluded, as extreme situations were considered particularly interesting for the analysis due to everyday hospital hygiene practice and their clinical relevance. Of the 7940 patient days included, according to the study protocol, 7037 had all the data required for statistically modeling the associated risk factors.

Statistical Analysis

Following Barnett & Graves [32] and Breslow et al. [33], we mapped the temporal dynamics of the multivariate risk factor analysis for transmission events using a logistic regression model. This allowed for a much more differentiated view of relevant time periods before a transmission event could be detected. We also analyzed mixed models. However, the analysis was conducted based on logistic models since random parts (patients, time and both) were not significant.

Dependent Variable Y_t: Transmission Events

The dependent variable Y_{t+1} is a binary representation of the transmission events at time *t+1* ($TE_{(t+1)}$) at the individual patient level with a value of 1 for a day with positive MDRO+ screening and 0 on days with negative or no MDRO+ screening. The postponement of Y in the form of *t+1* instead of t has practical reasons: If a screening is positive on day *t+1*, with the incubation

medRxiv preprint doi: https://doi.org/10.1101/2024.11.22.24317721; this version posted November 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

period the transmission took place on day *t* at the latest, especially since a large proportion of the screenings took place in the morning.

Individual Variables – Per Patient		
Variable	***Type***	***Explanation***
Sex	Binary	Sex of patient
Delivery	Categorical	Type of birth (C-section emergency, C-section primary, C-section secondary, vaginal)
Birth Weight	Continuous	Birth weight of patient in grams
INPULS	Continuous	Based on categorical severity score INPULS® in intensive care units (treated as a continuous variable in order to be able to map different temporal dynamics via moving averages)
Breast Milk	Binary	1 if the patient received breast milk, 0 if not
Feeding Tube	Binary	1 if the patient had a feeding tube in place, 0 if not
Kangarooing	Binary	1 if the patient experienced kangarooing, 0 if not
Antibiotics	Continuous	Number of different antibiotics administered, regardless of dose and frequency (Ampicillin, Tobramycin, Piperacillin/Tazobactam, Meropenem, Vancomycin, Cefuroxim, Erythromycin, Rifampicin, Cefotaxim, Levofloxacin, Teicoplanin, Linezolid, Ceftazidin, Cotrim, Amphotericin B)
Catheter	Continuous	Number of different catheters (venous [peripheral, central] and arterial)
Invasive Ventilation	Binary	1 if the patient was ventilated with a breathing tube, 0 if not
Environmental Variables – Per Ward		
Variable	***Type***	***Explanation***
MDRO+	Continuous	Total number of MDRO+ present per day
Bed Occupancy	Continuous	Number of patients, averaged over all three shifts
Nurse Patient Ratio	Continuous	Nurse patient ratio, averaged over all three shifts
Staff: FTE	Continuous	Nursing staff totalled across all three shifts in full-time equivalents (FTE)
Staff: Deviance	Continuous	Deviation between the nursing staff required by the G-BA and the staff actually deployed in full-time equivalents, averaged over all three shifts
Staff: Understaffed	Continuous	Number of missing full-time nursing staff compared to the requirement determined by G-BA, totalled across all three shifts
Staff: Overstaffed	Continuous	Number of additional full-time nursing staff compared to the requirement determined by the G-BA, totalled across all three shifts

Table 1: *Candidate Set for Potential Confounders* $X_{(i,\,t)}$
Differentiation of potential confounders for multivariate analysis in individual (patient specific) and environmental variables (ward specific). Potential confounders were also considered as equally weighted moving averages over the time (up to 14 days in the past).

Independent Variables $X_{(i,\,t)}$

The candidate set for potential confounders $X_{(i,\,t)}$ with *i in {1, …, 18}* in the model included 18 variables (*Table 1*). The variables $X_{(i,\,t)}$ were also considered as equally weighted moving averages (MA) over the time periods *[t; t], [t; t-1], [t; t-2], …, [t; t-14]*, so that a total of 15 temporally nuanced variants of each variable were available for selection.
For *d* in *{0,1, …, 14},* the respective moving average was defined as:

$$\text{MA: d} := \sum_{j=t}^{t-d} \frac{X_{(i,j)}}{d+1}.$$

In our model e.g. "MA: 2" represents the unweighted average over the two days *t*, the previous day *t-1* and *t-2* (*Figure 2*).

medRxiv preprint doi: https://doi.org/10.1101/2024.11.22.24317721; this version posted November 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

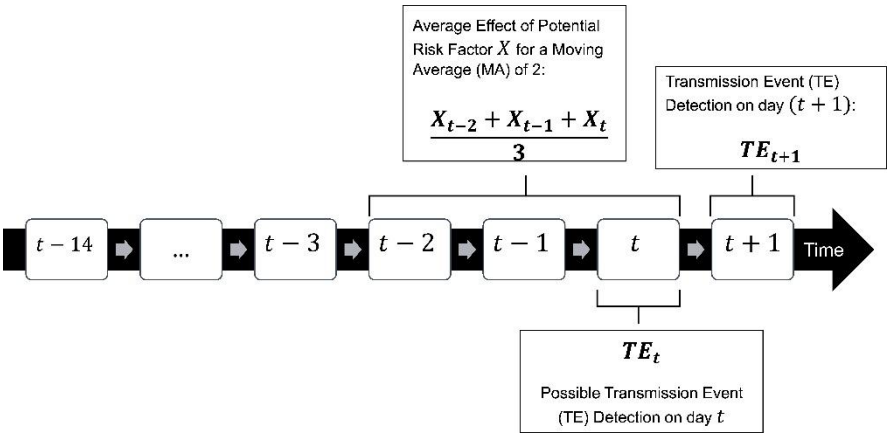

Figure 2: *Illustration of the temporal relationships between the detection of transmission events TE and potential influencing variables X_i*
The variable TE can only determine past transmissions, but not the exact time of the transmission event. In order to consider this measurement's inaccuracy, the various independent variables X_i are also considered in different time periods with the aid of moving averages.

The optimal model was chosen by the minimum Akaike's information criterion (AIC). Predictors were selected by using stepwise forward-backward selection: Starting from a null model, all univariate models were tested. The combination of confounder and moving average with the best AIC was selected. In order to obtain an interpretable model, we excluded all other moving averages of that predictor from the pool of available variables. If there were several variants of predictors that measured the same target variable (such as for the measurement of staff occupancies), all other predictors were also excluded from the set for further modelling when one variable was selected. After the inclusion of a new predictor in the model, each variable already included was also tested according to the 'leave-one-out' principle in order to see whether there were better alternatives in the current model that would lead to a lower AIC score, leading to the replacement of the corresponding variable by the better alternative. If this was the case, all combinations were tested again using the 'leave-one-out' method. If the model could not be further improved after the inclusion of at least two additional variables and corresponding 'leave-one-out' optimizations, the final model was determined based on the lowest AIC. The final model was then tested for multicollinearity using the variance inflation factor (VIF) and for linearity using a rainbow test [34,35] to ensure the necessary conditions for logistic regression.

Results

Demographics

During the study period, 434 out of 551 patients admitted to the NICU were included into the study cohort. 113 patients did not meet the inclusion criteria. Four patients were not included due to missing data. We observed 7940 patient days. Among all enrolled patients, 121 [27.9% (95% CI, 23.9%-32.3%)] had a birth weight of < 1,500 g (median birth weight 980.0 g, median gestational age of 28.7 weeks) with primary caesarian section as the most common mode of delivery [52.1% (95% CI, 42.8%-61.2%)]. 313 [72.1% (95% CI, 67.7%-76.1%)] patients had a birth weight ≥ 1,500 g (median birth weight 2,590.0 g, median gestational age of 36.7 weeks). The duration of hospital stay was longer in patients with a birth weight < 1,500 g (median of 29.0 days) as compared to patients with a birth weight ≥ 1,500 g (median 6.0 days) (*eTable 2 in Supplement*).

medRxiv preprint doi: https://doi.org/10.1101/2024.11.22.24317721; this version posted November 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

Colonization

225 patients were colonized with at least one MDRO+, equaling to 51.8 % (95% CI, 47.1%-56.5%) among all enrolled patients (n = 434). The MDRO+ colonization rate with at least one MDRO+ was 28.3 (95% CI, 24.8-32.3) per 1,000 patient days. The rate of MDRO+ colonizations per 1,000 patient days was 53.7 (95% CI, 48.7-59.0). Among MDRO+, *E. coli* was the most frequent species (n = 114) with a colonization rate of 14.4 (95% CI, 12.0%-17.2%) per 1,000 patient days (*Figure 3A*). Apart from *E. coli*, other MDRO+ with a frequency between 5 and 10 % included MSSA (n = 59), *K. oxytoca* (n = 54), *E. cloacae* (n = 53), *K. pneumoniae* (n = 49) and *K. oxytoca / Raoultella spp*. (n = 44) (*Figure 3A*). Patients with a birth weight < 1,500 g were most commonly colonized by two different MDRO+ species. Patients with a birth weight ≥ 1,500 g were most commonly colonized by zero MDRO+ species. Five patients were colonized with up to six different MDRO+ (*Figure 3B*).

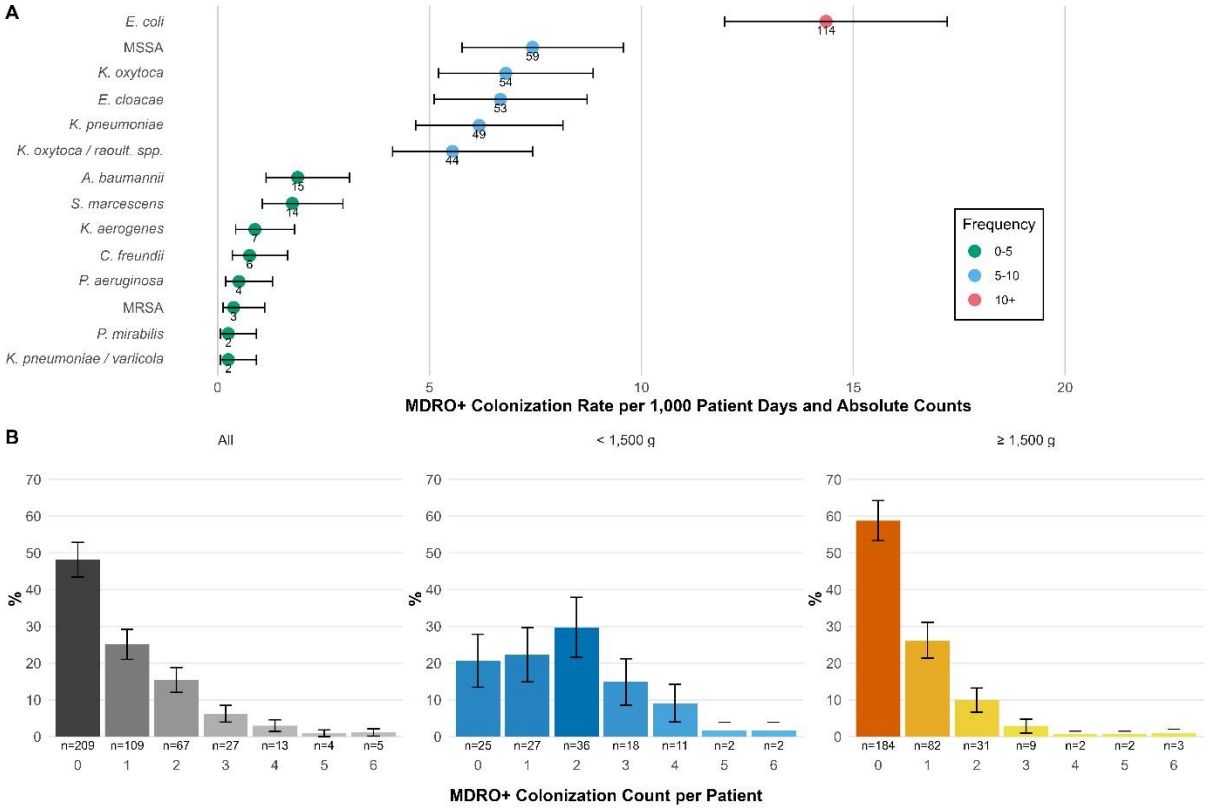

Figure 3: *MDRO+ Colonization – Descriptive*
A) Stratification of colonization rates per 1,000 patient days of MDRO+ by species with 95% CIs. Rates were attributed to frequency categories with color indicators. For every rate, the absolute count of respective MDRO+ isolates were stated. **B)** Distribution and counts of patients that experienced non-, single- or multiple MDRO+ colonizations with 95% CIs. Patient representation in groups: all (grey), birth weight < 1,500 g (blue), birth weight ≥ 1, 500 g (orange).

medRxiv preprint doi: https://doi.org/10.1101/2024.11.22.24317721; this version posted November 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

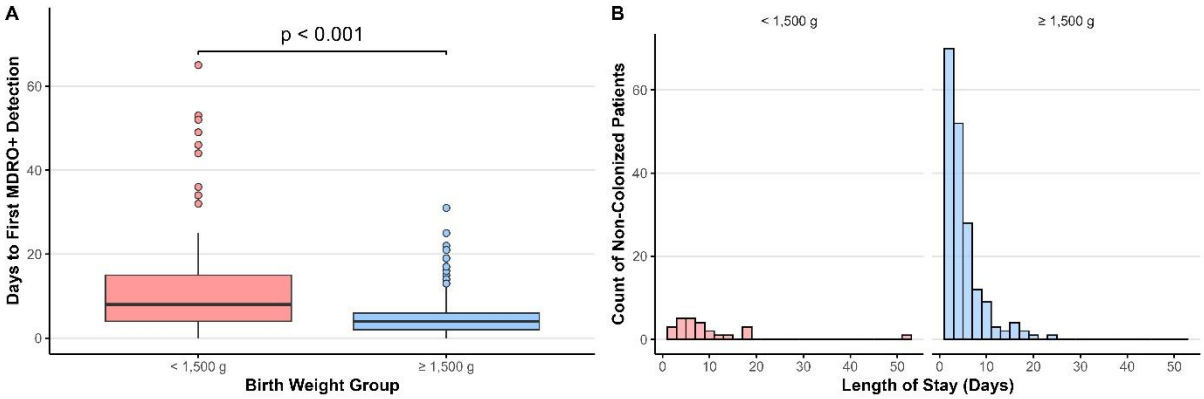

Figure 4: *MDRO+ Colonization – Dynamics*
A) Number of days from admission to first MDRO+ detection (regardless of species) by birth weight groups. Patients with birth weight ≥ 1,500 g were colonized with their first MDRO+ significantly earlier (Wilcoxon Rank-Sum Test, p < 0.001). **B)** Distribution of length of stay of patients that were not colonized by MDRO+ and stratification by birth weight.

In patients with a birth weight ≥ 1,500 g first detection of a MDRO+ in screening occurred significantly earlier [median 4 days] than in patients with a birth weight of < 1,500 g [median 8 days] (*Figure 4A*). Among patients that were never colonized by any MDRO+, the mean length of stay on the ward was 5.37 (95% CI, 4.79-5.95) days for patients with a birth weight of ≥ 1,500 g and 9.96 (95% CI, 6.01-13.9) days with a birth weight of < 1,500 g respectively (*Figure 4B*). When stratifying for MDRO+ species, there was no significant difference in days to first MDRO+ detection (*eFigure 1 in Supplement*).

Transmission Clusters

WGS identified 38 unique transmission clusters. *E. coli* clusters were most frequent (n = 11). The largest cluster was formed by *K. oxytoca* including 19 patients. *S. marcescens* formed a single cluster of seven patients (*eTable 3 in Supplement*). The timely distribution of the occurrence of a TE is represented in the supplemental content (*eFigure 2A in Supplement*). The overall transmission rate was 17.8 (95% CI, 15.0-20.9) per 1,000 patient days, where *E. coli* had the highest individual rate of 6.9 (95% CI, 5.3-9.0) per 1,000 patient days (*eFigure 2B in Supplement*). Overall, WGS identified 157 bacterial isolates that were not part of transmission clusters (i.e. “singleton”) and 179 isolates that contributed to clusters (*eTable 4 in Supplement*).
The phylogeny of *E. coli* isolates (dominant MDRO+) detected in the study is represented in supplemental content (*eFigure 3 in Supplement)*.

Discriminatory Capabilities of WGS

We compared the precision and validity of AFLP and WGS methods in our cohort. 317 out of 348 isolates that underwent WGS were also analyzed by AFLP- or *spa*-typing. For example, AFLP-Type “D” of *E. coli* that corresponded to Sequence Type (ST) 141 in all cases could be further resolved into three distinct WGS clusters 201903_ST141, 201905_ST141 and 201908_ST141 (*Figure 5A, eTable 5 in Supplement*).
Regardless of cluster composition, *E. coli* ST 141 was the predominant sequence type (*Figure 5B*). It accounted for 42.4% (95% CI, 31.2%-54.4%) of cluster-contributing *E. coli* isolates, followed by *K. oxytoca* ST 176, accounting for 52.6% (95% CI, 37.3%-67.5%) of cluster contributing *K. oxytoca* isolates.

medRxiv preprint doi: https://doi.org/10.1101/2024.11.22.24317721; this version posted November 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

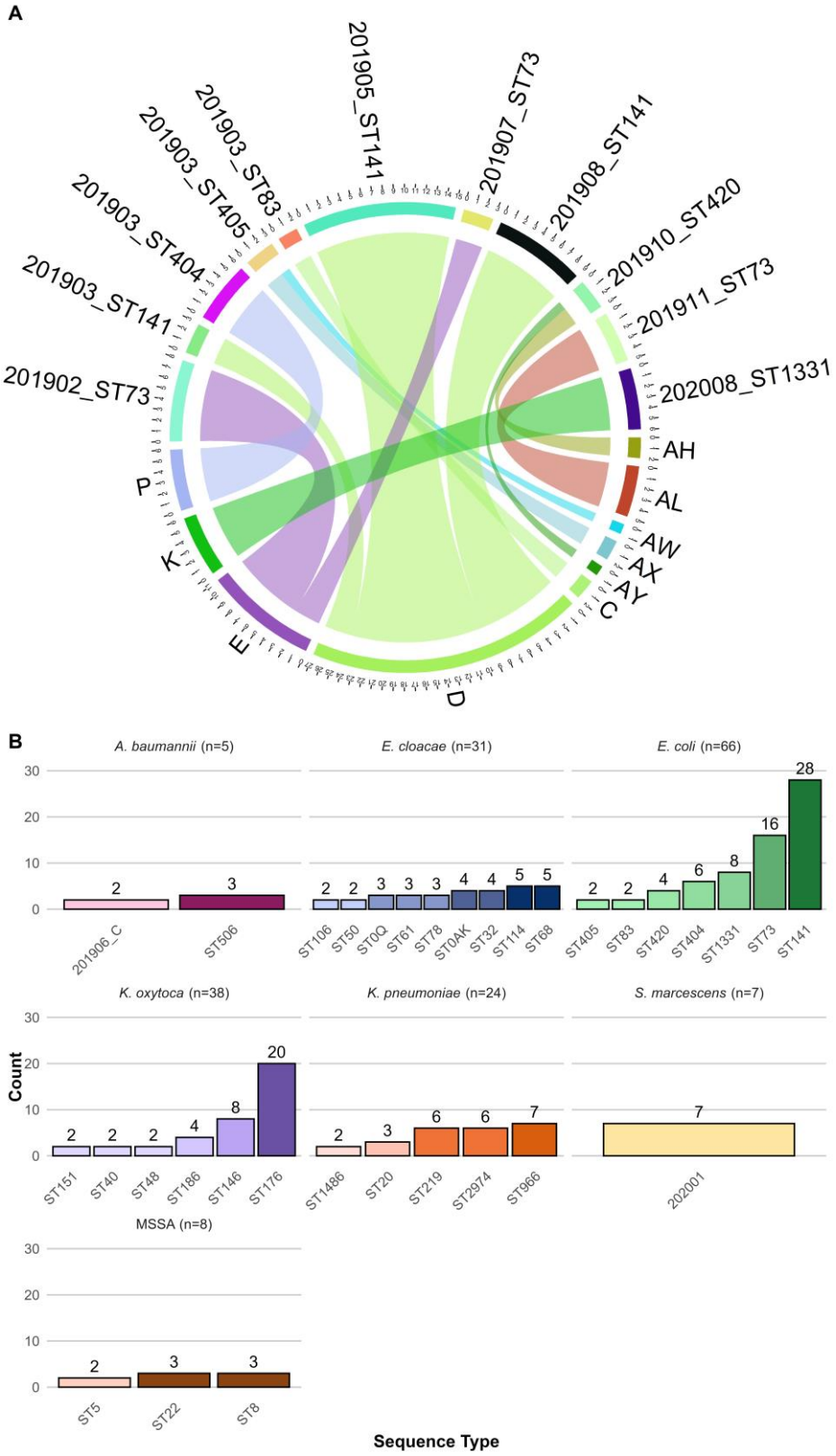

Figure 5: *Variable Cluster Definitions*

A) AFLP types (single- and dual-letters codes) were matched to their respective WGS clusters (year/month/underscore/ST/Number) in the chord diagram. Allocation of one AFLP type to various WGS-clusters was defined as a mismatch. For example: Perceivable mismatches for AFLP type "D" (green) and "E" (purple), where one AFLP type was attributed to multiple WGS-clusters; perceivable mismatches for AFLP-type "AY" and "AH", where both AFLP-types belonged to one WGS-cluster. **B)** Stratification of cluster contributing MDRO+ isolates on ST level. The dominant species in our study were (descending): *E. coli*, *Klebsiella spp.* and *E. cloacae*.

medRxiv preprint doi: https://doi.org/10.1101/2024.11.22.24317721; this version posted November 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

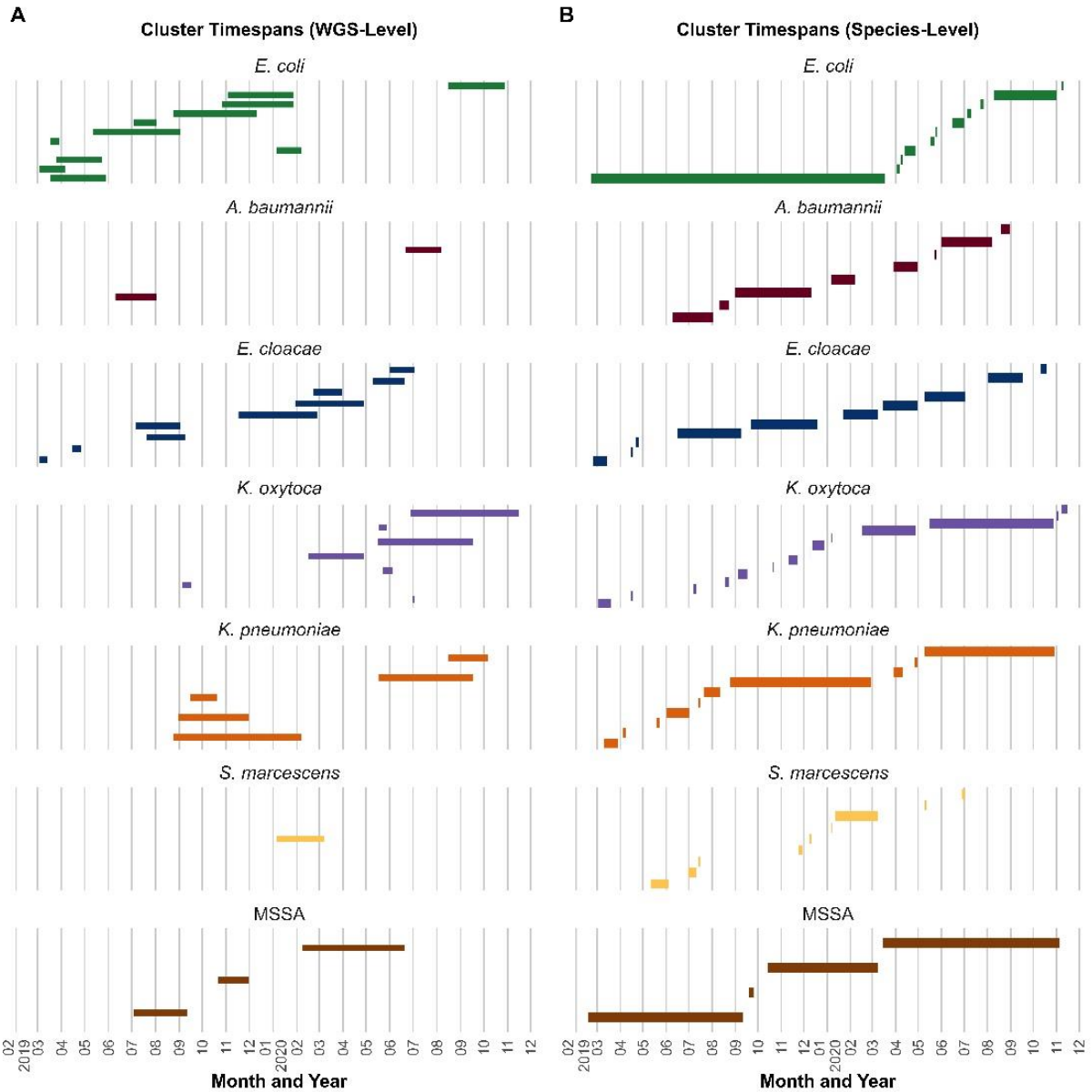

Figure 6: *Occurrence of MDRO+ Clusters According to Different Definitions*
A) Cluster definition according to WGS: The combined timeframes of residing patients colonized with WGS-cluster-contributing isolates were represented over time and indicated by horizontal bars. The day of index case detection defines the start of the timeline. For respective clusters, the discharge of the last patient colonized with a WGS cluster contributing isolate defined the end of the timeframe. Singletons (non-cluster contributing isolates) were excluded in this representation. **B)** Cluster definition according to species alone: Overlapping timeframes from both figures for respective MDRO+ indicated a gain in resolution in favor of WGS when compared to cluster determination by species alone, best exemplified by *E. coli*.

When comparing AFLP typing to WGS cluster identification mismatch rates between the methods were 0.10, 0.09 and 0.06 for *K. oxytoca*, *E. cloacae* and *E. coli* respectively (*eTable 6 in Supplement*).
Same species MDRO+ can cluster differently depending on the strain definition. For example, "*E. coli* Cluster 1" (lowest green bar in *Figure 6B*) is defined as the timely overlap of patients colonized with *E. coli* on the ward and formed a large cluster which can be further divided into distinct clusters as defined by WGS (*Figure 6A, 5B, eTable 7 in Supplement*), leading to an additional resolution of 9 clusters. Increases in cluster resolution were observed for other

medRxiv preprint doi: https://doi.org/10.1101/2024.11.22.24317721; this version posted November 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

species clusters as well, e.g. "*K. oxytoca* cluster 12", leading to an additional resolution of 4 clusters (*eTable 7 in Supplement*).

MDRO+ Blood Stream Infections

In the study period, we detected 10 MDRO+ bloodstream infections (BSI) with *E. coli* [5], *K. pneumonia* [3] and *K. oxytoca* [2]. Four patients had BSI derived from a TE (as revealed by WGS), of which three were detected with a same strain MDRO+ beforehand and admitted with a birth weight of < 1,000 g (*Figure 7*). The BSI isolates belonged to WGS clusters 201907_ST73, 201911_ST73, 201908_ST219, 202006_ST966 (*eTable 3 in Supplement*). The overall rate of BSI was 1.3 (95% CI, 0.6-2.3) per 1,000 patient days and 0.5 (95% CI, 0.1-1.3) for infections resulting from a TE.

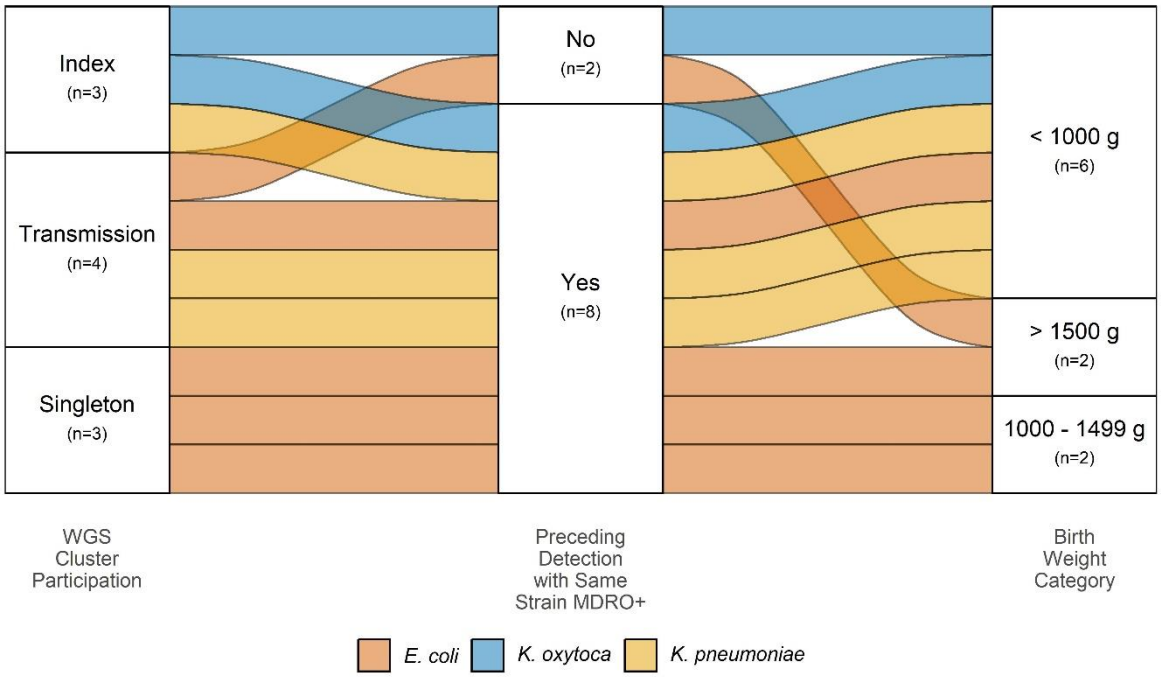

Figure 7: *Alluvial Diagram of 10 observed MDRO+ Blood Stream Infections*
Every combination of three categories resembles one patient. Seven BSI could be attributed to TC as defined by WGS (Index, Transmission), of which 4 derived from a TE. Of these, three were preceded with the detection of same strain MDRO+ in materials other than blood culture samples.

Antibiotic usage

We investigated antibiotic usage for frequently used substances and combinations. When comparing birth weight groups < 1,500 g and ≥ 1,500 g, we found that patients with a birth weight < 1,500 g received selected antibiotics or their combinations proportionally and significantly more often (*eFigure 4A in Supplement*). Among patients treated with selected antibiotics or a combination of these, we did not observe a significant difference in the total number of administrations across birth weight groups of < 1,500 g and ≥ 1,500 g (*eFigure 4B in Supplement*).

Multivariate Model Analysis

According to the study outline (*Figure 1*) we integrated multiple data points by multivariate statistical analysis. We combined patient as well as genomic information with clinical and

medRxiv preprint doi: https://doi.org/10.1101/2024.11.22.24317721; this version posted November 25, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

staffing data at high resolution in order to associate the genomic findings with clinical and environmental parameters on the NICU. The presented logistic regression model aims to predict the probability of becoming part of a TE (outcome variable) based on five predictor variables (*Figure 8A*).

Three out of five predictor variables were significantly associated with the outcome variable (*Figure 8A, 8B*): (1) Every one-unit increase of the *Number of Full-Time Nurses* decreases the odds of becoming part of a transmission cluster by 72%. (2) Every one-unit increase of *Antibiotic Prescriptions* with a moving average of 6 decreases the odds by 57%. (3) Every one-unit increase of *Vascular Catheter Usage* with a moving average of 12 increases the odds by 62%. *Kangarooing* with a moving average of 3 and the *MDRO+ daily presence count* (i.e. 'colonization pressure') with a moving average of 6 were not significantly associated with the outcome variable.

With a p-value < 0.001 for $\chi^2(5)$, the final model, with its predictors, was significantly better overall than the null model with only one constant. The general model fit can be described as weak to moderate due to the limited meaningfulness of the pseudo-R^2 of McFadden and Cragg-Uhlmann in logistic regressions. Tests using the VIF indicated no multicollinearity in the final model. A rainbow test for linearity in the final model was significant, indicating that this condition of logistic models cannot be hold. This means that not all of the necessary conditions for logistic regression were met for these data, limiting the significance of the results (*Figure 8A*).

A

Variable	Odds Ratio	Lower 95% CI	Upper 95% CI	p-value	Significance
(Intercept)	15,98	2,58	100,89	0,003	**
Vascular Catheter Usage Moving Average: 12	1,62	1,23	2,13	0,0006	***
Kangarooing Moving Average: 3	1,54	0,94	2,61	0,0982	
MDRO+ daily presence count Moving Average: 6	1,03	0,99	1,06	0,1329	
Antibiotic Prescription Moving Average: 6	0,43	0,28	0,65	0,0001	***
Number of Full-Time Nurse	0,28	0,21	0,37	0	***

Model Information	Observations	7037
	Dependent Variable	Transmission Event (Next)
	Type	Generalized Linear Model
	Family	Binomial
	Link Function	Logit
Model Fit	$\chi^2(5) = 115.50$, $p = 0.00$	
	Pseudo-R^2 (Cragg-Uhler) = 0.10	
	Pseudo-R^2 (McFadden) = 0.09	
	AIC = 1135.84, BIC = 1177.00	
Standard Errors	MLE	

B

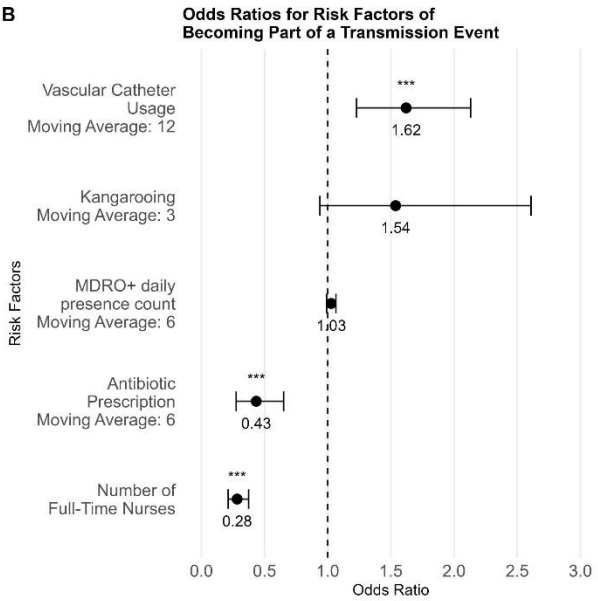

Figure 8: *Multivariate Logistic Regression Model*

A) Description of our proposed logistic regression model that aims to predict the probability of becoming part of a transmission cluster (outcome variable) based on five predictor variables with three significant associations out of five selected risk factors. The model type is a generalized linear model (GLM) with a binomial family and logit link function. The model shows statistical significance ($\chi^2(5) = 115.50$, $p < 0.001$), indicating that the predictors, as a group, are generally associated with the outcome. **B)** Forest plot of odds ratios according to the logistic regression model.

*: p-value < 0.05, **: p-value < 0.01, ***: p-value < 0.001

medRxiv preprint doi: https://doi.org/10.1101/2024.11.22.24317721; this version posted November 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

Discussion

This study identifies prospective WGS of routine screening isolates from newborns in intensive care as a powerful tool to systematically track and resolve transmission chains with MDRO+ at highest precision.

The transmission of – in part hospital adapted – microorganisms between patients is seen as a failure of basic hygiene procedures, especially in high-risk individuals such as extremely premature infants. However, since very preterm infants typically depend on life-supporting measures immediately after birth, separation from the mother is the norm, i.e. the inevitable build-up of the microbiome largely relies on input from staff skin and inanimate surfaces [36]. In this context, even the most advanced hygiene measures cannot prevent the transmission of potentially harmful enterobacteria and firmicutes like staphylococci and streptococci [1]. It is not even clear whether this would be desirable, since early exposure to the respective genus may be a prerequisite for long term tolerance [37].

On the other hand, phenotypic selection will lead to the occasional emergence of strains integrating traits facilitating inter-individual spread, in-host expansion and invasion. This will lead to the dreaded NICU outbreaks that are particularly feared in case bacteria carry resistance against first-line antibiotics [38]. It seems reasonable to suggest that real-time knowledge of the number of patients carrying a particular strain in parallel, the length of time the strain has been in a neonatal intensive care unit, and the number of infants developing clinical symptoms associated with the strain are proxies for the threat posed by the transmissions. Accordingly, accuracy and speed in the typing of bacterial strains are the theoretical basis for predicting the risk potential of an outbreak [18]. Considering the routine turnover times of bacterial typing (including material shipment, DNA-preparation and the genetic typing itself), it seems advisable to treat the occurrence of same-species-clustering with vigilance and initiate IPC-measures early. However, relying solely on cluster definition by species will inevitably lead to incorrect definition or elongation of transmission events and leading to unnecessary and unfavorable IPC-measures such as ward closures. Of note, the scientific evidence for the effectiveness of ward closures remains unclear, yet the potentially detrimental impact on treatment quality seems apparent [39,40].

Here we found AFLP to resolve TE in the majority of cases. As a possible alternative to AFLP, fourier-transformed-infrared-spectroscopy-based typing methods (FTIR) may aid in nosocomial outbreak identification (22–25). However, the comparative nature of both methods limits characterizing clusters over long periods and for longitudinal surveillance purposes. Additionally, they carry considerable mismatch rates. Accordingly, AFLP and FTIR lack resolution to resolve strain heterogeneity in all cases. Therefore we asked whether genomic cluster definition by WGS, as the gold standard for bacterial transmission analysis [41,42], would be of potential relevance in a clinical setting of a level III NICU. We found several cases, where WGS uncovered AFLP-defined TC as pseudo-clusters. As examples *E. coli* formed multiple transmission clusters of ST 73 and ST 141 as revealed by WGS. Upon in-depth-analysis, these clusters fell into distinct clusters based on their phylogenetic profile. This highlighted the discriminating capabilities of WGS as compared to methods that focus on housekeeping genes.

Our study is in line with national data regarding the importance of gram-negative bacteria, especially *E. coli*, *E. cloacae*, *K. oxytoca* and *K. pneumoniae* as screening targets and causes for nosocomial infections [5,43]. In few cases, WGS delineated the “microbial biography” from admission, to MDRO+ colonization and detection of MDRO+ in clinical samples (blood culture and tracheal sample) with the same strains.

medRxiv preprint doi: https://doi.org/10.1101/2024.11.22.24317721; this version posted November 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

It seems important to note that swap-based screenings have limitations in sensitivity for detecting transmissions since they rely on strains occurring at considerable density on the mucosa or skin-mucosa transition zone. Moreover, given the requirement of minimal handling for the most fragile very preterm infants, sampling quality will vary [44].
Any typing method relies on sufficient DNA quantity. To achieve this enrichment of picked colonies by liquid culture was required in this study. This method may select for strains that have growth advantages in culture (in the case of mixed strains of the same species).
The precision in TC definition enabled us to reliably identify significant factors in multivariate analysis affecting transmission despite the still number of events in our analysis.
Most notably, the addition of only one full time nurse on the ward decreases a patient's risk of becoming part of a TC. This seems by no means self-evident, since increasing personnel on a NICU may increase transmissions as well, e.g. due to more patient procedures being executed by two staff members, longer incubator opening times etc. Notably, staffing data included nurses declared as state-examined with high-qualification and did not account for nurses in training. Other training levels and professions should be included in further research.
In order to investigate the potential of preventing or increasing the chances of bacterial transmissions by increasing staffing levels, objective and highly granular quantification of hand hygiene measures is needed. In our observer-dependent hand hygiene compliance observations, we did not detect major changes in hand hygiene compliance over time, however, the data in its form was not suitable for the proposed multivariate analysis. It is worth mentioning that the addition of only one full time nurse on the ward is not explicitly listed as a quality indicator in the G-BA guidelines since they do not focus on IPC principles but rather focus on outlining staffing levels that are of relevance to intensive care medicine.
Presence of an indwelling catheter in patients increased the risk of becoming part of a transmission cluster. We hypothesize that vascular catheter presence may be linked to increased patient contact with staff (e.g. catheter care, drug administration).
Finally and to our surprise, we identified prior administration of antibiotics to reduce the risk of being part of a TC. Several factors may underlie this interrelation. First, the culture based screening method, preceding typing, maybe impacted by antibiotics decreasing bacterial densities at the swabbed areas since antibiotic therapy – even if administered parenterally, this can have profound effects on intestinal colonization [45,46]. Probably, this is mediated by effective antibiotic levels on mucosal surfaces [47], although this has not been explicitly shown for the gut. These considerations are in line with the fact that patients with birth weight < 1,500 g proportionally received more (commonly used) antibiotics and were detected with their first MDRO+ species later than patients with birth weight ≥ 1,500 g.
Given the evidence highlighting the potential short- and long-term effects of inappropriate or prolonged antibiotic treatment, e.g. an increased risk for bronchopulmonary dysplasia (BPD) [48], necrotizing enterocolitis (NEC) or invasive fungal infections (IFI) [49], atopy, asthma and even mortality, this finding needs exploration in independent studies. In contrast, number of NICU patients colonized with MDRO+ at any given time point ('colonization pressure'), kangarooing, the INPULS score were not significantly associated with TC.
Within our model, a necessary condition for logistic models cannot be maintained in the final model due to the significance of the rainbow test for linearity. To restore the condition of linearity, transformations such as Box-Cox [50] or Yeo-Johnson would be necessary [51]. However, back transformations of the variables are not valid and the cumbersome interpretations of transformations prevent a meaningful application in clinical practice. Further analysis in this area is needed to identify valid methods that can adequately represent complex temporal dynamics while providing interpretable risk factors that allow practical conclusions to

medRxiv preprint doi: https://doi.org/10.1101/2024.11.22.24317721; this version posted November 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

be drawn in hospitals. Alternatively, more complex methods for nonlinear data such as fractional polynomials [52–54], splines [55] or distributed lag non-linear models [56] could be used, but these also deviate significantly from the usual representation and interpretability of risk factors. Further insights could be gained through additional variability in the temporal relationships, for example by combining lags and moving averages. An analysis based on data that also contains negative screening results or a multicenter study design would certainly also provide exciting insights. Due to the monocentric approach, further investigations are required. In summary, we identify WGS, as part of the colonization screening, to be very useful in systematically resolving MDRO+ transmission chains with highest precision. Prospective integration of granular patient and staffing data with multivariate model analysis builds the foundation for rational, data-driven and actionable conceptualizations of IPC-measures in the future.

Acknowledgements

We sincerely thank Leonardo Duarte dos Santos from Group Reuter and Sabine Weber, Marion Buck, Sophie Sträb, Doris Scheibert and Eva Zimmermann from Group Jonas for their invaluable contributions and support.

Ethics and dissemination

The study was approved by the Ethics Committee, Medical Centre—University of Freiburg, on 28 August 2018 (registration number 287/18).

Registration

The study was registered in the German Clinical Trials Register (DRKS) under the registration ID DRKS00017733.

Funding

This study was funded by the programme 'Clinical Studies' of the Medical Faculty of the Albert-Ludwig University of Freiburg, Freiburg, Germany.
Philipp Henneke, Timmy Nguyen and Martin Kuntz were funded by the German Research Foundation (DFG) within TRR359 (Project ID 491676693).
Sandra Reuter was funded by the German Federal Ministry of Education and Research (BMBF; TAPIR 01KI2018).
Sandra Reuter and Tjibbe Donker were funded by the BMBF Network of University Medicine 2.0 (NUM 2.0, 01KX2121, GenSurv).

Data sharing statement

Due to confidentiality restrictions, the data supporting this study are not publicly available. However, they are available from the corresponding author upon reasonable request.

Conflict of interests

The authors declare no competing interests.

medRxiv preprint doi: https://doi.org/10.1101/2024.11.22.24317721; this version posted November 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

Abbreviations

Abbreviation	***Meaning***
NICU	Neonatal Intensive Care Unit
TE	Transmission Event
MDR	Multi drug resistant / multi drug resistance
NI	Nosocomial infection
IPC	Infection prevention control
KRINKO	Commission for Hospital Hygiene and Infection Prevention
RKI	Robert-Koch-Institute
MRGN	Multi resistant gram negative
MRSA	Methicillin resistant *Staphylococcus aureus*
MDRO+	Multidrug resistant organism plus
AFLP	Amplified fragment length polymorphism
WGS	Whole genome sequencing
TC	Transmission cluster
ESBL	Extended spectrum betalactamase
MALDI-TOF	Matrix-assisted-laser desorption-ionization time-of-flight
MIC	Minimum inhibitory concentration
EUCAST	European Committee on Antimicrobial Susceptibility Testing
MSSA	Methicillin susceptible *Staphylococcus aureus*
G-BA	Gemeinsamer Bundesausschuss
DNA	Deoxyribonucleic acid
PCR	Polymerase chain reaction
MLST	Multi locus sequence typing
SNP	Single-nucleotide polymorphism
MA	Moving averages
AIC	Akaike’s information criterion
VIF	Variance inflation factor
ST	Sequence type
BSI	Blood stream infection
FTIR	Fourier-transform infrared spectroscopy
BPD	Bronchopulmonary dysplasia
NEC	Necrotizing enterocolitis
IFI	Invasive fungal infections

medRxiv preprint doi: https://doi.org/10.1101/2024.11.22.24317721; this version posted November 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

References

[1] Van Rossum T, Haiß A, Knoll RL, Marißen J, Podlesny D, Pagel J, et al. Bifidobacterium and Lactobacillus Probiotics and Gut Dysbiosis in Preterm Infants: The PRIMAL Randomized Clinical Trial. JAMA Pediatrics 2024;178:985–95. https://doi.org/10.1001/jamapediatrics.2024.2626.

[2] Seidel J, Haller S, Eckmanns T, Harder T. Routine screening for colonization by Gram-negative bacteria in neonates at intensive care units for the prediction of sepsis: systematic review and meta-analysis. Journal of Hospital Infection 2018;99:367–80. https://doi.org/10.1016/j.jhin.2018.03.017.

[3] Goldmann DA, Leclair J, Macone A. Bacterial colonization of neonates admitted to an intensive care environment. The Journal of Pediatrics 1978;93:288–93. https://doi.org/10.1016/s0022-3476(78)80523-x.

[4] Klassert TE, Leistner R, Zubiria-Barrera C, Stock M, López M, Neubert R, et al. Bacterial colonization dynamics and antibiotic resistance gene dissemination in the hospital environment after first patient occupancy: a longitudinal metagenetic study. Microbiome 2021;9:169. https://doi.org/10.1186/s40168-021-01109-7.

[5] Litz JE, Goedicke-Fritz S, Härtel C, Wagenpfeil G, Zemlin M, Simon A. Umsetzung des mikrobiologischen Kolonisationsscreenings: Umfrage an 80 neonatologischen Intensivstationen 2019. https://doi.org/10.25646/6251.

[6] Folgori L, Tersigni C, Hsia Y, Kortsalioudaki C, Heath P, Sharland M, et al. The relationship between Gram-negative colonization and bloodstream infections in neonates: a systematic review and meta-analysis. Clinical Microbiology and Infection 2018;24:251–7. https://doi.org/10.1016/j.cmi.2017.08.008.

[7] Karagiannidou S, Triantafyllou C, Zaoutis TE, Papaevangelou V, Maniadakis N, Kourlaba G. Length of stay, cost, and mortality of healthcare-acquired bloodstream infections in children and neonates: A systematic review and meta-analysis. Infection Control & Hospital Epidemiology 2020;41:342–54. https://doi.org/10.1017/ice.2019.353.

[8] Hibbert JE, Currie A, Strunk T. Sepsis-Induced Immunosuppression in Neonates. Front Pediatr 2018;6:357. https://doi.org/10.3389/fped.2018.00357.

[9] Parra-Llorca A, Pinilla-Gonzlez A, Torrejón-Rodríguez L, Lara-Cantón I, Kuligowski J, Collado MC, et al. Effects of Sepsis on Immune Response, Microbiome and Oxidative Metabolism in Preterm Infants. Children (Basel) 2023;10:602. https://doi.org/10.3390/children10030602.

[10] Houghteling PD, Walker WA. Why is initial bacterial colonization of the intestine important to the infant's and child's health? J Pediatr Gastroenterol Nutr 2015;60:294–307. https://doi.org/10.1097/MPG.0000000000000597.

[11] Walker WA. The importance of appropriate initial bacterial colonization of the intestine in newborn, child, and adult health. Pediatr Res 2017;82:387–95. https://doi.org/10.1038/pr.2017.111.

[12] Clock SA, Ferng Y-H, Tabibi S, Alba L, Patel SJ, Jia H, et al. Colonization With Antimicrobial-Resistant Gram-Negative Bacilli at Neonatal Intensive Care Unit Discharge. Journal of the Pediatric Infectious Diseases Society 2017;6:219–26. https://doi.org/10.1093/jpids/piw014.

[13] Bubser C, Liese J, Serna-Higuita LM, Müller A, Vochem M, Arand J, et al. Impact of early antibiotic exposure on the risk of colonization with potential pathogens in very preterm infants: a retrospective cohort analysis. Antimicrob Resist Infect Control 2022;11:72. https://doi.org/10.1186/s13756-022-01110-1.

medRxiv preprint doi: https://doi.org/10.1101/2024.11.22.24317721; this version posted November 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

[14] Bhatta DR, Hosuru Subramanya S, Hamal D, Shrestha R, Gauchan E, Basnet S, et al. Bacterial contamination of neonatal intensive care units: How safe are the neonates? Antimicrobial Resistance & Infection Control 2021;10:26. https://doi.org/10.1186/s13756-021-00901-2.

[15] Decembrino L, Maini A, Decembrino N, Maggi I, Lacerenza S. Management of outbreaks in neonatal intensive care units. Early Human Development 2014;90:S54–6. https://doi.org/10.1016/S0378-3782(14)70018-0.

[16] Koch-Institut R. Epidemiologisches Bulletin 42/2013. Oktober 2013:16. https://doi.org/DOI: 10.25646/463.

[17] Gastmeier P, Loui A, Stamm-Balderjahn S, Hansen S, Zuschneid I, Sohr D, et al. Outbreaks in neonatal intensive care units—They are not like others. American Journal of Infection Control 2007;35:172–6. https://doi.org/10.1016/j.ajic.2006.07.007.

[18] Do SES, Logan LK, Green SJ, Moore NM, Hayden MK. Whole-genome sequencing for neonatal intensive care unit outbreak investigations: Insights and lessons learned n.d. https://doi.org/doi:10.1017/ash.2021.161.

[19] Koch-Institut R. Epidemiologisches Bulletin 2/2012 2012:6.

[20] Hygienemaßnahmen bei Infektionen oder Besiedlung mit multiresistenten gramnegativen Stäbchen. Bundesgesundheitsbl 2012;55:1311–54. https://doi.org/10.1007/s00103-012-1549-5.

[21] Götting T, Reuter S, Jonas D, Hentschel R, Henneke P, Klotz D, et al. Protocol for a prospective cohort study: Prevention of Transmissions by Effective Colonisation Tracking in Neonates (PROTECT-Neo). BMJ Open 2020;10:e034068. https://doi.org/10.1136/bmjopen-2019-034068.

[22] Epidemiologisches Bulletin 2/2012 2012.

[23] Gemeinsamer Bundesausschuss. Beschluss des Gemeinsamen Bundesausschusses über eine Änderung der Qualitätssicherungsrichtlinie Früh- und Reifgeborene/QFR-RL: Erstfassung der Datenfelder der Strukturabfrage als Anlage 6. Bundesanzeiger. 2019 May 28. Available from: https://www.g-ba.de/downloads/39-261-3581/2018-11-22_QFR-RL_Erstfassung-Anlage-6_BAnz.pdf

[24] Jonas D, Spitzmüller B, Daschner FD, Verhoef J, Brisse S. Discrimination of Klebsiella pneumoniae and Klebsiella oxytoca phylogenetic groups and other Klebsiella species by use of amplified fragment length polymorphism. Research in Microbiology 2004;155:17–23. https://doi.org/10.1016/j.resmic.2003.09.011.

[25] Jonas D, Speck M, Daschner FD, Grundmann H. Rapid PCR-Based Identification of Methicillin-Resistant *Staphylococcus aureus* from Screening Swabs. J Clin Microbiol 2002;40:1821–3. https://doi.org/10.1128/JCM.40.5.1821-1823.2002.

[26] Ponstingl H, Ning Z. SMALT - A new mapper for DNA sequencing reads. Wellcome Trust Sanger Institute; 2010. Available from: ftp://ftp.sanger.ac.uk/pub/hp3/smalt.tgz. DOI: not available.

[27] Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. GigaScience 2021;10:giab008. https://doi.org/10.1093/gigascience/giab008.

[28] McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010;20:1297–303. https://doi.org/10.1101/gr.107524.110.

[29] Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. Journal of Computational Biology 2012;19:455–77. https://doi.org/10.1089/cmb.2012.0021.

[30] Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol 2014;15:R46. https://doi.org/10.1186/gb-2014-15-3-r46.

medRxiv preprint doi: https://doi.org/10.1101/2024.11.22.24317721; this version posted November 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

[31] Seemann T. MLST: Multi-locus sequence typing. [software on the internet]. GitHub; 2017 [cited 2024 Nov 18]. Available from: https://github.com/tseemann/mlst.

[32] Barnett A, Graves N. Competing risks models and time-dependent covariates. Crit Care 2008;12:134. https://doi.org/10.1186/cc6840.

[33] Statistical methods in cancer research. 2: The design and analysis of cohort studies. Reprinted. Lyon: 1991.

[34] Utts JM. The rainbow test for lack of fit in regression. Communications in Statistics - Theory and Methods 1982;11:2801–15. https://doi.org/10.1080/03610928208828423.

[35] Krämer W, Sonnberger H. The Linear Regression Model Under Test. Heidelberg: Physica-Verlag HD; 1986. https://doi.org/10.1007/978-3-642-95876-2.

[36] Jara J, Alba C, Campo RD, Fernández L, Sáenz de Pipaón M, Rodríguez JM, et al. Linking preterm infant gut microbiota to nasograstric enteral feeding tubes: exploring potential interactions and microbial strain transmission. Frontiers in Pediatrics n.d.:Article ID 1397398. https://doi.org/DOI 10.3389/fped.2024.1397398.

[37] Freudenhammer M, Karampatsas K, Le Doare K, Henneke P, Heath PT. Invasive Group B Streptococcus Disease With Recurrence and in Multiples: Towards a Better Understanding of GBS Late-Onset Sepsis. Frontiers in Immunology 2021:Article ID 617925. https://doi.org/10.3389/fimmu.2021.617925.

[38] Verani JR. Child deaths caused by Klebsiella pneumoniae in sub-Saharan Africa and south Asia: a secondary analysis of Child Health and Mortality Prevention Surveillance (CHAMPS) data. The Lancet Microbe 2024;e131–e141 (February 2024):e131–41. https://doi.org/10.1016/S2666-5247(23)00290-2.

[39] Wong H, Eso K, Ip A, Jones J, Kwon Y, Powelson S, et al. Use of ward closure to control outbreaks among hospitalized patients in acute care settings: a systematic review. Syst Rev 2015;4:152. https://doi.org/10.1186/s13643-015-0131-2.

[40] Redondo-Bravo L, Gutiérrez-González E, San Juan-Sanz I, Fernández-Jiménez I, Ruiz-Carrascoso G, Gallego-Lombardo S, et al. Serratia marcescens outbreak in a neonatology unit of a Spanish tertiary hospital: Risk factors and control measures. American Journal of Infection Control 2019;47:271–9. https://doi.org/10.1016/j.ajic.2018.08.026.

[41] Janes VA, Notermans DW, Spijkerman IJB, Visser CE, Jakobs ME, van Houdt R, et al. Amplified fragment length polymorphism and whole genome sequencing: a comparison of methods in the investigation of a nosocomial outbreak with vancomycin resistant enterococci. Antimicrobial Resistance & Infection Control 2019;8:153. https://doi.org/10.1186/s13756-019-0604-5.

[42] Quainoo S, Coolen JPM, van Hijum SAFT, Huynen MA, Melchers WJG, van Schaik W, et al. Whole-Genome Sequencing of Bacterial Pathogens: the Future of Nosocomial Outbreak Analysis. Clin Microbiol Rev 2017;30:1015–63. https://doi.org/10.1128/CMR.00016-17.

[43] Artelt T, Kaase M, Küster H, Paul T, Eiffert H, Groß U, et al. INSIST - Impact des Neonatologie-Screenings auf Infektionsprädiktion / Senkung von Transmission. Gemeinsamer Bundesausschuss (G-BA) 2017.

[44] Warnke P, Johanna Pohl FP, Kundt G, Podbielski A. Screening for Gram-negative bacteria: Impact of preanalytical parameters. Sci Rep 2016;6:30427. https://doi.org/10.1038/srep30427.

[45] Klassert TE, Zubiria-Barrera C, Kankel S, Stock M, Neubert R, Lorenzo-Diaz F, et al. Early Bacterial Colonization and Antibiotic Resistance Gene Acquisition in Newborns. Front Cell Infect Microbiol 2020;10:332. https://doi.org/10.3389/fcimb.2020.00332.

medRxiv preprint doi: https://doi.org/10.1101/2024.11.22.24317721; this version posted November 25, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY-NC 4.0 International license .

[46] Heimdahl A, Kager L, Malmborg AS, Nord CE. Impact of different betalactam antibiotics on the normal human flora, and colonization of the oral cavity, throat and colon. Infection 1982;10:120–4. https://doi.org/10.1007/BF01816741.

[47] Havard CW, Bax RP, Samanta TC, Pearson RM, Brumfitt W, Hamilton-Miller JM, et al. Sputum and blood concentrations of cefuroxime in lower respiratory tract infection. Thorax 1980;35:379–83. https://doi.org/10.1136/thx.35.5.379.

[48] Shi W, Chen Z, Shi L, Jiang S, Zhou J, Gu X, et al. Early Antibiotic Exposure and Bronchopulmonary Dysplasia in Very Preterm Infants at Low Risk of Early-Onset Sepsis. JAMA Network Open 2024;7:e2418831. https://doi.org/10.1001/jamanetworkopen.2024.18831.

[49] Esaiassen E, Fjalstad JW, Juvet LK, Van Den Anker JN, Klingenberg C. Antibiotic exposure in neonates and early adverse outcomes: a systematic review and meta-analysis. Journal of Antimicrobial Chemotherapy 2017;72:1858–70. https://doi.org/10.1093/jac/dkx088.

[50] Box GEP. Analysis of Transformations | Journal of the Royal Statistical Society Series B: Statistical Methodology | Oxford Academic n.d. https://academic.oup.com/jrsssb/article-abstract/26/2/211/7028064?redirectedFrom=fulltext (accessed October 24, 2024).

[51] Yeo I-K. A new family of power transformations to improve normality or symmetry. Biometrika 2000;87:954–9. https://doi.org/10.1093/biomet/87.4.954.

[52] Faes C, Aerts M, Geys H, Molenberghs G. Model Averaging Using Fractional Polynomials to Estimate a Safe Level of Exposure. Risk Analysis 2007;27:111–23. https://doi.org/10.1111/j.1539-6924.2006.00863.x.

[53] Royston P, Altman DG. Regression Using Fractional Polynomials of Continuous Covariates: Parsimonious Parametric Modelling. Applied Statistics 1994;43:429. https://doi.org/10.2307/2986270.

[54] Sauerbrei W, Royston P. Building Multivariable Prognostic and Diagnostic Models: Transformation of the Predictors by Using Fractional Polynomials. Journal of the Royal Statistical Society Series A: Statistics in Society 1999;162:71–94. https://doi.org/10.1111/1467-985X.00122.

[55] Wouthuyzen-Bakker M, Benito N, Soriano A. The Effect of Preoperative Antimicrobial Prophylaxis on Intraoperative Culture Results in Patients with a Suspected or Confirmed Prosthetic Joint Infection: a Systematic Review. Journal of Clinical Microbiology 2017;55:JCM.00640-17. https://doi.org/10.1128/JCM.00640-17.

[56] Gasparrini A, Armstrong B, Kenward MG. Distributed lag non-linear models. Statistics in Medicine 2010;29:2224–34. https://doi.org/10.1002/sim.3940.

Supplement

Supplemental_Tables_Figures.pdf
PRJEB81699_accessions_SAM.xlsx