Zala PZ, Ouedraogo S, Schumacher S, Ouedraogo P, Rosa-Mangeret F, Pfister RE. A minimal neonatal dataset (mND) for low- and middle-income countries as a tool to record, analyse, prevent and follow-up neonatal morbidity and mortality. *Journal of Global Health Reports*. 2023;7:e2023015. doi:10.29392/001c.75151

Research Article

# A minimal neonatal dataset (mND) for low- and middle-income countries as a tool to record, analyse, prevent and follow-up neonatal morbidity and mortality

Persis Zokara Zala[1,2], Solange Ouedraogo[3], Sofia Schumacher[1], Paul Ouedraogo[4], Flavia Rosa-Mangeret[1,5], Riccardo E. Pfister[1]

[1] Neonatology and Paediatric Intensive Care Unit, Geneva University Hospitals and Geneva University, Geneva, Switzerland , [2] Centre Médico Chirurgical Pédiatrique Persis, [3] Neonatal Unit, Bogodogo University Hospital Centre, Ouagadougou, Burkina Faso, [4] Saint Camille Hospital, Ouagadougou, Burkina Faso, [5] Global Health Institute University of Geneva, Geneva, Switzerland

Keywords: neonate, mortality, dataset, low- and middle-income

https://doi.org/10.29392/001c.75151

Journal of Global Health Reports

Vol. 7, 2023

## Background

Neonatal mortality accounts for the most significant and today increasing proportion of under-5 mortality, especially in sub-Saharan Africa. The neonatal population is a sharp target for intervention for these 2.5 million annual deaths. The limited availability of quality data on morbidities leading up to this mortality hampers the development and follow-up of effective interventions. For leverage, undoubtedly more detailed and standardized data adapted to low and middle-income countries (LMICs) is urgently needed.

## Methods

Drawing on existing databases such as the Swiss Neonatal Network and Vermont Oxford Network, 267 clinical, administrative, and structural variables of neonatal health and healthcare services were selected and submitted for ranking to 42 experts through two Delphi rounds. An empirically limited number of variables with the highest ranking for availability and relevance in low and middle-income countries were field-tested in three centres in Burkina Faso during one year for improvement and practicality.

## Results

We report the database development process according to the Standards for Quality Improvement Reporting Excellence (SQUIRE 2.0) recommendations. The final dataset is composed of 73 clinical and 6 administrative patient variables, and 21 structural healthcare center variables. Two-thirds of clinical variables maintain matching definitions with high-income countries.

## Conclusions

The developed minimal neonatal dataset is standardized and field-tested for relevance and availability in LMICs allowing south-south and some south-north cross-comparison.

With 2.5 million deaths[1,2] no age group other than the neonatal period has higher mortality whilst remaining largely unrecognized.[3] Despite a consistent achievement in under-5 mortality, the Millennium Development Goal reached least for newborns.[4] Neonatal mortality (during the first 28 days) still represents the most significant fraction of the under-five mortality and its contribution to infant mortality is even on the rise: 40% in 1990, 45% in 2015[5] and 47% in 2020.[2] Regional differences, unrealistic numbers acquired under governmental pressure and artificially decreased neonatal mortality by late declaration of surviving newborns only[6,7] put doubt on reported figures and likely underestimate official mortality figures. Although 'late abortion' rates should mirror such practice, these notoriously remain un- or underreported.[6,8,9]

Most neonatal deaths occur in low and middle-income countries (LMIC), close to 80% in two single areas: sub-Saharan Africa (43%) and Central & South Asia (36%).[2] In 2019, neonatal mortality was reported 7 to 13 times higher in low compared to high-income countries.[1] Close to 80% of all neonatal deaths are linked to three leading conditions, prematurity and low birth weight, perinatal complications and asphyxia, as well as sepsis and infection.[5,10-12] However, associations are tightly interlinked and not nec-

essarily causes. Prematurity, for instance, the most significant contributor to neonatal death, is not per se a treatable cause of death.[13] Quality data on specific underlying or contributing morbidities, and precise circumstances of death are urgently needed, a prerogative for any targeted intervention. Neonatal mortality (the first 28 days) is a particularly narrow-defined temporal intervention target, and three quarters of deaths even concentrate during the first week of life.[4,10] Mortality, however, is only the tip of a much greater underlying morbidity[11,14-17] on which limited quality data is available in LMIC.[18-22] Facing neonatal morbidity will also influence morbidity and mortality after the neonatal period, and long-term handicap and adult disease resulting from neonatal disease.

Whether due to lack of resources,[23] political or academic interest, current neonatal LMIC data are frequently unavailable, incomplete, or non-standardized. Where data are available, they are generally collected retrospectively, regionally (clusters) or in areas of practicality, and extrapolated on a broader scale. Such data only poorly represent the intra- and inter-regional and national variabilities to geographically target health strategies.[9,24] As a result, neonates are systematically unrepresented, even more so during epidemics and war. Despite top ranking in mortality, neonates usually get the least attention and the slimmest support of Maternal-Newborn and Child Health programs. Within these programs, reports rarely focus on anything other than crude neonatal mortality. Lack of quality data may be one of the reasons why financial supports largely skip this most vulnerable age group where the burden remains unmeasured and thus hard to defend. To promote, target, leverage, and follow-up interventions, standardized data representative of the specific health challenges of the neonatal populations in LMIC are urgently required.[25] Such data must be locally available, context-adapted, and assembling feasible. Standardized data should in priority target south-south quality comparison and improvement. Our report is structured according to the Standards for Quality Improvement Reporting Excellence (SQUIRE 2.0).

To design and development a quality improvement dataset for prospective collection of standardized data on neonatal health and healthcare structures representative of LMIC.

## METHODS

### CONTEXT

Regionalization in high-competence centers is essential to improve population-based neonatal care,[26] and requires excellence in terms of competence, education, and reputation. Today however, to treat newborns, many reference centers in LMIC do hardly have more staff, equipment, and infrastructure than lower-level centers, despite their obligation to accept the sickest. Delays in decision making in the community before referral, inappropriate transport conditions and delays at admission to the reference centers further worsen outcomes. The resulting very high mortality undermines the population's trust in these reference centers and with it, often also government support.

Focusing on high-level healthcare facilities can generate high-quality data that supports public understanding of the risks leading to this concentrated mortality and morbidity and advocates government support in a virtuous quality improvement circle.

### INTERVENTION: DEVELOPMENT OF VARIABLE SETS

In 2018, 267 newborn variables collected through a literature search and from two established high-income neonatal databases, the Swiss Neonatal Network (SNN) data collection and Vermont Oxford Network (VON), were submitted to a two-step Delphi selection process. Experts were all except one, physicians or nurses with broad experience in LMIC neonatal care, 44% actively practicing in such countries. One expert was a neonatal database manager. Of the invited 42 experts, 28 answered to the first and 25 to the second Delphi round.

Experts graded variables according to feasibility, reliability, and representability. We retained variables with a mean greater than or equal to +1 SD of its respective subgroup (the 14 subgroups and rating scale for variable selection are presented in Appendices S1 and S2 in the **Online Supplementary Document**), and discarded if less than -2 SD. The in-between variables were re-submitted for the second round, together with experts' suggestions of new items. The empirical target was a dataset of 60 clinical variables and 20 structural health care facility variables.

Variable definitions were maintained equivalent to the original databases (SNN, VON) whenever possible.

### STUDY OF THE INTERVENTION: FIELD TESTING

We identified three demographically diverse centers in Burkina Faso for field-testing and obtained local ethics committee clearance in August 2020. All centers cared for in- and out-born neonatal patients at high level (table 1). Two centers are in the capital city of Ouagadougou, one a university unit and the other a confessional maternity hospital with the highest premature admission rate in the country. The third is a maternity hospital in northern Ouahigouya.

We instructed part-time local data managers (1 midwife; 2 physicians) and provided continuous remote support with two-weekly checks. Training of local managers included one-to-one guided recording of 10 cases and regular focused discussions of problematic item definitions.

Data sets were entered 'continuously' into a spreadsheet according to the predefined written data definitions and structure. During the first 3 months, feed-back was provided 'on demand' often several times a week. Recurrent definition issues were discussed with participating centers, and definitions improved if necessary. In addition, they were invited to propose new variables that they felt were necessary. Data completeness/availability was also reviewed with local data managers and heads of units during three on-site visits to determine appropriateness of identified variables and common data handling strategies.

**Table 1. Unit characteristics for database testing***

| Neonatal care center | Center 1 | Center 2 | Center 3 |
|---|---|---|---|
| Births per year (2020) | 1230 | 6234 | 3523 |
| National level | highest | highest | highest |
| Active neonatal beds | 15 | 25 | 30 |
| Running water | cold/warm | cold | cold/warm |
| Wall oxygen | yes | yes | yes |
| Electric suction devices | 1 | 2 | 1 |
| Incubators | 1 | 0 | 1 |
| Radiant warmers | 4 | 1 | 1 |
| Heated beds | 0 | 0 | 6 |
| Phototherapy units | 3 | 1 | 6 |
| CPAP devices | 0 | 0 | 5 |
| Medical staff | 2 | 10 | 2 |
| Nursing staff | 14 | 35 | 39 |

*Details are presented in Appendix S3 in the Online Supplementary Document

## MEASURES: VARIABLE RELIABILITY ASSESSMENT

At 3 months of data collection, a quality report of selected variables (table 2) was shared with participating centers. One of the Geneva investigators, a Burkina Faso national physician, visited all participating centers in January 2021 for consolidation of the definitions. Issues with data plausibility, structure, and some definitions were individually discussed based on concrete examples to confirm/optimize data completeness and quality. Systematic definition issues, detected from the mismatch between expected and reported results, were also clarified with clinicians on site.

Thereafter, improved definitions and local expertise allowed data submission in bunches every two weeks for an interim plausibility and quality check, and data back-up in Geneva. Co-operative epidemiological work rounds (virtual and on-site) were held on a 3-monthly schedule to smooth identified or potential hurdles.

After 12 months of data entry, we requested data managers to report time requirements for data gathering, recording and transmission, including the time for regular briefings. In addition, all questions/answers that circulated between Burkina Faso and Geneva essentially through smartphone messaging, were accounted for, estimating 5 minutes per message.

## ANALYSIS

The analysis focused on quality improvement of the database and did not evaluate patient or population outcome. Database content and quality were first analyzed quantitatively and qualitatively through a two-round Delphi process, followed by qualitative focus group evaluations, fed by individual problem-solving sessions.

**Table 2. Quality parameters of the preliminary unit summary**

| Domains | Main variables |
|---|---|
| Period covered | admission date and DOB |
| Population demographics | gender, gestational age, weight |
| Reasons for admission | RDS, prematurity, infection, asphyxia and "other" |
| Delivery & adaptation | mode of delivery, Apgar score, resuscitation requirements |
| Temperature | admission temperature, recurrent hypo-/hyperthermia |
| Neurology | asphyxia, convulsions |
| Infections | EOS, LOS |
| Supports | antibiotic exposure, respiratory support, catheter, phototherapy |
| Cause of death | respiratory, neurological, digestive, infectious, hypothermia related, bleeding and "other" |
| Data completeness | percentage of unknown data: 5' Apgar score, gestational age, BW |

## ETHICS CONSIDERATIONS

The protocol was submitted to the Ethics Committee of Burkina Faso, specifying that anonymity of the patients would be preserved by means of an individual code and the participating centers would remain anonymous. We received written clearance without need to request patient consent, on August 12, 2020 (No. 2020-8-173). This work does not report patient data in any way, but merely used it

to identify and improve variable definitions. Any personal patient data remained within the hospital files. Nominative data was available to participating local investigators for their patients only, and to the two primary investigators (PZ, RP) in Geneva during the test period. Participating centers gave their written agreement to participate and were anonymized for comparison of center data against both other centers. All authors and heads of participating centers gave consent for publication.

## RESULTS

### VARIABLE SELECTION (PHASE 1)

The two-round Delphi analysis identified 71 patient variables (7 administrative and 64 clinical) with roughly one-third that required context-adapted re-definition from the original VON or SNN definitions, 2/7 (28%) for administrative and 27/64 (42%) for clinical variables.

For structural unit data, the Delphi ranking identified 20 variables, 9/20 (45%) identical to the VON and SNN definitions and 11/20 (55%) defined to account for typical LMIC unit structure characteristics.

Appendices S1-S2 in the **Online Supplementary Document** report all variable subgroups and rating criteria that identified the 71 clinical and administrative, and 20 structural variables submitted to the subsequent two improvement steps.

### PRIMARY VARIABLE IMPROVEMENT (PHASE 2)

On-site focus groups concluded after 3 months of data collection on addition and clarification of some clinical and administrative variables, yielding a total of 78 variables. One variable was entirely recoded, 14 variable definitions were adapted, 5 variables split into two, 4 variables merged into two, 5 variables deleted as generally unavailable, and 9 new variables added.

For the structural unit data, the definition of one variable was adapted, and one new variable was added based on a convergent needs-assessment of the three local co-researchers. As a result, the structural unit data are now composed of 21 variables. For details see Appendix S4 in the **Online Supplementary Document**.

### SECONDARY VARIABLE IMPROVEMENT (PHASE 3)

After 9 months of data entry, it became clear that the sepsis definition needed specifications with perceived severity of the infection being an essential element. Therefore, antibiotic therapy duration and "circulatory compromise" were added. We also improved the variable definition of "pregnancy dating" and two unit structure variables. For details see Appendix S4 in the **Online Supplementary Document**.

Finally, our ***mND*** for LMIC is composed of 79 patient data, 73 clinical and 6 administrative. The additional 21 unit structure variables need update after unit modifications only (table 3).

### ESTIMATED WORKTIME FOR DATA MANAGERS

After 12 months of use, overall 2039 newborns were included. The average data entry time for 100 newborn files was approximatively 40 hours (table 4).

## DISCUSSION

### SUMMARY

Our standardized neonatal database (***mND***) is one of the first neonatal databases designed specifically for LMIC. A pilot project seems to be ongoing in Ethiopia, though little information and no published reference is available so far. Our ***mND*** neonatal database is inspired by neonatal databases used in high-income countries, such as the SNN and VON. After a selection and improvement process, it is composed of 79 clinical and 21 structural variables. Although we report the English version of the ***mND***, the original version is in French. French-speaking African countries generally receive less international attention for linguistic reasons, limiting their access to the broader Anglo-Saxon support.

The strength of the developed ***mND*** is its context-adapted set of variables that remain internationally comparable for 2/3 of them with identical definitions. They were chosen for their availability and completed by context-representative and reliable variables for neonatal health in LMIC context. The dataset has been thoroughly field-tested with incremental quality improvement phases within three geographic areas of Burkina Faso with differing neonatal populations.

### INTERPRETATION

The ***mND*** provides a standardized frame for south-south comparison between neonatal healthcare facilities in LMIC, including French speaking units. It presently provides a tool for cross-sectional quality control to identify weaknesses and strengths in neonatal facility care. As neonatal units vary considerably and currently do not use international level of care definitions, the structural unit characteristics, developed explicitly for LMIC, support inter-unit comparability, and allow disease-targeted appreciation of structural deficiencies. Over time, longitudinal data will provide healthcare facilities with follow-up, particularly after corrective interventions.

When data entry is cumbersome and time consuming, data quality suffers. 'Real-life' studies have been done to assess the time required to complete a form and longitudinal studies are planned for exactly this purpose. The present spreadsheet entry required close to 40 hours of work per 100 cases. We opted within the testing phase to hire dedicated data managers on site and evaluated their required time investment, allowing comparison with future data entry developments. Indeed, we are currently working on facilitated data entry routines to replace the spreadsheet data entry with a branching algorithm that should provide more efficient and faster, whilst secure, electronic data entry.

**Table 3. mND administrative, clinical and structural variable characteristics and definitions**

| N° | Variable | Definition | Format | Type | Min | Max | Missing value |
|---|---|---|---|---|---|---|---|
| | Administrative patient variables | | | | | | |
| 1a | Patient's family name | | | free text | | | |
| 1b | Patient's first name | | | free text | | | |
| 2 | Patient code* | automatically generated (for the moment manually) | | free text (number) | | | 999999 |
| 3 | Gender* | 0=undetermined; 1=boy; 2=girl | | ordinal | 0 | 2 | 9 |
| 4a | Date of birth* | date - calendar | DD.MM.YYYY | date | Entry date - 1 month | < to the date of entry | 99.99.9999 |
| 4b | Time of birth* | time | HH:MM | time | | | 99:99 |
| 5a | Place of birth* | 0=outborn, 1=inborn=born where hospitalized | | boolean | 0 | 1 | 9 |
| 5b | Place of birth (level)* | 0=home; 1=level 1=no doctor, but midwife or birth attendant; 2=level 2=with doctors 3=highest level with paediatricians | | ordinal | 0 | 3 | 9 |
| 6 | Date of admission* | date - calendar | DD.MM.YYYY | date | Entry date - 1 month | < to the date of entry | 99.99.9999 |
| | Clinical patient variables | | | | | | |
| 7a | Gestational age* | Best estimate of GA in weeks | | integer | 20 | 44 | 99 |
| 7b | Gestational age* | GA in days over weeks | | integer | 0 | 6 | 9 |
| 7c | Pregnancy dating | unknown=0, vague=date of last menstrual period known or dating by Ballard score=1, certain=US before 20 SA=2 | | integer | 0 | 2 | 9 |
| 8 | Mother's year of birth | To be completed in years according to the format | YYYY | date | Entry date -50 years | < to the date of entry | 9999 |
| 9 | Mother's education* | illiterate=0, basic or elementary school=1, non-university college=2, university=3 | | ordinal | 0 | 3 | 9 |
| 10 | Housing area | rural=village=0, urban=city=1 | | boolean | 0 | 1 | 9 |
| 11 | Type of dwelling | sum of the facilities in the house: running water, electricity, toilet, separate kitchen (if 0, 1, 2 or 3 items=0, if all=1) | | boolean | 0 | 1 | 9 |
| 12a | Gravidity* | number of pregnancies without the current | | integer | 0 | 20 | 99 |
| 12b | Parity* | number of live or deceased births without the current | | integer | 0 | 20 | 99 |

| N° | Variable | Definition | Format | Type | Min | Max | Missing value |
|---|---|---|---|---|---|---|---|
| 12c | Twinning* | no=0, yes=1 twins or more | | boolean | 0 | 1 | 9 |
| 12d | Birth sequence* | if multiple foetuses, birth order | numbers and letters IIA, IIB, IIIA, IIIB, IIIC | integer/ letter | IIA | IIIC | 9 |
| 13 | Number of antenatal controls* | number of check-ups or consultations during this pregnancy, not only ultrasound | | integer | 0 | 20 | 99 |
| 14 | Maternal fever | ≥ 38.5°C outside immediate labour: no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 15 | Maternal HIV | Infection during pregnancy: no=0; yes treated (HAART)=1; yes untreated=2 | | ordinal | 0 | 2 | 9 |
| 16 | Maternal syphilis | Infection during pregnancy: no=0; yes treated=1; yes untreated=2 | | ordinal | 0 | 2 | 9 |
| 17 | Maternal hepatitis B | Maternal hepatitis B active during pregnancy: no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 18 | Maternal malaria | Maternal malaria active during pregnancy: no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 19 | Duration of rupture of membranes | hours (if >4 days=99) | | number | 0 | 99 | 999 |
| 20 | Quality of the amniotic fluid | clear=0, stained=1, meconium=2, purulent and/or malodorous=3 | | ordinal | 0 | 3 | 9 |
| 21 | Maternal antibiotics during delivery* | no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 22 | Maternal hypertension/pre-eclampsia/eclampsia* | One of the 3 situations: Maternal hypertension/pre-eclampsia/ eclampsia: no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 23 | Maternal diabetes | Maternal diabetes none=0, gestational=1, other=2 (type 1/2...) | | ordinal | 0 | 2 | 9 |
| 24 | Antenatal steroid treatment* | none=0, incomplete=1, complete >24h before birth=2 | | ordinal | 0 | 2 | 9 |
| 25 | Birth weight* | weight in grams | | integer | 500 | 6500 | 9999 |
| 26 | Admission weight* | weight in grams | | integer | 500 | 6500 | 9999 |
| 27 | Admission length* | measured in cm (if less than 1 week of life, take birth size) | | decimal | 30.0 | 60.0 | 99.9 |
| 28 | Admission head circumference* | measured in cm (if less than 1 week of life, take birth size) | | decimal | 30.0 | 60.0 | 99.9 |
| 29 | Reason for admission* | prematurity=1, infection=2, asphyxia=3, RDS=4, other=5 (put 5 and write the cause beside) | | ordinal | 1 | 5 | 9 |
| 30 | General condition on admission | bad=0 (e.g. hypotonia, apathy, other...), good=1 | | boolean | 0 | 1 | 9 |

| N° | Variable | Definition | Format | Type | Min | Max | Missing value |
|---|---|---|---|---|---|---|---|
| 31 | Cyanosis on admission | central cyanosis on admission: no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 32 | Central temperature on admission* | first central T° (axillary/rectal) measured at admission in °C | | decimal | 30.0 | 41.5 | 99.9 |
| 33 | Cold on touch at admission | trunk cold to the touch: no=0, yes=1, to be done if admission temperature not measured | | boolean | 0 | 1 | 9 |
| 34 | Foetal risk factors | no=0, yes=1 (foetal distress and/or umbilical cord prolapse and/or antepartum haemorrhage and/or other) | | boolean | 0 | 1 | 9 |
| 35 | Mode of delivery* | 0=spontaneous vaginal delivery<br>1=instrumented vaginal delivery<br>2=caesarean section before the onset of labor<br>3=caesarean section after the start of labour | | ordinal | 0 | 3 | 9 |
| 36 | Foetal trauma at birth | obstetrical trauma (specify type): no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 37a | Apgar score at 1 min* | Apgar score at 1 min | | integer | 0 | 10 | 99 |
| 37b | Apgar score at 5 min* | Apgar score at 5 min | | integer | 0 | 10 | 99 |
| 37c | Apgar score at 10 min* | Apgar score at 10 min | | integer | 0 | 10 | 99 |
| 38 | Bag & mask ventilation at birth* | during first hour of life: no=0, yes <5 min=1, yes >5 min=2 | | ordinal | 0 | 2 | 9 |
| 39 | Administration of oxygen at birth* | during first hour of life: no=0, yes <5 min=1, yes >5 min=2 | | ordinal | 0 | 2 | 9 |
| 40 | Thoracic compressions at birth* | During first hour of life: no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 41 | Vascular access at birth | access during resuscitation (in the first hour of life): none=0, peripheral venous access=1, central venous access=2 | | ordinal | 0 | 2 | 9 |
| 42 | Hypothermia <36.5 °C during hospital stay | T°≥36.5°C=0, T°<36.5°C=1 | | ordinal | 0 | 1 | 9 |
| 43 | Hyperthermia >38.5 °C during hospital stay | T°≤38.5°C=0, T°>38.5°C=1 | | ordinal | 0 | 1 | 9 |
| 44 | Signs of respiratory distress* | respiratory rate>60/min, recessions, expiratory grunting, nasal flaring, cyanosis in room air: <2 signs=0, >2 signs=1 | | ordinal | 0 | 1 | 9 |
| 45 | Circulatory compromise | skin re-capillarisation time >3 seconds and/or significant skin pallor and/or very marked mottling, hypotension no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 46a | Asphyxia | (only if Apgar score<5 at 5 min of life) no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 46b | Sarnat score (the worst if asphyxia) | normal=0, moderate encephalopathy=1, severe encephalopathy=2 | | ordinal | 0 | 2 | 9 |

| N° | Variable | Definition | Format | Type | Min | Max | Missing value |
|---|---|---|---|---|---|---|---|
| 47 | Seizures | none=0, yes=1 | | boolean | 0 | 1 | 9 |
| 48 | Congenital HIV | to be asked only if mother is HIV+: no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 49 | Congenital malaria | no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 50 | Omphalitis | no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 51a | Neonatal infection* | no=0; yes generalized clinical (hypotonia, apathy, hypoperfusion signs) =1, yes clinical and culture proven=2, yes clinical and localized=3 | | ordinal | 0 | 3 | 9 |
| 51b | Date of diagnosis of neonatal infection requiring antibiotics* | date - calendar | DD.MM.YYYY | date | date of admission | date of discharge | 99.99.9999 |
| 51c | Central catheter in place when neonatal infection diagnosed | no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 52 | Haemoglobin | Haemoglobin determined: yes=1, no=0 | | boolean | 0 | 1 | 9 |
| 53 | Hyperbilirubinemia requiring phototherapy | Clinical or biological no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 54 | Hypoglycaemia requiring intravascular (IV) treatment | no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 55 | Congenital malformation* | absent=0 or minor, present=1 requiring treatment within 1 month | | boolean | 0 | 1 | 9 |
| 56 | Minimum weight during hospitalization | in grams | | integer | 500 | 6500 | 9999 |
| 57 | Use of an external heating device | Use of an external heating device such as heat lamp, radiant table, heating bed, incubator: none=0, yes=1 | | boolean | 0 | 1 | 9 |
| 58 | Use of skin-to-skin | no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 59 | Maximum respiratory support except mechanical ventilation* | no=0, oxygen=1, caffeine=2, oxygen + caffeine=5 | | ordinal | 0 | 5 | 9 |
| 60 | Mechanical ventilation* | no=0, yes HiFlow=1, yes CPAP=2, yes invasive mechanical ventilation=3 | | ordinal | 0 | 3 | 9 |
| 61 | Vascular access during the stay | none=0, peripheral=1, central=2 | | ordinal | 0 | 2 | 9 |
| 62 | Blood transfusion* | no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 63a | Received antibiotics* | no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 63b | Start date of antibiotics | date - calendar | DD.MM.YYYY | date | date of admission | date of discharge | 99.99.9999 |

| N° | Variable | Definition | Format | Type | Min | Max | Missing value |
|---|---|---|---|---|---|---|---|
| 63c | End date of antibiotics | date - calendar | DD.MM.YYYY | date | date of admission | date of discharge +1 month | 99.99.9999 |
| 64 | Administration of anti-malaria drugs | no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 65 | Type of milk given during the stay* | exclusive breast milk=0, mixed=1, artificial=2 none=3 | | ordinal | 0 | 3 | 9 |
| 66 | Feed stop >12 hours | no=0, yes only once=1, yes several times=2 | | ordinal | 0 | 2 | 9 |
| 67 | Administration of vitamin K (im/iv/ PO)* | no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 68 | Vaccinations in the first month* | no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 69 | Bilirubinaemia | not measured=0, yes normal=1, yes pathological=2 | | ordinal | 0 | 2 | 9 |
| 70 | Blood glucose | not measured=0, yes normal > 2.5mmol/L=1, yes hypoglycaemia (<2.6 mmol/L)=2, yes hyperglycaemia ( > 10mmol/L)=3 | | ordinal | 0 | 3 | 9 |
| 71 | CRP | not measured=0, yes normal=1 yes pathological=2 | | ordinal | 0 | 2 | 9 |
| 72a | Bacterial culture | none=0, blood=1, CSF=2, CSF and blood=3 | | ordinal | 0 | 3 | 9 |
| 72b | If positive culture: Isolated Bacteria | Name of microorganism isolated: Escherichia Coli=1 Klebsiella spp=2, Group B Streptococcus=3, Staphylococcus Aureus=4, Staphylococcus Coagulase Negative=5, Enterococcus spp=6, Enterobacter spp=7, Pseudomonas aeruginosa=8, Proteus spp=9, Citrobacter spp=10, Acinetobacter spp=11 and other=12 | | ordinal | 1 | 12 | 99 |
| 72c | If positive culture: antibiogram | Note only the antibiotics to which the bacteria are resistant Ampicillin=1, Amoxicillin=2, CoAmoxicillin=3, Piperacillin/ Tazobactam=4, Ceftriaxone=5, Cefotaxime=6, Cefuroxime=7, Cefepime=8, Gentamicin=9, Amikacin=10, Cloxacillin=11, Flucloxacilin=12, Vanco=13, Meropenem=14, Imipenem=15, Others=16 | | ordinal | 1 | 16 | 99 |
| 73 | Date of discharge* | date - calendar | DD.MM.YYYY | date | date of admission | date of discharge | 99.99.9999 |
| 74 | Type of discharge* | home=0, home against medical advice=1, internal transfer=2, external transfer=3, death=4 | | ordinal | 0 | 4 | 9 |
| 75 | Discharge weight (grams)* | Weight at discharge in grams | | integer | 500 | 6500 | 9999 |
| 76 | Feeding at discharge* | exclusive breastfeeding=0; exclusive breast milk=1, donor milk=2, mixed=3, artificial=4 none=5 | | ordinal | 0 | 5 | 9 |

| N° | Variable | Definition | Format | Type | Min | Max | Missing value |
|---|---|---|---|---|---|---|---|
| 77 | Skin-to-skin at discharge | recommended skin-to-skin after discharge: no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 78a | Date of death* | date - calendar | DD.MM.YYYY | date | date of admission | date of discharge | 99.99.9999 |
| 78b | Time of death* | time | HH:MM | time | | | 99.9 |
| 79 | Probable cause of death* | respiratory=1, neurological=2, digestive=3, bleeding=4, hypothermia=5, infection=6, other=8 | | ordinal | 1 | 8 | 9 |
| | Structural unit variable characteristics | | | | | | |
| 1 | Name of the hospital* | manually generated | | text | | | |
| 2 | Obstetrics department in same hospital* | no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 3 | Number of births in the center per year* | number of births in the previous year | | integer | 300 | 6000 | 9999 |
| 4 | National neonatal level* | Level 1=no physician but midwife or birth attendant, the lowest would be a district maternity unit=1, level 2=with physicians=2, level 3=the highest with paediatricians=3, other=4 | | ordinal | 1 | 4 | 9 |
| 5 | Number of active neonatal beds* | average over the year | | integer | 1 | 80 | 99 |
| 6 | Dedicated neonatal resuscitation table | no=0, yes in delivery room=1, yes only in the unit=2 | | ordinal | 0 | 2 | 9 |
| 7 | Running water in the unit | no=0, cold=1, cold and hot=2 | | ordinal | 0 | 2 | 9 |
| 8 | Toilets with WC available for parents | no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 9a | Number of incubators | number | | integer | 0 | 30 | 99 |
| 9b | Number of radiant tables | number | | integer | 0 | 30 | 99 |
| 9c | Number of heated beds | number | | integer | 0 | 30 | 99 |
| 10 | Electric or wall suction | no=0, yes=1 | | boolean | 0 | 1 | 9 |
| 11 | Number of phototherapy devices | number | | integer | 0 | 30 | 99 |
| 12 | Oxygen availability | none=0, extractor=1, cylinder=2 wall=3 | | ordinal | 0 | 3 | 9 |
| 13 | Number of simultaneous functional oxygen treatments | number | | integer | 0 | 30 | 99 |

| N° | Variable | Definition | Format | Type | Min | Max | Missing value |
|---|---|---|---|---|---|---|---|
| 14 | Number of functional CPAP | number | | integer | 0 | 20 | 99 |
| 15 | Number of functional respirators (for invasive ventilation) | number | | integer | 0 | 10 | 99 |
| 16 | X-available on unit* | no=0, fixed in radiology=1, mobile in neonatology=2 | | ordinal | 0 | 2 | 9 |
| 17 | Ultrasonography available* | no=0, fixed in radiology=1, mobile in neonatology=2 | | ordinal | 0 | 2 | 9 |
| 18 | Number of effective positions for general practitioners | example: 1 physician at 100% and 1 physician at 60%=1.6 | | decimal | 0.0 | 80.0 | 99.99 |
| 19 | Number of effective positions for paediatricians or neonatologist* | example: 1 physician at 100% and 1 physician at 60%=1.6 | | decimal | 0.0 | 80.0 | 99.99 |
| 20 | Number of effective positions of caregivers with training in nursing and midwifery (state or private with diploma)* | example: 1 caregiver at 100% and 1 caregiver at 60%=1.6 | | decimal | 0.0 | 100.0 | 999.99 |
| 21 | Number of effective positions for caregivers without training in nursing and midwifery (=other caregivers who are physicians, nurses or midwives)* | example: 1 caregiver at 100% and 1 caregiver at 60%=1.6 | | decimal | 0.0 | 80.0 | 99.99 |

*Data with corresponding definitions in high income countries

**Table 4. Time requirements for data entry**

| | Center 1 | Center 2 | Center 3 | Overall |
|---|---|---|---|---|
| Number of neonates included | 428 | 918 | 693 | 2039 |
| Time to gather records (min/case) | 1.0 | 1.0 | 0.5 | 0.8 |
| Time to enter data (min/case) | 20 | 20 | 25 | 21.7 |
| Communication time with center (min/case) | 1.4 | 1.2 | 1.0 | 1.2 |
| Time per case (min) | 22.4 | 22.2 | 26.5 | 23.7 |
| Time per 100 entries (hours) | 37.3 | 37.0 | 44.2 | 39.5 |

The challenges of patient data confidentiality were solved through secondary anonymization during the testing and will now be encrypted with a clear separation of administrative patient identifications and clinical data. Data ownership entirely remains within participating centers.

### LIMITATIONS

As increasing numbers of variables tends to reduce data quality, we empirically targeted 60 clinical variables. We finished up with 73 clinical and 6 administrative patient variables, a number slightly higher than planned, due to definition issues and local requests. In comparison, most databases in high-income countries, such as the SNN and the VON, present quite larger variable numbers and tend to increase their variable numbers over time. Within the database development process and field-testing, dedicated data managers on site and a tight feed-back loop with the central data management ensured variable definition and data quality. Extension of the ***mND*** to additional centers will need continuous quality monitoring by increasing automatic plausibility routines such as used in high-income databases.

For patient data, the Delphi selection of the variables was based on a broad contribution of diverse experts from various French-speaking LMIC. The final field-testing and definition improvement was performed in one single African country, Burkina Faso. This choice was favored by an excellent clinical, educational and academic interaction between the Geneva University Hospitals and the three Burkina Faso centers, allowing easy and frequent team exchanges thought essential for the test phase. Although all definition issues have been addressed extensively during field-testing with over 2000 clinical data entries, our ***mND*** may still have some limitations for generalizability. Some of the French and possibly locally developed definitions may not directly translate to other LMIC. However, two-thirds of the variables maintained "international" definitions used in high-income countries and already have English and French translations.

We acknowledge that structural data was tested on three units only. After the Delphi process, few adaptations were made to these variables during field-testing. As the units considerably differ in terms of geographical area, administrative context, and populations characteristics, we feel the structural variables to be representative too. However, the small number and single-country (Burkina Faso) field-testing may limit its generalizability.

## CONCLUSIONS

In conclusion, we provide with our ***mND*** probably the first comprehensive, standardized neonatal database conceived and field-tested for LMIC, including French-speaking sub-Saharan Africa. It is a tool for comparative south-south quality control, for improved leverage on bottlenecks in neonatal care, support, and follow-up. Despite LMIC specificity, the elevated comparability with high-income neonatal databases allows for international comparability and development.

We are currently working on automated data plausibility controls and an online life streaming quality control output for participating centers. Gradual inclusion of additional centers will increase inter-center comparability and simultaneously increase center confidentiality. The use of comparative data for quality control, publications and research will follow an established request and acknowledgment process similar to the ones currently used by the SNN. A broad use of the data will be encouraged and supported by the Geneva University team with optional possibilities to add limited research variables for specific projects.

Our ***mND*** should provide, through high-quality data, a better understanding of in-hospital neonatal morbidity and mortality, allowing through south-south comparison, identification and follow-up of the most cost-effective interventions for each user's healthcare setting.

### ACKNOWLEDGEMENTS

We thank local data managers Francois Niada in Ouahigouya, Eulalie Kaboré, Maïmouna Samandoulougou and Fiacre Agbodossindji in Ouagadougou for data entry, regular feed-back and active variable improvement collaboration.

We are also grateful to our experts from Africa: Aminata Sow, Anne Ester Njom Nlend, Cyprien Kouakou, Daniele Kedy Koum, Eliane Kuissi, Evelyn Amine, Solange Ouédraogo, François Dianouni Niada, Jonas Ayeroue, Madeleine Folquet, Nadia Slitine, Serge Berlin Dzeukou, Mrs Endygil and experts from Switzerland: Catherine Jorgensen, Cristina Exhenry, Eric Giannoni, Flavia Rosa-Mangeret,

Francisca Barcos, Jean-Claude Fauchère, Maria Julia Rodriguez, Marie Müller, Mario Gehri, Stephane Sizonenko, Thomas Berger et Vincent Muehlethaler, Myrtha Martinet et Sylvie Loiseau and Yann Levy-Jamet for their time given to rank and select the most relevant and available variables.

Finally, we thank Mark Adams, data-manager of the SNN for his support in establishing the original data set and participation to the Delphi analysis.

## DISCLAIMER

none.

## ETHICS STATEMENT

The protocol was submitted to the Ethics Committee of Burkina Faso, specifying that anonymity of the patients would be preserved by means of an individual code and the participating centers would remain anonymous. We received written clearance without need to request patient consent, on August 12, 2020 (No. 2020-8-173). This work does not report patient data in any way, but merely used it to identify and improve variable definitions. Any personal patient data remained within the hospital files. Nominative data was available to participating local investigators for their patients only, and to the two primary investigators (PZ, RP) in Geneva during the test period. Participating centers gave their written agreement to participate and were anonymized for comparison of center data against both other centers. All authors and heads of participating centers gave consent for publication.

## DATA AVAILABILITY

Our database ***mND*** composed of 100 variables (79 clinical and administrative and 21 structural) figures in table 3 but individual data is not publicly available.

## FUNDING

The project was supported by a Swiss Government grant (Bourse d'Excellence de la Confédération Suisse 2019.0800 / Burkina Faso / OP), and the non-profit 4earlylife Association Geneva, Switzerland for financially supporting local data managers. The content is solely the responsibility of the authors and does not reflect the view of the funding bodies. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.The APC was funded by Geneva University.

## AUTHORSHIP CONTRIBUTIONS

Design: ZP, SS, RP. Supervision: RP. Contact with local data managers: ZP. Contribution to the improvement of the variables: ZP, RP, SS, SO, FRM, PO. Writing: PZ, RP. Proofreading and corrections: RP, SO, FRM, SS, PO. All authors have read and approved the final manuscript.

## DISCLOSURE OF INTEREST

The authors completed the ICMJE Disclosure of Interest Form (available upon request from the corresponding author) and disclose no relevant interests.

RP is secretary of the non-profit Association 4earlylife, Geneva, an NGO with the sole aim to support neonatal health.

## ADDITIONAL MATERIAL

Our article contains additional information as an Online Supplementary Document.

## CORRESPONDENCE TO:

Pfister Riccardo E
Maternity Hospital
32 Boulevard de la Cluse, 1211 Geneva
Swizerland
Riccardo.pfister@unige.ch
Riccardo.pfister@hcuge.ch

Submitted: October 18, 2022 GMT, Accepted: February 02, 2023 GMT

# REFERENCES


1. UNICEF, UN IGME. Child mortality report 2020. Published 2020. https://www.unicef.org/media/60561/file/UN-IGME-child-mortality-report-2019.pdf

2. UN Inter-agency Group for Child Mortality Estimation. *Levels & Trends in Child Mortality: Report 2021*.

3. Rosa-Mangeret F, Benski AC, Golaz A, et al. 2.5 Million Annual Deaths—Are Neonates in Low- and Middle-Income Countries Too Small to Be Seen? A Bottom-Up Overview on Neonatal Morbi-Mortality. *Tropical Medicine and Infectious Disease*. 2022;7(5):64. doi:10.3390/tropicalmed7050064

4. United Nations. *The Millennium Development Goals Report*. United Nations; 2015:72.

5. Liu L, Oza S, Hogan D, et al. Global, regional, and national causes of under-5 mortality in 2000–15: an updated systematic analysis with implications for the Sustainable Development Goals. *Lancet*. 2016;388(10063):3027-3035. doi:10.1016/s0140-6736(16)31593-8

6. Målqvist M, Eriksson L, Nga NT, et al. Unreported births and deaths, a severe obstacle for improved neonatal survival in low-income countries; a population based study. *BMC Int Health Hum Rights*. 2008;8(1):1-7. doi:10.1186/1472-698x-8-4

7. Pais MS. Birth registration: Right from the start. *UNICEF Innocenti Digest*. 2002;(9):1-32. http://cskc.daleel-madani.org/sites/default/files/resources/birthregistration_Digestenglish.pdf

8. Burke M, Heft-Neal S, Bendavid E. Sources of variation in under-5 mortality across sub-Saharan Africa: a spatial analysis. *Lancet Glob Health*. 2016;4(12):e936-e945. doi:10.1016/s2214-109x(16)30212-1

9. Duke T, Yano E, Hutchinson A, et al. Large-scale data reporting of paediatric morbidity and mortality in developing countries: It can be done. *Arch Dis Child*. 2016;101(4):392-397. doi:10.1136/archdischild-2015-309353

10. Fottrell E, Osrin D, Alcock G, et al. Cause-specific neonatal mortality: Analysis of 3772 neonatal deaths in Nepal, Bangladesh, Malawi and India. *Arch Dis Child Fetal Neonatal Ed*. 2015;100(5):F439-F447. doi:10.1136/archdischild-2014-307636

11. Kedy Koum D, Exhenry C, Penda CI, Nzima Nzima V, Pfister RE. Morbidité et mortalité néonatale dans un hÔpital de district urbain à ressources limitées à Douala, Cameroun. *Archives de Pédiatrie*. 2014;21(2):147-156. doi:10.1016/j.arcped.2013.11.014

12. Oza S, Lawn JE, Hogan DR, Mathers C, Cousens SN. Neonatal cause-of-death estimates for the early and late neonatal periods for 194 countries: 2000–2013. *Bull World Health Organ*. 2014;93(1):19-28. doi:10.2471/blt.14.139790

13. Muhe LM, McClure EM, Nigussie AK, et al. Major causes of death in preterm infants in selected hospitals in Ethiopia (SIP): a prospective, cross-sectional, observational study. *Lancet Glob Health*. 2019;7(8):e1130-e1138. doi:10.1016/s2214-109x(19)30220-7

14. Ouédraogo Yugbaré SO, Coulibaly G, Kouéta F, et al. Profil à risque et pronostic néonatal de l'asphyxie périnatale en milieu hospitalier pédiatrique à Ouagadougou. *J Pediatr Pueric*. 2015;28(2):64-70. doi:10.1016/j.jpp.2015.02.003

15. Nagalo K, Dao F, Tall FH, Yé D. Morbidité et mortalité des nouveau-nés hospitalisés sur 10 années à la Clinique El Fateh-Suka (Ouagadougou, Burkina Faso). *Pan Afr Med J*. 2013;14:153. doi:10.11604/pamj.2013.14.153.2022

16. Kouéta F, Ouédraogo Yugbaré SO, Dao L, Dao F, Yé D, Kam KL. Medical audit of neonatal deaths with the "three delay" model in a pediatric hospital in Ouagadougou. *Cahiers Santé*. 2011;21(4):209-214. doi:10.1684/san.2011.0271

17. Coulibaly G, Ouédraogo-Yugbaré SO, Kouéta F, et al. Asphyxie périnatale et insuffisance rénale aiguë à Ouagadougou. *Archives de Pédiatrie*. 2016;23(3):249-254. doi:10.1016/j.arcped.2015.12.002

18. Ouédraogo Yugbaré SO, Yaméogo P, Gombané B, Zombré V, Kam L. Evaluation of the human and material resources for the resuscitation and the essential care of newborns in public reference hospitals in Burkina Faso. *Internet Journal of Pediatrics and Neonatology*. 2015;18(1):1-6.

19. Ouédraogo SOY, et al. Analyse de la prise en charge du nouveau-né dans le cadre de la stratégie nationale de subvention des accouchements et des soins obstétricaux et néonatals d'urgence au Centre Hospitalier Universitaire Pédiatrique Charles de gaulle, Ouagadougou (Burkina fa. *Pan African Medical Journal*. 2015;20:1-7.

20. Boureima K, Solange O, Aïcha K, et al. Détresse vitale du nouveau-né en salle de naissance au chu-yo de Ouagadougou: Intérêt de la réanimation néo natale. *Pan Afr Med J*. 2016;23:1-8. doi:10.11604/pamj.2016.23.234.8579

21. de Graft-Johnson J, Vesel L, Rosen HE, et al. Cross-sectional observational assessment of quality of newborn care immediately after birth in health facilities across six sub-Saharan African countries. *BMJ Open*. 2017;7(3):e014680. doi:10.1136/bmjopen-2016-014680

22. Berrueta M, Bardach A, Ciaponni A, et al. Maternal and neonatal data collection systems in low- and middle-income countries: Scoping review protocol. *Gates Open Res*. 2020;4:18. doi:10.12688/gatesopenres.13106.1

23. Sa'avu M, Duke T, Matai S. Improving paediatric and neonatal care in rural district hospitals in the highlands of Papua New Guinea: A quality improvement approach. *Paediatr Int Child Health*. 2014;34(2):75-83. doi:10.1179/2046905513y.0000000081

24. Burke M, Heft-Neal S, Bendavid E. Sources of variation in under-5 mortality across sub-Saharan Africa: a spatial analysis. *Lancet Glob Health*. 2016;4(12):e936-e945. doi:10.1016/s2214-109x(16)30212-1

25. Stevenson AG, Tooke L, Edwards EM, et al. The use of data in resource limited settings to improve quality of care. *Seminars in Fetal and Neonatal Medicine*. 2021;26(1):101204. doi:10.1016/j.siny.2021.101204

26. Shrime MG, Iverson KR, Yorlets R, et al. Predicted effect of regionalised delivery care on neonatal mortality, utilisation, financial risk, and patient utility in Malawi: an agent-based modelling analysis. *Lancet Glob Health*. 2019;7(7):e932-e939. doi:10.1016/s2214-109x(19)30170-6

# SUPPLEMENTARY MATERIALS

## Online Supplementary Document

Download: https://www.joghr.org/article/75151-a-minimal-neonatal-dataset-mnd-for-low-and-middle-income-countries-as-a-tool-to-record-analyse-prevent-and-follow-up-neonatal-morbidity-and-morta/attachment/158677.docx