

1 **Table 1. Outcome measurements**

Outcomes	Tool	Measurement strategy	Rationale
Primary outcome			
Developmental delay	Bayley Scales of Infant and Toddler Development III (BSID III). The scale is used to assess language, cognitive, motor, social-emotional, and general adaptive function.	Assessment by Bayley trained clinical psychologist / clinician between 24-to-36-month chronological age. The BSID III scores will be calculated using corrected gestational age. For the BSID III, composite scores in different domains range from: Cognitive: 55-145 Language: 47-153 Motor: 46-154 Social-Emotional: 55-145 Adaptive: 40 -160. Composite score in each individual domain may be converted into scaled scores	BSID III has been developed in USA and used in LMIC countries, including India [28-33]. BSID III, in comparison to BSID II, requires a higher cutoff score for comparable sensitivity in detecting developmental delay. BSID III combined scores < 80 and cognitive and language scores <85 were found to correlate with BSID II Mental Development Index (MDI) scores <70 and clinically to moderate-severe developmental delay [31].

		that range from 1-19. Average scaled scores are between 8-12. A composite score of < 80 will indicate developmental delay.	
Cerebral palsy	Hammersmith Infant Neurological Examination (HINE)	Standardized neurological assessment using HINE at 6, 12, and 24 months. According to a recent systematic review [32], HINE scores have 90% predictive accuracy of cerebral palsy. By 12 months, score of <40 is predictive of non-ambulatory cerebral palsy [plegic, not walking] and a score of 40-60 is predictive of ambulatory cerebral palsy [able to walk], with any score <70 considered	HINE is a strongly recommended scale for early diagnosis of cerebral palsy [32]. It is a standardized and scorable clinical neurological examination that is applicable for infants aged between 2 and 24 months and has been shown to be an easy to train tool with interobserver reliability even in less experienced staff [38].

		abnormal. Maximum score is 78. The accepted norms will be used for cutoff in this study, which are >70 at 6 month and >73 at 12 months and older [32-37]. Due to the risk of more subtle findings, particularly in premature infants, there is increased sensitivity to the tool when used at multiple time points, therefore, the examination will be performed at an additional time point of 2 years.	
Hearing impairment	Hearing assessment measured ideally at discharge or as soon as possible at enrolment or first	First assessment by screening auditory brainstem responses (ABR). If screening ABR is	International guidelines recommend hearing assessment within the first month of life whenever feasible [39].

	follow-up	passed, no need for further hearing assessment tests. In case screening ABR is failed, the participant is referred at the earliest possible opportunity to audiologist or specialist for re-screening, including, diagnostic ABR as applicable.	While Otoacoustic Emission (OAE) screening can be used in settings where ABR is not available, it is less preferred due to the potential for missing retro-cochlear hearing loss. Therefore, in studying high risk infants, ABR is preferred [39-42].
Vision impairment	Structural visual examinations, including for retinopathy of prematurity, will be performed between discharge and 3 months, with visual acuity measured at 12 months.	Visual acuity measured by Teller Acuity Cards at the facility. Impairment defined by measurements 2 lines less than age defined normative values.	Teller Acuity Cards will be used due to availability of age-specific reference ranges [43,44].
Epilepsy	Screen for seizure	The operational definition	Screening questionnaires for

	at each follow-up visit. If screening positive refer to clinician.	of epilepsy per the International League against Epilepsy (ILAE) has room for interpretation with the expectation of implementation by a provider with expertise in epilepsy and with the understanding that there is no single clinical factor that can convey this risk across all patients, and each case must be assessed individually [45]. To reduce variability in diagnosis for purposes of this study, the ILAE diagnostic criterion has been operationalized for this study in the screening questionnaire as either 2 unprovoked seizures separated by at least 24	pediatric epilepsy are limited due to small populations and limited age range and have demonstrated moderate validity [46, 47]. A questionnaire comprised from validated pediatric epilepsy screening questionnaires and agreed upon by pediatric neurologists from the participating sites, has been employed after local pilot testing.
--	---	--	---

		hours or 1 unprovoked seizure and one other positive finding of neurodevelopmental impairment.	
Secondary outcomes			
Mortality	Screen for death of enrolled child occurring any time between 3 months of age and 3 years of age	Mortality/survival status at each screening point. Assessments performed every 3 months until the age of 3 years.	
Growth	Anthropometric parameters of weight (W), length (L) and head circumference (HC) for age using standardized equipment.	Anthropometry at 6, 12, and 24 months of age. Parameters are plotted as per corrected age on WHO Multicentre Growth Reference Study (MGRS) charts and interpreted as per z scores. Postnatal growth failure defined as fall in z scores	Measurements are obtained during facility visits. In case participant unable to visit facility, we perform the measurement at the home of the participant.

		or z scores < 2 SD.	
Feeding practices	Assessment of breastfeeding, expressed breast feeds, complementary feeds at 6 months and adequacy of feeding beyond infancy.	Feeding practices measured through standard questionnaire at each follow-up visit- every 3 months until the age of 3 years.	
Maternal depression	Patient Health Questionnaire 9 (PHQ-9)	Maternal depression measured using PHQ-9 at 6 and 12 months of age.	Literature suggests, while early screening for post-partum depression is important, additional screening points at either 6 or 12 months can capture additional women who may have initially screened negative at earlier time points. Therefore, we have included these time points

			in this study [48].
Home environment	HOME	Assessment of home environment measured using standard HOME - SF scale at 18 months by a home visit.	HOME is a validated tool across cultures. The MAL-ED study has also identified that a three-factor structure- emotional and verbal responsiveness, clean and safe environment, child cleanliness can be examined cross culturally [49].

2

3

4 Table 2: Outcome measurement schedule

[illegible]

admission at least overnight		X	X	X	X	X	X	X	X
Home environment							X		
Maternal Depression			X		X				
Approximate duration of study visit	30 min	30 min	45 min	20 min	60 min	20 min	30 min	20 min	60 - 90 min