

# An Approach for Measuring Developmental Outcomes in Neonatal Trials.

## Assessment of the ASQ-3 and BSID-III

Kristy P Robledo<sup>1, ID</sup>  
@redheadkristy  
kristyrobledo@ctc.usyd.edu.au

Ingrid Rieger<sup>2</sup> William Tarnow-Mordi<sup>1</sup> Sarah Finlayson<sup>1</sup> Andrew J Martin<sup>1</sup>  
<sup>1</sup> NHMRC Clinical Trials Centre, University of Sydney  
<sup>2</sup> Department of Newborn care, RPA Women and Babies

### Introduction

The Bayley Scales of Infant Development III (BSID-III) is a criterion assessment of the developmental trajectory of infants<sup>1</sup> and a valid, but labour intensive and relatively expensive, tool for assessing long-term developmental outcomes in neonatal trials. The Ages and Stages Questionnaire (ASQ-3) is a parent-completed questionnaire that has potential for use as an affordable tool in this setting, but it was originally developed as a screening tool and it has not been validated against the BSID-III.

### Objectives

1. Compare ASQ-3 to BSID-III
2. Assess current ASQ-3 cutpoints for predicting disability according to BSID-III
3. Identify optimal cutpoints for ASQ-3

### Methods

Babies born <30 weeks gestational age and enrolled in the Australian Placental Transfusion Study (Tarnow-Mordi et al. 2017) were assessed at 24 months of age with the ASQ-3 and BSID-III. Both instruments assess:

- cognitive,
- language,
- fine motor,
- gross motor and
- social domains.

**Moderate Delay:** >2 standard deviations below average.

### Analyses

- Spearman correlations ( $r$ ),
- kappa ( $k$ ) to assess agreement
- Receiver Operating Characteristic curves (ROC) to assess ASQ-3 to predict moderate delay on BSID-III
- Area under the ROC curve (AUC)
- Sensitivity and specificity statistics.

### Results

405 infants have BSID-III performed  $\pm$  3 months of ASQ-3. The median corrected age was 24 months (range: 22- 42). The mean gestation of babies born was 27.6 weeks (SD 1.6) with mean birthweight 1016g (SD 262). 50% were treated with deferred cord clamping, 58% were male and 79% were singleton births.

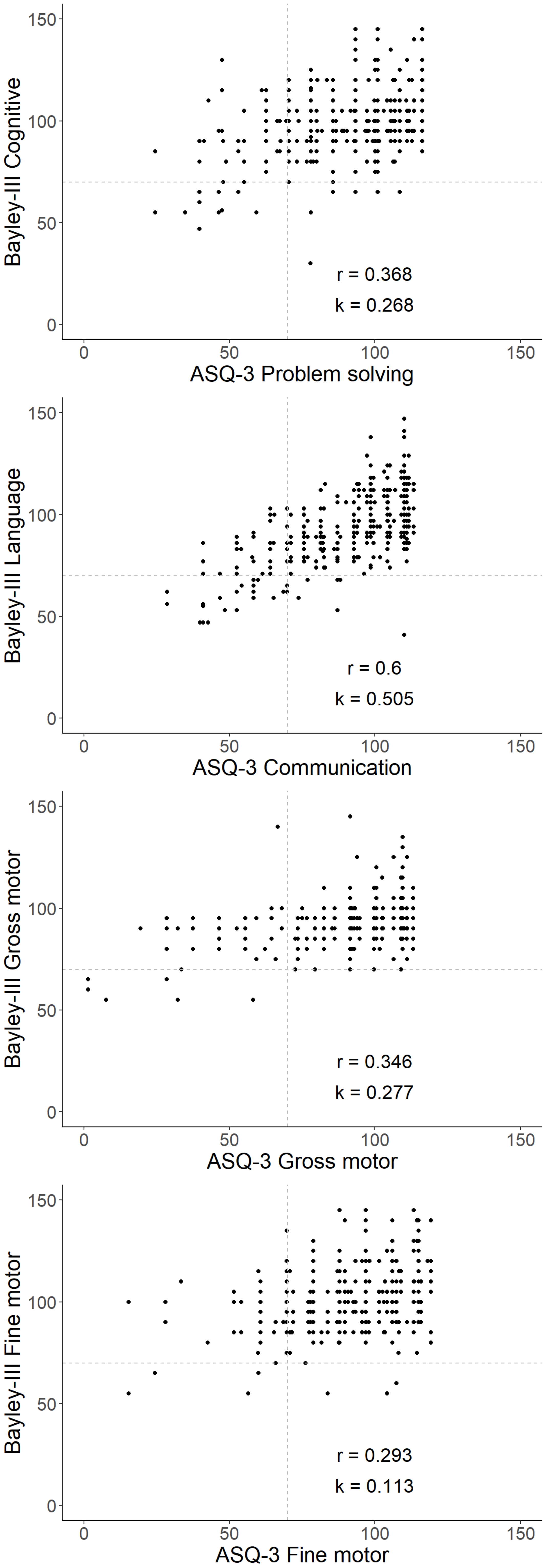


Table 1: Paired ttests for each domain

Domain	ASQ mean (SD)	BSID mean (SD)	Mean difference (95%CI)	P-value
Cognition	89.8 (19.5)	99.5 (16.2)	-9.7 (-11.6 to -7.8)	< 0.001
Language	91.9 (20)	94 (17.6)	-1.7 (-3.3 to -0.1)	0.033
Fine motor	93.2 (18.9)	103.3 (15.6)	-10.2 (-12.2 to -8.2)	< 0.001
Gross motor	92.3 (22)	93.2 (11.1)	-0.8 (-2.8 to 1.1)	0.4

Table 2: Sensitivity and Specificity for ASQ-3 traditional cutpoints

Domain	Sensitivity	Specificity
Cognition	65% (38-86)	89% (86-92)
Language	81% (64-93)	90% (86-93)
Fine motor	57% (18-90)	88% (85-91)
Gross motor	100% (66-100)	89% (86-92)

Table 3: Sensitivity and Specificity for ASQ-3 optimal cutpoints

Domain	Sensitivity	Specificity
Cognition	65% (38-86)	95% (92-97)
Language	84% (67-95)	88% (84-91)
Fine motor	71% (29-96)	71% (66-75)
Gross motor	100% (66-100)	92% (89-95)

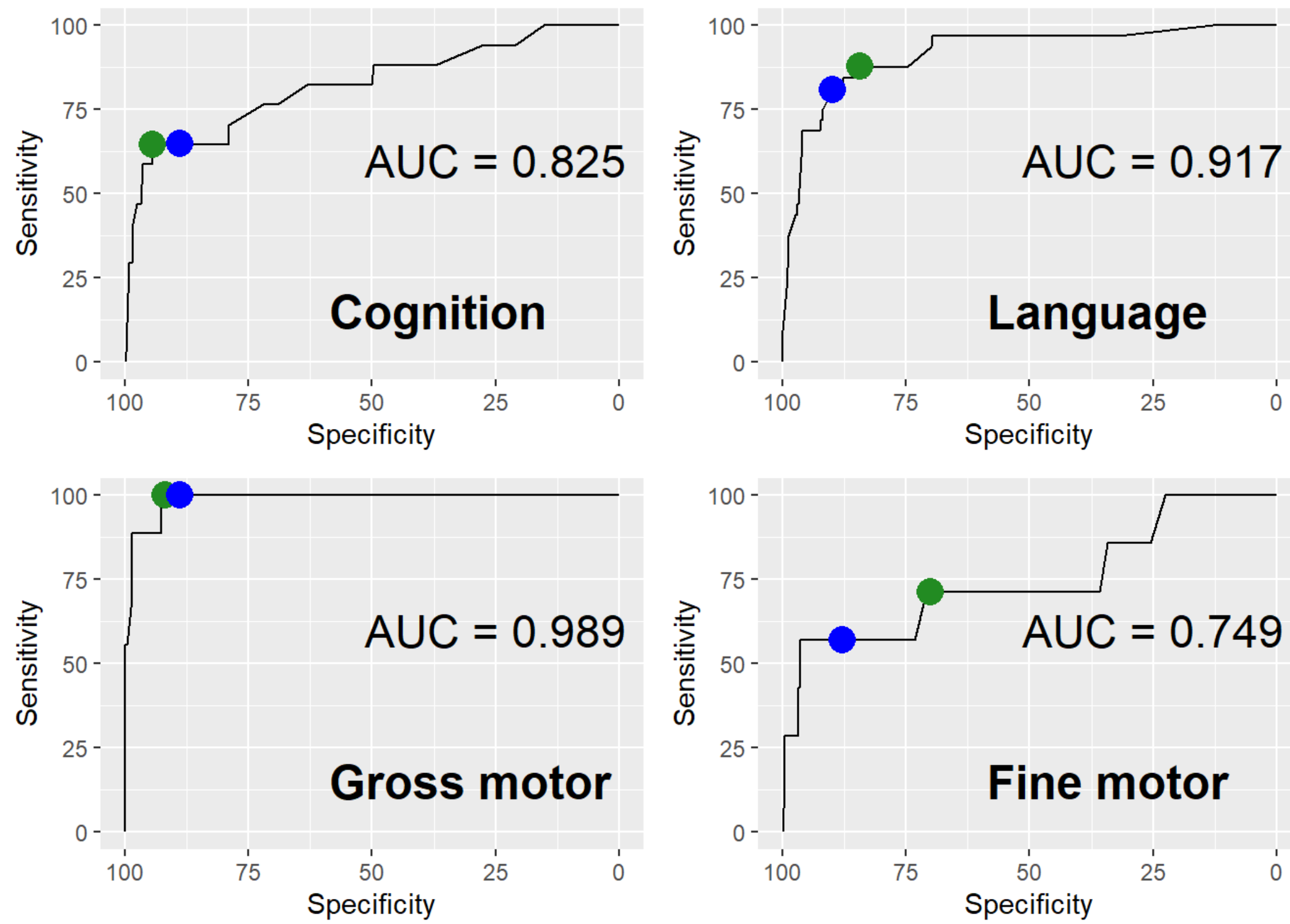


Figure 1: ROC for ASQ-3 domains. Blue indicates traditional cutpoints. Green indicates optimal cutpoints.

### Conclusions

1. ASQ-3 and BSID-III are moderately well correlated
2. Alternative ASQ-3 cutpoints with higher sensitivity are suggested to avoid missing children with moderate delays.
3. The ASQ-3 is a reasonable tool for assessing outcomes in RCTs, but should not replace diagnostic assessments of delay.

### References

Tarnow-Mordi, William, Jonathan Morris, Adrienne Kirby, Kristy Robledo, Lisa Askie, Rebecca Brown, Nicholas Evans, et al. 2017. "Delayed Versus Immediate Cord Clamping in Preterm Infants." *New England Journal of Medicine* 377 (25): 2445–55. <https://doi.org/10.1056/NEJMoa1711281>.

