Calibration Performance Across UQ Methods

Method	90% CI Coverage (%)
Multi-output BNN	92.0%
Quantile Regression	89.4%
Conformal Prediction	84.5%
BNN + Opponent	31.3%
BNN + Environment	29.7%
BNN + Vegas	29.7%
BNN Baseline	26.2%

Legend:
- Severely Under-calibrated (<50%)
- Under-calibrated (50-80%)
- Approaching Target (80-88%)
- Well-calibrated (≥88%)
- Target (90%)