**Prognostic effect of molecular subgroups**

To assess principal molecular subgroup outcomes derived from the WHO guidelines, patients were classified by principal molecular subtype as described in the 2021 WHO Classification (i.e., WNT, SHH TP53+, SHH TP53-, non-WNT/non-SHH). This was further compared to results when classified according to the 2016 WHO Classification (i.e., WNT, SHH, G3, and G4).

We first assessed survival in a Kaplan–Meier analysis of the two classification schemes. The log-rank test revealed highly significant differences in survival across subgroups for both molecular classification schemes, with the 2016 classification demonstrating more statistically significant between subgroup differences (P < 0.0001 versus P = 0.0043; Figure 2A and 2B). Meanwhile, favourable survival outcomes for WNT subtype were confirmed in Kaplan–Meier analysis. We then examined the Kaplan-Meier estimates of survival probability, specifically focusing on G3 versus G4 and SHH TP53- versus SHH TP53+ subgroups. The results are displayed in Figure 2C and 2D. There was no statistically significant difference in survival distributions between SHH TP53- and SHH TP53+ (P = 0.081); the five-year survival probabilities were 65.4% (95% CI, 51.8% to 76.0%) in SHH TP53- and 36.5% (95% CI, 9.3% to 65.2%) in SHH TP53+. Distinct survival outcomes were observed for G3 and G4 (P = 0.0013), with a five-year overall survival probability of 54.3% (95% CI, 43.1% to 68.4%) for G3 and 71.7% (95% CI, 65.7% to 78.2%) for G4.

We then used Cox proportional hazard model to evaluate simultaneous effect of risk factors on survival rates. The hazard ratio (G3 vs G4) was 1.89 (95% CI, 1.27 to 2.81) in univariable analysis (Figure 3A). The elevated hazard for G3 remained significant in multivariable analysis (HR, 1.90; 95% CI, 1.27 to 2.83) after adjusting for prognostic markers that were significant in univariable Cox analyses, including age, hydrocephalus before surgery, and LCA-pathology (Figure 3A and 3B). In multivariable Cox regression analysis, age (P = 0.04; HR, 1.87; 95% CI, 1.04 to 3.37) was the only remaining significant independent prognostic factor in non-WNT/non-SHH patients (Figure 3B). Moreover, there was no statistically significant difference in hazard between SHH TP53+ and TP53- subgroups in both univariable analysis (HR, 1.71; 95% CI, 0.93 to 3.15) (Figure 3A) and multivariable analysis (HR, 1.58; 95% CI, 0.84 to 2.97) (Figure 3C).

The WHO CNS5 principal classification (2021 WHO) achieved a C-index of 0.55 (95% CI, 0.49 to 0.60) in the univariable Cox model analysis of predicting overall survival, indicating slightly less discriminatory capability compared to the 2016 principal classification, which had a C-index of 0.59 (95% CI, 0.52 to 0.65). The results remained consistent in the multivariable Cox analysis, with a C-index of 0.59 (95% CI, 0.52 to 0.66) for the WHO CNS5 classification and a C-index of 0.61 (95% CI, 0.53 to 0.68) for the 2016 classification. When combining both molecular classifications, the five-molecular-subgroup scheme did not outperform the previous classification, with a C-index of 0.58 (95% CI, 0.51 to 0.64) in the univariable analysis and 0.60 (95% CI, 0.53 to 0.67) in the multivariable analysis.

**STATISTICAL ANALYSIS**

The prognostic value of the two molecular classification schemes was evaluated by three metrics: 1) Kaplan-Meier curves with log-rank tests to compare patient subgroups; 2) Hazard ratios (HR) for overall survival, including 95% confidence intervals (CIs), from univariate and multivariable Cox proportional hazard regression; 3) Concordance index (Harrel’s C-index) from the Cox models. The C-index from multivariate Cox model was estimated via cross-validating bootstrapped datasets to avoid the optimism bias. We used forest plots to visualize the HRs and their CIs from Cox model stratified by subgroups. In addition, univariate analyses were conducted to compare patient characteristics between two groups, using the t test for continuous variables and chi-squared test for categorical variables. The significance threshold was set at P < 0.05 for all statistical tests in this study.



