The Relationship Between Overall Survival (OS), Progression-Free Survival (PFS), and Objective Response Rate (ORR) in Immune Checkpoint Inhibitor Clinical Trials of Head and Neck Squamous Cell Carcinoma (HNSCC): A Systematic Review and Meta-analysis

Poster No: 951P

Wang X[1], Ballas M[1], Chen H[1], Bell KF[1], Slowley A[2], Ellis C[1], Turner DC[1], Zhou H[1]

[1]GSK, Collegeville, PA, USA; [2]GSK, Brentford, Middlesex, UK

Aims

Immunotherapies for recurrent/metastatic (R/M) HNSCC have improved OS; however, many patients (80%) with metastatic HNSCC do not respond to, or do not benefit from, programmed cell death protein 1 (PD-1) blockade. Therefore, there remains an unmet need for new treatment approaches in this patient population.[1,2]

OS is the gold standard endpoint used to measure efficacy of novel therapeutics in oncology pivotal trials; however, long-term follow-up is required to establish OS benefit.[3,4]

Tumor-based assessments such as ORR and PFS can be used as measures of clinical benefit and are used extensively as surrogate endpoints in early-phase trials to inform and expedite drug development decisions in lieu of OS.[3,4]

Prior studies suggest that the relationship between ORR or PFS and OS may vary by cancer type, treatment mechanism of action and disease setting.[3,4]

Direct investigation of these endpoints as surrogate markers of efficacy in R/M HNSCC is needed to allow for faster patient access to novel therapies.

Objective

Investigate the association of surrogate markers PFS and/or ORR with OS for immuno-oncology (I-O)-based regimens in patients with R/M HNSCC by meta-analysis of aggregate data identified in a systematic literature review (SLR), with the aim of informing future clinical trial planning.

Methods

The workflow for this study is shown in **Figure 1.**

An SLR was performed to identify relevant Phase II and III trials that investigated I-O treatments and I-O/chemotherapy combinations for R/M HNSCC.

- Quality assessment for each randomized controlled trial was performed to the standards recommended by Cochrane.[5] Data for OS, PFS and ORR were reported in each trial.

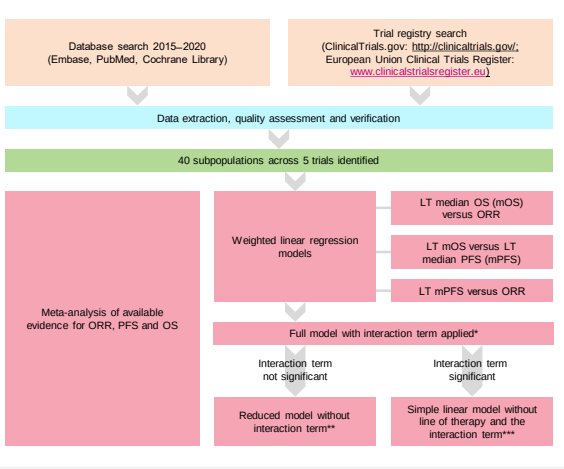
The meta-analysis to investigate the association among ORR, log-transformed (LT) median PFS (mPFS) and LT median OS (mOS) used a fixed-effect, weighted linear regression model, with weights equal to sample size of each subpopulation, and included line of therapy as a covariate.

The full model with interaction term was applied.

- Statistical significance of interaction terms between lines of therapy and ORR or LT mPFS determined whether the reduced model without interaction term or the simple linear model without line of therapy and the interaction term was fitted (**Figure 1**).

This model was found to be relevant for I-O alone; the change in PFS and ORR for chemotherapy had a much smaller effect on OS. As such, these data were not pooled with I-O data.

Figure 1. SLR and meta-analysis workflow for data extraction from selected trials



*Full term: $Y \sim \beta_0 + \beta_1{}^*X + \beta_{LINE}{}^*\text{Line} + \beta_{interaction}{}^*\text{Line}_i{}^*X$, where i=1 or ≥2; **Reduced model: $Y \sim \beta_0 + \beta_1{}^*X + \beta_{Line\_i}$, where i=1 or ≥2; ***Simple linear: $Y \sim \beta_0 + \beta_1{}^*X$ 3 sets of Y versus X: LT mOS versus ORR; LT mOS versus LT mPFS; LT mPFS versus ORR

Summary of SLR Efficacy Outcomes

Overall, 4 trials in the ≥2 treatment line setting and 1 trial in the first-line R/M HNSCC setting were identified for data extraction and analysis.

- The overall patient population is shown in **Table 1** as reported for I-O monotherapy or I-O agent combination therapy for the selected trials.

Table 1. Clinical outcomes of SLR screened trials

Trial	Treatment	Subgroup	mOS (month)	mPFS (month)	ORR (%)
KEYNOTE-40	Pembro	Overall	8.4	2.1	14.6
KEYNOTE-48	Pembro	Overall	11.5	2.3	17.0
EAGLE	Durva+treme	Overall	6.5	2.0	18.2
	Durva	Overall	7.6	2.1	17.9
CheckMate-141	Nivo	Overall	7.7	2	13.3
CONDOR	Durva+treme	Overall	7.6	2	7.8
	Treme	Overall	5.5	1.9	1.6
	Durva	Overall	6	1.9	9.2

Durva, durvalumab; nivo, nivolumab; pembro, pembrolizumab; treme, tremelimumab

Results

Meta-analysis Modelling

Representations for each study group in the meta-analysis model relationships are shown by circle colour and size:

- Grey-filled circles represent KEYNOTE-48 in the first-line R/M HNSCC setting
- Open circles represent the other 4 studies in the ≥2 lines of treatment setting.

For LT mOS versus ORR and LT mOS versus LT mPFS, the interaction terms were not statistically significant, with p=0.33 and 0.51, respectively; so the reduced model with no interaction term was fitted.

Model fitting results showed a strong relationship between LT mOS and ORR, and LT mOS and LT mPFS, with squared correlation coefficients (r^2) of 0.77 (**Figure 2**) and 0.78 (**Figure 3**), respectively.

In the weighted simple linear model for LT mPFS versus ORR, the coefficient for ORR was statistically significant (p=0.013); however, the association between LT mPFS and ORR was weak, with r^2 of 0.26 (**Figure 4**).

The linear regression model demonstrated the following:

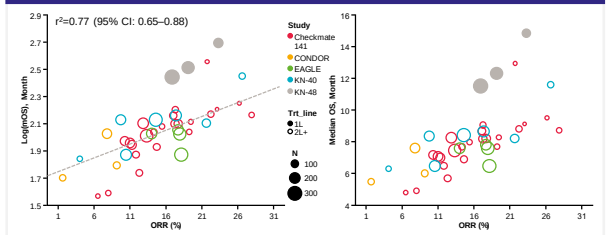
- LT mOS (month) versus ORR (%) - the regression coefficient for ORR was estimated to be 0.02 (95% CI: 0.012–0.028), indicating that a 10% ORR improvement will lead to LT mOS improvement of 0.2 (95% CI: 0.12–0.28), or a 1.22-fold (95% CI: 1.13–1.32) improvement in mOS (mOS_i/mOS_j) if converting LT mOS to the original scale.
- LT mOS (month) versus LT mPFS (month) - the slope was estimated to be 0.54 (95% CI: 0.08–1.00), indicating that a 1-unit improvement of LT mPFS will lead to a 0.54-unit improvement of LT mOS. Converting to the original scale, a 2-fold improvement of mPFS ($mPFS_i/mPFS_j$) will lead to a 1.22-fold (95% CI: 1.06– 2.00) improvement of mOS (mOS_i/mOS_j).
- LT mPFS (month) versus ORR (%) - the slope was estimated to be 0.01 (95% CI: 0.002–0.017), which implies 10% ORR improvement will lead to LT mPFS improvement of 0.1 (95% CI: 0.02–0.17), or a 1.1-fold (95% CI: 1.02–1.19) improvement in mPFS ($mPFS_i/mPFS_j$) if converting LT mPFS to the original scale.

Exploratory Analysis

The results above are based on the fixed-effect model. An exploratory analysis was performed using the random-effect model.[6] In the random-effect model, "study" was entered as a random effect, taking into account the variation between and within the trials.

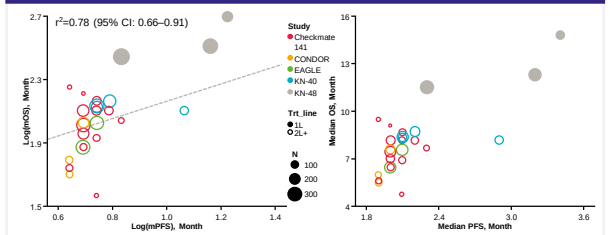
- The r^2 for LT mOS versus ORR was 0.72 (95% CI: 0.59–0.85) and the r^2 for LT mOS versus LT mPFS was 0.56 (95% CI: 0.35–0.78). Even though r^2 estimated from the random-effect model was smaller than the fixed-effect model, the magnitude was still considered to be moderate to strong.

Figure 2. Relationship between (A) LT mOS and ORR and between (B) mOS and ORR



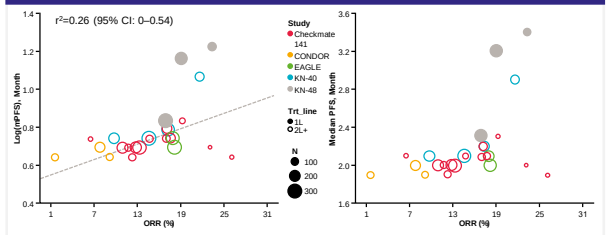
There are 40 data points presented in each panel of Figure 2. The regression line for ≥2 R/M HNSCC treatment line settings based on the reduced model is shown in the left panel. A statistically significant association between LT mOS and ORR (p<0.0001) was demonstrated.
CM-141, CheckMate-141; KN-40, KEYNOTE-040; KN-48, KEYNOTE-048

Figure 3. Relationship between (A) LT mOS and LT mPFS and between (B) mOS and mPFS



Due to PFS not reported for some of the subgroups, there are 25 data points presented in each panel of Figure 3. The regression line for ≥2 R/M HNSCC treatment line settings based on the reduced model is shown in the left panel. A statistically significant association between LT mOS and LT mPFS (p<0.0001) was demonstrated.

Figure 4. Relationship between (A) LT mPFS and ORR and between (B) mPFS and ORR



Due to PFS not reported for some of the subgroups, there are 25 data points presented in each panel of Figure 4. The regression line for ≥2 R/M HNSCC treatment line settings based on weighted simple linear model is shown in the left panel. A statistically significant interaction term for LT mPFS versus ORR was detected in the full model (p=0.003). As such, the weighted simple linear regression model was fitted by including trials in the R/M HNSCC ≥2 treatment line setting only.

Discussion

Trials in the first-line and ≥second-line R/M HNSCC treatment settings were pooled together to estimate the regression coefficients as the interaction terms for LT mOS versus ORR or LT mPFS were not statistically significant.

- This meta-analysis demonstrates strong positive correlation between ORR and LT mOS and between LT mPFS and LT mOS. The regression coefficients in the fitted model enable the prediction of the endpoint OS to predict therapeutic efficacy based on endpoint ORR or PFS.
- This meta-analysis only included one clinical trial in the first-line R/M HNSCC setting (KEYNOTE-48); the result for the interaction term and the regression coefficients could change if re-evaluating the relationship by including more clinical trials in this setting if available in the future.

This meta-analysis included patients within subgroups of the same study that are not mutually exclusive, potentially yielding stronger associations between the predicted and the regressor endpoints than is true. Ideally, mutually exclusive subgroups are preferred for this type of analysis and this could be accounted for in the future by appropriate partial de-weighting for the non-exclusive subgroups.

- To account for this limitation, sensitivity analysis that would include only mutually exclusive subgroups in each trial could be performed, though the number of observations would greatly decrease.

Conclusions

The meta-analysis, based on a systematic review of R/M HNSCC I-O clinical trials, showed significant correlation between ORR and LT mOS and between LT mPFS and LT mOS.

- These findings support ORR and PFS as surrogate endpoints to predict OS effect in early phases of clinical trials evaluating I-O treatments in R/M HNSCC.

For late-stage clinical trials in the first-line R/M HNSCC setting, ORR or PFS may be used as adaptive decision criteria to inform the design when trials are adopting OS as a primary endpoint. Early assessment tools such as ORR and PFS have been implemented in a Phase II/III trial in R/M HNSCC sponsored by GlaxoSmithKline (INDUCE-3, NCT04128696).

Disclosures

XW is employed by and is a shareholder with GlaxoSmithKline (GSK). **MB** is an employee and holds a leadership role, has received travel/accommodation expense payments, is a shareholder with licensing/royalties at GSK; is a shareholder with Bristol-Myers Squibb; has licensing/royalties with AstraZeneca. **HC** is an employee and shareholder with GSK. **KFB** is an employee and shareholder with GSK; is a spouse/financial dependent with Humana Pharmacy; is a shareholder with CVS. **AS** is an employee and shareholder with GSK. **CE** is an employee and shareholder with GSK. **DCT** is an employee and shareholder with GSK. **HZ** is an employee and shareholder with GSK.

Acknowledgments

We thank Daniel Parks, Director VEO DMA, who was assisted by Jessica Chao, Manager VEO DMA, for conducting the quality control (QC). We also want to thank Ruth Bull, Emma Hawe, Jean-Gabrial Le Moine and Outi Ahdesmaki at RTI-Health Solutions for completing the SLR portion of this study and Carlos Baccan, MD, for his contribution to the original abstract for this poster presentation. Funding for this study 214479 was provided by GSK. Editorial support was provided by Shawna Graves, at Fishawack Indicia Ltd, UK, and was funded by GSK.

References

1. Bauml JM, et al. *Ann Transl Med*. 2019;7:S75.
2. Kao HF and Lou PJ. *Head Neck*. 2019;41:Suppl 1:4–18.
3. Shukuya T, et al. *J Thorac Oncol*. 2016;11:1927–1939.
4. Zhao S, et al. *Int J Cancer*. 2019;144:2854–2866.
5. Higgins JPT and Green S. March 2011. http://handbook-5-1.cochrane.org/.
6. Buyse M, et al. *Biostatistics*. 2000;1:49–67.

Presenting author email address: xiaowei.x.wang@gsk.com

Presented at the European Society for Medical Oncology Congress (ESMO), Virtual Format, 19–21 September, 2020