The combined use of propensity score matching and a joint tumor growth dynamics (TGD) - Overall Survival (OS) model to benchmark the efficacy of new treatments for advanced renal cell carcinoma (RCC)

Mayu Osawa\textsuperscript{1}, Martin Winiger\textsuperscript{2}, Ramon Garcia\textsuperscript{3}, Jonathan French\textsuperscript{3}, Anna Kondic\textsuperscript{1}, Bauke Stegenga\textsuperscript{2}, and Amit Roy\textsuperscript{1}

\textsuperscript{1}Clinical Pharmacology and Pharmacometrics, Bristol Myers Squibb, Princeton, NJ, United States
\textsuperscript{2}Worldwide Scientific Collaborations, Global Medical, Bristol-Myers Squibb, Princeton, NJ, United States
\textsuperscript{3}Metrum Research Group, Tariffville, CT, United States
Background

• Non-registrational data generation (NRDG) and investigator sponsored research (ISR) studies typically have a small number of patients, are single arm studies, and have short-term follow-up of overall survival (OS) due to the nature of the research.

• These limitations make it difficult to compare overall survival data from NRDG/ISR studies to an appropriate benchmark treatment.

• The present work is to establish feasibility of using a joint tumor growth dynamic model (TGD-OS) to extrapolate OS and generate a synthetic control arm from historical data by propensity score modeling.
# Introduction to disease and treatments

- **Renal cell carcinoma (RCC)**
  - 90% of kidney cancer
  - 30% manifest itself as metastatic disease (mRCC)
  - Karnofsky performance score (KPS) 0-100 is a measure of functional impairment
  - MSKCC/Motzer risk score (0-5) is calculated based on 5 independent measurements
  - IMDC—International metastatic RCC database consortium score (0-4)

- **Treatment paradigms for mRCC**
  - Tyrosine kinase inhibitors (TKIs): sunitinib (anti-VEGFR2, PDGFR and KIT) …
  - Immuno-oncology (IO) therapies: ipilimumab (Ipi,I) and nivolumab (Nivo,N) …

---

**Diagram:**

- **Dendritic cell**
  - MHC
  - CD28
  - CTLA-4
- **T cell**
  - TCR
  - CTLA-4 is expressed on T cells and inhibits T-cell activation
- **Anti-CTLA-4**
- **Anti-PD-1**
- **PD-1** expression on tumor-infiltrating lymphocytes interacting with PD-L1/PD-L2 is associated with reduced effector function
Combined Modeling Approach (TGD-OS + PSM)

- A previously developed joint TGD-OS model for advanced RCC was refined (slide 4 and 5)
- The fitted joint TGD-OS model was then used to predict OS of RCC patients in Cohort 1 of NRDG study CA209-920 (NCT02982954) in which subjects received an investigational dosing regimen (N6I1)
- The predicted OS for CA209-920 (Cohort 1) was benchmarked against the approved NIVO + IPI treatment regimen (N3I1) investigated in the registrational study of CA209-214 by utilizing a propensity score model (PSM)
Joint TGD-OS model (1/2)

Joint TGD-OS model was developed with data from 1275 subjects from historical RCC studies, including CA209-214

Longitudinal Model for SLD

• Mixture Wang model (Feng et al. CPT PSP 8 (2019):825-834)

\[
\hat{y}_{i,j,t} = TB0_{i,l} \times \exp(-TS_{i,l}t_{ij}) + TG_{i,l}t_{ij} + TLIM_{i,l}
\]

Where \( i \) indexes patients, \( j \) indexes observations, \( l = 1,2,3 \) indexes sub-population

• \( TB0_{i,l} \) - is the baseline tumor burden
• \( TS_{i,l} \) - tumor shrink rate (1/week)
• \( TG_{i,l} \) - tumor growth rate (cm/week)
• \( TLIM \) - the approximate tumor limit of patient

• The following covariates were included based on either improved model performance on BIC or covariates of clinical interest
  • \( TB0 \) - Albumin + KPS ((< 90, ≥ 90)
  • \( TS \) - Line of therapy + Number of Index Lesions + PDL1 status
  • \( TG \) - Albumin + Line of therapy + Number of Index Lesions + PDL1 status
Joint TGD-OS model (2/2)

Model for OS

• Parametric (log-logistic) model was selected as it showed lowest BIC compared with the other models (Weibull and Gompertz)
• Time-varying tumor size and Tumor growth and shrinkage were included on the hazard
• Covariates were included by backward elimination from the full model based on BIC
  • MSKCC risk score
  • KPS (< 90, ≥ 90)

Hazard of death at time $t$ for patient $i$

$$h_i(t | x_i) = \frac{\left( \frac{T}{\lambda_i(t)} \right) \exp \left( (\gamma - 1) \log \left( \frac{t}{\lambda_i(t)} \right) \right)}{1 + \exp \left( \gamma \log \left( \frac{t}{\lambda_i(t)} \right) \right)}$$

where $\log \lambda_i(t) = x_{OS,i}^T \beta + \text{TumEff}_i(t) + \text{NewEff}_i(t)$

- $\gamma$ : shape parameter
- $x_{OS,i}$ : baseline covariates of patient $i$
- TumEff($t$) : effect of time-varying absolute tumor size derived from the TGD model
- $\theta_{ATS \log} \left( \frac{y_{i,t}(t)}{10} \right)$
- NewEff($t$) : time-varying effect of the appearance of new lesions

\[
\text{NewEff}(t) = \begin{cases} 
0 & t < t_{i,\text{NEWL}} \\
\theta_{\text{NEWL}} & t \geq t_{i,\text{NEWL}} 
\end{cases}
\]
Observed and Predicted OS of CA209-920 (NRDG Study) using TGD-OS Model

- Observed overall survival in 2020
- Predicted OS using tumor size data in 2018 as input into the joint TGD-OS model

TGD-OS model using tumor size data in 2018 adequately predicted OS curve in 2020
KM Curve of OS in CA209-214 and 920 (Unweighted)

- OS of N6+I1 (CA209-920) appears less favorable than that of N3+I1 (CA209-214) based on cross-study comparisons.

- Results of cross-study comparison could be misleading, due to imbalances in subject characteristics.

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Categories</th>
<th>CA209214-Sunitinib</th>
<th>CA209214-N3I1</th>
<th>CA209920-N6I1</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMDC</td>
<td>poor</td>
<td>95 (17.8)</td>
<td>101 (18.5)</td>
<td>20 (18.9)</td>
</tr>
<tr>
<td>IMDC</td>
<td>intermediate</td>
<td>332 (62.3)</td>
<td>327 (59.8)</td>
<td>86 (81.1)</td>
</tr>
<tr>
<td>IMDC</td>
<td>favorable</td>
<td>106 (19.9)</td>
<td>119 (21.8)</td>
<td></td>
</tr>
<tr>
<td>Karnofsky</td>
<td>less than 80</td>
<td>52 (9.76)</td>
<td>54 (9.87)</td>
<td>1 (0.943)</td>
</tr>
<tr>
<td>Karnofsky</td>
<td>80</td>
<td>95 (17.8)</td>
<td>82 (15.0)</td>
<td>24 (22.6)</td>
</tr>
<tr>
<td>Karnofsky</td>
<td>greater than 80</td>
<td>386 (72.4)</td>
<td>411 (75.1)</td>
<td>81 (76.4)</td>
</tr>
<tr>
<td>Line of therapy</td>
<td>1st line</td>
<td>529 (99.2)</td>
<td>535 (97.8)</td>
<td>104 (98.1)</td>
</tr>
<tr>
<td>Line of therapy</td>
<td>2nd line</td>
<td>4 (0.750)</td>
<td>12 (2.19)</td>
<td>2 (1.99)</td>
</tr>
<tr>
<td>MSKCC</td>
<td>poor</td>
<td>50 (9.38)</td>
<td>42 (7.68)</td>
<td>30 (28.3)</td>
</tr>
<tr>
<td>MSKCC</td>
<td>intermediate</td>
<td>210 (39.4)</td>
<td>209 (38.2)</td>
<td>41 (38.7)</td>
</tr>
<tr>
<td>MSKCC</td>
<td>favorable</td>
<td>273 (51.2)</td>
<td>296 (54.1)</td>
<td>35 (33.0)</td>
</tr>
</tbody>
</table>

![KM Curve Image]

Number at risk
Generate a synthetic control arm by propensity score modeling

- Goal: Estimate the survival distribution and hazard ratio for the CA209-214 regimen *had it be tested in CA209-920 population and compared to CA209-920 regimen*

- Achieved this through a weighted Kaplan-Meier estimate using average treatment effect in the treated (ATT) weights

- The ATT weights are a function of
  - The conditional probability of a subject selected randomly from the pooled population coming from the CA209-920 study, conditional on baseline covariates (*propensity score*)
  - The proportion of subjects in the pooled population who were in the CA209-920 study

\[
    w_i = \frac{\Pr(Z_i = 1|x_i)\Pr(Z_i)}{1 - \Pr(Z_i = 1|x_i)}
\]

- ATT weights weight subjects in CA209-214 relative to how similar their baseline covariates are to patients in the CA209-920 population

- The propensity scores were estimated using a random forest model (input features shown on slide 10)

- To assess how well the original and weighted CA209-214 populations compare to the CA209-920 population, plots comparing the weighted and unweighted mean differences were generated
After Weighting, Patients Characteristics in CA209-214 More Closely Align with CA209-920

Continuous Covariates

- Age
- Albumin
- Abs lymphocyte ct
- Bodyweight
- BL SLD
- BLD/ULN
- Number of lesions

Categorical Covariates

- KPS (lt80)
- KPS (80)
- KPS (gt80)
- SEX (male)
- MSKCC (poor)
- MSKCC (intermediate)
- MSKCC (favorable)
- PDL-1 (negative)
- PDL-1 (positive)
- PDL-1 (unknown)
- IMDC (poor)
- IMDC (intermediate)
- IMDC (favorable)

BLD/ULN: Baseline LDH divided by Upper Limit of Normal
KPS: Karnofsky Performance Status
MSKCC: Memorial Sloan Kettering Cancer Center
IMDC: International Metastatic RCC Database Consortium
Weighted OS of CA209-214 reflects higher baseline severity in the CA209-920 population

Due to higher baseline disease severity in CA209-920 population, overall survival of the weighted CA209214 N3I1 population is lower than when this same patient population is not weighted.

Weighting CA209214 N3I1 population decreases weight of the favorable risk subgroup which is very small in study 920. The weighting effect is stronger on the sunitinib arm as favorable risk patients have the best outcomes on Sunitinib.
PSM provided the benchmark distribution of OS on N3I1 and Sunitinib in a similar patient population in Cohort 1 of CA209-920 N6I1
Using early endpoints, i.e. limited to radiographic tumor size measurements, of the CA209-920 patient population in 2018 and a joint TGD-OS model, this model was able to reliably predict future overall survival in 2020.

PSM generated a synthetic control arm from historical data by adjusting the baseline characteristics to more closely resemble the baseline characteristics of the CA209-920.

This work established the feasibility of predicting OS in advanced RCC with longitudinal TGD and immature OS data together with baseline covariates, and benchmarking the predicted OS with that of an established therapy using PSM weighting.

Conclusion
Acknowledgement

• Akintunde Bello
• Satyendra Suryawanshi
• Yan Feng
• Iryna Shnitsar
• Kald Abdallah
• Heddy Bartell
• Sebastian Garrido
backups
2018 data cut vs. 2022 data cut

[Graph showing survival probability and proportion of subjects remaining in study for 2018 and 2020 data cuts]

Number at risk:
- 2018 data cut: 106, 97, 80, 65, 65, 40, 24, 12
- 2020 data cut: 106, 97, 80, 65, 65, 40, 24, 12

Time (months):
- 0, 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 42
CA209-920 (BMS-Sponsored NRDG) Phase 3b/4 Safety Trial of Nivolumab Combined with Ipilimumab in Subjects with Previously Untreated, Advanced or Metastatic RCC

Treated N=200

Cohort 1: ccRCC KPS ≥70% (n=100)
- Nivolumab 6 mg/kg IV plus Ipilimumab 1 mg/kg IV
- Q8 weeks alternating with nivolumab 480 mg IV Q8 weeks, staggered Q4 weeks

Cohort 2: Non-ccRCC, KPS ≥70% (n=50)
- Nivolumab 3 mg/kg IV combined with Ipilimumab 1 mg/kg IV
- Q3 weeks for 4 doses

Cohort 3: RCC (any histology), with nonactive brain mets, KPS ≥70% (n=25)
- Nivolumab 3 mg/kg IV combined with Ipilimumab 1 mg/kg IV
- Q3 weeks for 4 doses

Cohort 4: RCC (regardless of any histology or existing non-active brain metastasis or no), KPS 50%-60% (n=25)
- Nivolumab 3 mg/kg IV combined with Ipilimumab 1 mg/kg IV
- Q3 weeks for 4 doses

Advanced/met RCC
- Previously untreated in advanced or metastatic setting
- Tissue for PD-L1 testing

Screened N=250

Until progression, unacceptable toxicity, withdrawal of consent, or end of trial (2 years)

Enrollment completed:
- Cohort 1, alternative dosing n=106
- Cohort 2, non clear cell n=52
- Cohort 3, brain mets n=28
- Cohort 4, low KPS n=25

George DJ et al., BMJ Open. 2022 Sep 14;12(9)
CA209-214 (Registrational Trial in 1L RCC)

Treatment

Randomize 1:1

Arm A
3 mg/kg nivolumab IV + 1 mg/kg ipilimumab IV Q3W for four doses, then 3 mg/kg nivolumab IV Q2W

Arm B
50 mg sunitinib orally once daily for 4 weeks (6-week cycles)

Stratified by
- IMDC prognostic score (0 vs 1–2 vs 3–6)
- Region (US vs Canada/Europe vs Rest of World)

Treatment until progression or unacceptable toxicity

Motzer RJ et al., Cancer. 2022 Jun 1;128(11):2085-2097.
Schematic of Modeling Strategy

Ref: Metrum Research Group
Subpopulation 1 (fast TG):
\[ TB_i(t) = TB_0 \times e^{-TS_i t} + TG_i \times t \]

Subpopulation 2 (no-growth):
\[ TB_i(t) = TB_0 \times e^{-TS_i t} + TB_{SS_i} \]

Subpopulation 3 (intermediate TG and tumor shrinkage (TS)):
\[ TB_i(t) = TB_0 \times e^{-TS_i t} + TG_i \times t \]

where \( TB_i(t) \) is the TB at time \( t \) for the \( i^{th} \) patient, and \( TB_0_i \), \( TS_i \), and \( TG_i \) represent baseline TB, TS rate constant, and linear TG rate for the \( i^{th} \) patient, respectively.
Before PSM ATT Weighting

Subjects in CA209214 with dissimilar BL disease severity given lower ATT weights

Subjects in CA209214 with similar BL disease severity given higher ATT weights

After PSM ATT Weighting

• Propensity Score model

• ATT weighting

After PSM ATT weighting, weighted CA209214 population is more similar to CA209920 with respect to BL disease severity
2018 data cut vs. 2022 data cut (Tumor size)