

<b>Formula</b>	<b>Our Method</b>	<b>SRK/T</b>	<b>Holladay 1</b>	<b>Haigis</b>	<b>Hoffer Q</b>
<b>SRK/T</b>	<u>1.45e-18</u>	/	/	/	/
<b>Holladay 1</b>	<u>4.99e-14</u>	6.82e-01	/	/	/
<b>Haigis</b>	<u>5.10e-15</u>	1.00e+00	1.00e+00	/	/
<b>Hoffer Q</b>	<u>6.74e-24</u>	<u>4.15e-02</u>	<u>6.65e-09</u>	<u>1.39e-09</u>	/
<b>Barrett</b>	<u>5.76e-04</u>	<u>3.76e-14</u>	<u>1.61e-09</u>	<u>5.92e-09</u>	<u>1.06e-22</u>

**Table S1** The p-values from the post-hoc paired Wilcoxon tests, following the Friedman test, for the comparison of the testing set performance between methods. The p-values were adjusted with Bonferroni correction. The p-values between our method and the conventional methods were shown in the first column. All significant p-values (p-value < 0.05) were underscored. The Friedman test statistic was 176.73, and the associated p-value was 2.67e-36, which was statistically significant.