Do NFL Combine Stats Predict Rookie WR Success?

Costa Tsakopoulos, Charlie Vitner, Jackson Weber, Ava Weisheit, Tyler Wirges

Research Question

The NFL Draft Combine is one of the most publicized evaluation events in sports. Every spring, teams, media, and fans obsess over 40-yard dash times, vertical jumps, and shuttle drills, treating these numbers as indicators of who will become an impact player on Day 1. For wide receivers especially, the Combine is often framed as a predictor of early success, with the idea being that elite athleticism should translate directly into rookie-year production.

But teams also invest millions into scouting, film study, interviews, and analytics because they know the Combine only captures a small slice of who a player is. And despite how much attention these metrics get, there is still a real question about whether they actually matter once a rookie steps onto an NFL field.

From a team and coaching perspective, this matters because draft decisions shape roster construction, cap strategy, and long-term development plans. If Combine metrics do not predict rookie performance, teams may be overweighting the wrong information. For players, the stakes are just as high. Their draft position affects contract value, opportunity, and perception. If Combine numbers are not strongly tied to early success, players might benefit more from focusing on development and scheme fit rather than chasing marginal improvements in testing drills. For the business side and fans, the Combine drives media narratives, betting markets, and draft-day expectations. Understanding what actually predicts rookie performance helps cut through the hype and gives a clearer picture of how teams should evaluate talent.

That brings us to our core research question: Are NFL Combine metrics actually correlated with rookie wide receiver performance, or are other factors like draft position or team context more meaningful?

Methods

All data comes from Pro-Football-Reference.com, which is widely considered the premier free online source for NFL statistics. It is highly reliable for box scores, career stats, and historical data, and is commonly used by both researchers and writers.

We used three years of data (2023-2025). Combine data for each year was uploaded into its own tab and required some cleaning. Heights had been converted to dates on import, and the drafted team, round, and pick were stored in a single column that needed to be separated. Name formatting presented another challenge. The WR statistics page omitted suffixes like "Jr." and "III," and in one case Nathaniel Dell Jr. was listed under his nickname "Tank." We

cross-referenced names manually, using both the data and our own football knowledge to ensure every player who should have had statistics was correctly matched. To capture rookie performance, we pulled statistics from every player in the year they were drafted and used an XLOOKUP to match rookie WRs in our draft tabs to the comprehensive WR statistics page.

To create a single measurable outcome, we built a composite Score as our target variable using a fantasy scoring structure: 1 point per reception, 0.1 points per yard, and 6 points per touchdown. This gave us one continuous variable to predict, making correlation and regression analysis more manageable.

We briefly explored adjusting Score for quarterback quality by adding a TeamQB tab with each team's end-of-season ranking, primary QB, and that QB's QBR. However, the correlation between Score and QBR was r = 0.099, and between Score and team rank was r = -0.029, both essentially zero. There was no meaningful relationship between quarterback environment and rookie WR performance in our dataset, so we kept the raw Score as our target variable.

From there, our analysis had two goals: to determine whether Draft Combine metrics predict a contributing rookie season at all, and to identify whether anything predicts performance better than raw athleticism. We used a correlation matrix to examine relationships between each Draft Combine metric and Score, and to flag multicollinearity between predictors. We then used multiple regression to quantify the individual contribution of each significant predictor while

Findings

Correlation analysis was conducted to evaluate the relationship between NFL Combine metrics and rookie wide receiver performance, measured by a composite Score variable. The results indicate that most Draft Combine metrics have weak linear relationships with rookie performance, with correlation coefficients generally ranging between -0.20 and +0.20. For example, the 40-yard dash (r = -0.11), vertical jump (r = 0.17), and broad jump (r = 0.11) show minimal association with performance, indicating limited predictive ability.

Agility-based metrics show stronger relationships relative to other Combine variables. The 3-cone drill demonstrates a moderate negative correlation with Score (r = -0.50), and shuttle time shows a weaker but still notable negative relationship (r = -0.32). These relationships indicate that lower times in agility drills are associated with higher rookie performance.

The strongest relationship observed in the dataset is between draft position (Pick) and rookie performance (r = -0.53), which exceeds the magnitude of all combine metrics. This indicates that earlier draft selection is more strongly associated with higher rookie production than any individual athletic measurement. Additional variables including quarterback quality (QBR) and team rank show near-zero correlations with rookie performance (r = 0.099 and r = -0.029 respectively), indicating no meaningful linear relationship within this dataset.

To further investigate these relationships, we ran a series of multiple regression models using our composite Score as the target variable, building each model progressively to isolate the strongest predictors and maximize sample size.

Our first model (Figure 2) included all available combine metrics: height, weight, 40-yard dash, vertical jump, broad jump, shuttle, and draft pick. While this model produced the highest raw R^2 of 0.569, it was limited to only 21 observations due to missing combine data. With 7 predictors and only 21 players, the model was too small to be reliable, and the Adjusted R^2 of 0.337 reflected this. The overall model was not statistically significant (Significance F = 0.077).

To increase our sample size, we removed shuttle since it had the most missing data (Figure 3). This brought our observations up to 57. R^2 dropped to 0.369 and Adjusted R^2 to 0.294, suggesting shuttle had been inflating the previous model. Draft pick was the only statistically significant predictor in this model (p = 4.46E-06), while all combine metrics remained insignificant.

We then ran a regression using only combine-based physical metrics, removing draft pick entirely to isolate the predictive power of athleticism alone (Figure 4). With 67 observations, this model produced an R^2 of only 0.048 and an Adjusted R^2 of -0.030, meaning combine metrics alone explained essentially none of the variance in rookie performance. No predictor was statistically significant and the overall model was not meaningful (Significance F = 0.686). Removing height, the highest p-value predictor, did not improve results (Figure 5). R^2 remained at 0.048 and Adjusted R^2 stayed negative at -0.013.

Our final and strongest model used draft pick as the sole predictor (Figure 6). With 77 observations, this model produced an R^2 of 0.279 and an Adjusted R^2 of 0.270, with a p-value of 7.79E-07. Draft pick was highly statistically significant, and the minimal gap between R^2 and Adjusted R^2 indicates no overfitting. This model outperformed every other configuration on both reliability and statistical validity.

Overall, NFL Draft Combine metrics have limited explanatory power in predicting rookie wide receiver performance, while draft position demonstrates the strongest relationship among the variables analyzed.

Implications

Our findings suggest that rookie wide receiver performance is far less dependent on external factors like QB play or team quality than originally believed. The near-zero correlations between Score and both QBR and team rank show no meaningful relationship between a player's early production and the situation they are drafted into. Most of the variation seen in rookie performances appears to be driven by the players' own development, opportunity, and adaptation

to the NFL level. This tells us that attempting to adjust for team or QB context adds noise rather than improving predictive accuracy.

These findings are useful across the board. For coaches, they suggest focusing on individual player development rather than linking outcomes to outside factors. For players, the data supports the idea that landing in the best developmental environment matters more than landing on the best team. For front offices, the findings support a more disciplined, talent-focused evaluation process that avoids over-adjusting for situational factors that the data shows do not meaningfully predict rookie performance.

Conclusions and Suggestions

Our analysis shows that NFL Draft Combine metrics alone are poor predictors of rookie wide receiver performance. When used without draft pick position, combine metrics explained essentially none of the variance in rookie Score ($R^2 = 0.048$). Draft pick position was the strongest predictor in our dataset, explaining approximately 27% of rookie WR variance ($R^2 = 0.279$, $p = 7.79E-07$). For every spot earlier in the draft, a rookie WR's Score increases by approximately 0.568 points. Even so, 73% of rookie WR performance remains unexplained, pointing to factors like scheme fit, coaching, and football IQ that combine drills do not capture.

NFL teams and coaches should use Draft Combine metrics to supplement scouting rather than drive player selection. Coaching staffs should be patient with rookies and build individualized development plans, understanding that natural performance fluctuation is expected and that different players adjust to the NFL on different timelines. Organizations should also be cautious about overvaluing prospects who benefited in college from elite systems or quarterback play, as those advantages will not carry over.

We recommend that teams also hold private workouts and attend collegiate pro days for players they are seriously considering. With high-profile talent increasingly declining combine invitations, these metrics may be most applicable to later-round prospects and undrafted free agents. Future research should incorporate more years of data, and categorizing wide receivers by archetype and offensive scheme fit could reveal more nuanced relationships that our current model was unable to capture.

Tables and Figures

	Ht.in	*Wt*	*40yd*	*Vertical*	*Broad Jum*	*3Cone*	*Shuttle*	*Rec*	*Yds*	*TD*	*Score*	*Pick*
Ht.in	1											
Wt	0.342963	1										
40yd	-0.0093	0.294359	1									
Vertical	-0.18493	0.024148	-0.18374	1								
Broad Jum	-0.13135	0.007711	-0.16358	0.622448	1							
3Cone	0.157204	0.159281	0.277202	-0.37504	-0.29678	1						
Shuttle	-0.15504	-0.24726	0.227558	-0.29986	-0.17143	0.684529	1					
Rec	0.016906	0.03004	-0.14218	0.200582	0.152588	-0.5423	-0.26803	1				
Yds	-0.00608	0.049532	-0.10833	0.184298	0.127928	-0.47777	-0.28815	0.974613	1			
TD	0.014942	0.035718	-0.05543	0.069936	-0.00283	-0.31476	-0.34552	0.764713	0.787615	1		
Score	0.006509	0.041706	-0.11296	0.171935	0.111006	-0.50132	-0.32209	0.979567	0.9875	0.861659	1	
Pick	-0.23025	-0.1752	0.323666	-0.30383	-0.21346	0.484619	0.466955	-0.53153	-0.53376	-0.41286	-0.52852	1

Figure 1 - Correlation table

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.75444259
R Square	0.56918363
Adjusted R S	0.33720558
Standard Err	44.9229025
Observations	21

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	7	34660.8468	4951.54955	2.45360989	0.07704843
Residual	13	26234.8732	2018.06717		
Total	20	60895.72			

	Coefficients	*Standard Erro*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-1555.122	1196.07983	-1.3001825	0.21611802	-4139.0954	1028.85135	-4139.0954	1028.85135
Ht.in	0.96816226	5.29770643	0.18275121	0.85781246	-10.476837	12.4131612	-10.476837	12.4131612
Wt	-0.7592303	0.82735348	-0.9176614	0.37550756	-2.5466189	1.02815819	-2.5466189	1.02815819
40yd	324.307317	184.516955	1.75760172	0.10232233	-74.317328	722.931962	-74.317328	722.931962
Vertical	-0.9226628	5.59495502	-0.1649098	0.87155168	-13.009828	11.1645027	-13.009828	11.1645027
Broad Jump	1.31679563	2.70956035	0.48598129	0.63506944	-4.5368536	7.17044489	-4.5368536	7.17044489
Shuttle	51.2197244	109.636806	0.46717636	0.64810735	-185.6362	288.075644	-185.6362	288.075644
pick	-0.7592475	0.19986247	-3.7988499	0.00221216	-1.1910241	-0.3274709	-1.1910241	-0.3274709

Figure 2 - Full regression : All variables (excluding Broad Jump & 3 Cone)

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.60766506
R Square	0.36925683
Adjusted R Sc	0.29356765
Standard Erro	57.5717546
Observations	57

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	6	97020.8128	16170.1355	4.87859457	0.00054084
Residual	50	165725.346	3314.50692		
Total	56	262746.159			

	Coefficients	*Standard Erro*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	95.7273749	618.563522	0.15475755	0.87763549	-1146.694	1338.14877	-1146.694	1338.14877
Ht.in	-4.4178035	4.12585982	-1.0707595	0.28941921	-12.704837	3.86922979	-12.704837	3.86922979
Wt	-0.0956543	0.60654568	-0.1577034	0.87532567	-1.3139372	1.12262852	-1.3139372	1.12262852
40yd	111.890696	105.185105	1.06375038	0.29255234	-99.379806	323.161197	-99.379806	323.161197
Vertical	-0.962056	3.84559724	-0.2501708	0.80348037	-8.6861654	6.76205337	-8.6861654	6.76205337
Broad Jump	-0.5351796	2.05993401	-0.2598042	0.79608161	-4.6726788	3.60231963	-4.6726788	3.60231963
Pick	-0.6778255	0.13168238	-5.1474276	4.4588E-06	-0.9423173	-0.4133337	-0.9423173	-0.4133337

Figure 3 - Regression : No Shuttle

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.21965274
R Square	0.04824733
Adjusted R Sc	-0.0297652
Standard Erro	66.8796953
Observations	67

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	5	13831.4447	2766.28894	0.61845623	0.68616945
Residual	61	272846.512	4472.89364		
Total	66	286677.957			

	Coefficients	*Standard Erro*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	191.13571	627.245578	0.30472229	0.76161491	-1063.1193	1445.39076	-1063.1193	1445.39076
Ht.in	0.55330773	4.28047733	0.12926309	0.89757476	-8.0060357	9.11265116	-8.0060357	9.11265116
Wt	0.21501652	0.64261839	0.33459441	0.73907884	-1.0699784	1.5000114	-1.0699784	1.5000114
40yd	-88.465642	100.399715	-0.8811344	0.38170376	-289.22728	112.295996	-289.22728	112.295996
Vertical	3.4799181	4.01053685	0.86769383	0.3889627	-4.539646	11.4994822	-4.539646	11.4994822
Broad Jump	0.46951271	2.20872679	0.21257166	0.83236984	-3.9471095	4.8861349	-3.9471095	4.8861349

Figure 4 : Regression : Combine Metrics Only

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.21905849
R Square	0.04798662
Adjusted R Sc	-0.0134336
Standard Erro	66.3472353
Observations	67

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	13756.7074	3439.17685	0.78128385	0.54164227
Residual	62	272921.249	4401.95563		
Total	66	286677.957			

	Coefficients	*Standard Erro*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	239.719019	498.172371	0.48119694	0.63207049	-756.11336	1235.5514	-756.11336	1235.5514
Wt	0.25213569	0.57031225	0.44210113	0.65995437	-0.8879023	1.39217364	-0.8879023	1.39217364
40yd	-90.944015	97.7673761	-0.9302082	0.35587114	-286.37822	104.490185	-286.37822	104.490185
Vertical	3.38260034	3.90787629	0.86558532	0.39005472	-4.4291331	11.1943337	-4.4291331	11.1943337
Broad Jump	0.46247345	2.19047602	0.2111292	0.83347888	-3.9162257	4.84117263	-3.9162257	4.84117263

Figure 5 - Regression : Combine Metrics minus Height

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.52851627
R Square	0.27932945
Adjusted R Sc	0.26972051
Standard Erro	64.2533359
Observations	77

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	120014.185	120014.185	29.0697448	7.7895E-07
Residual	75	309636.838	4128.49117		
Total	76	429651.023			

	Coefficients	*Standard Erro*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	142.517908	13.3433934	10.6807844	1.0226E-16	115.936505	169.09931	115.936505	169.09931
Pick	-0.5680789	0.10536297	-5.3916366	7.7895E-07	-0.7779727	-0.3581851	-0.7779727	-0.3581851

Figure 6 - Regression : Pick Only

Data Sources

Combine Data : https://www.pro-football-reference.com/draft/2025-combine.htm

Performance Data : https://www.pro-football-reference.com/years/2025/receiving.htm

Team Data : https://www.nfl.com/standings/league/2025/reg

QB Data : https://www.pro-football-reference.com/years/2025/passing.htm