Improving Outcomes in a Home Health / Chronic Care Settings

Enabling an "Aging-in-Place" Strategy

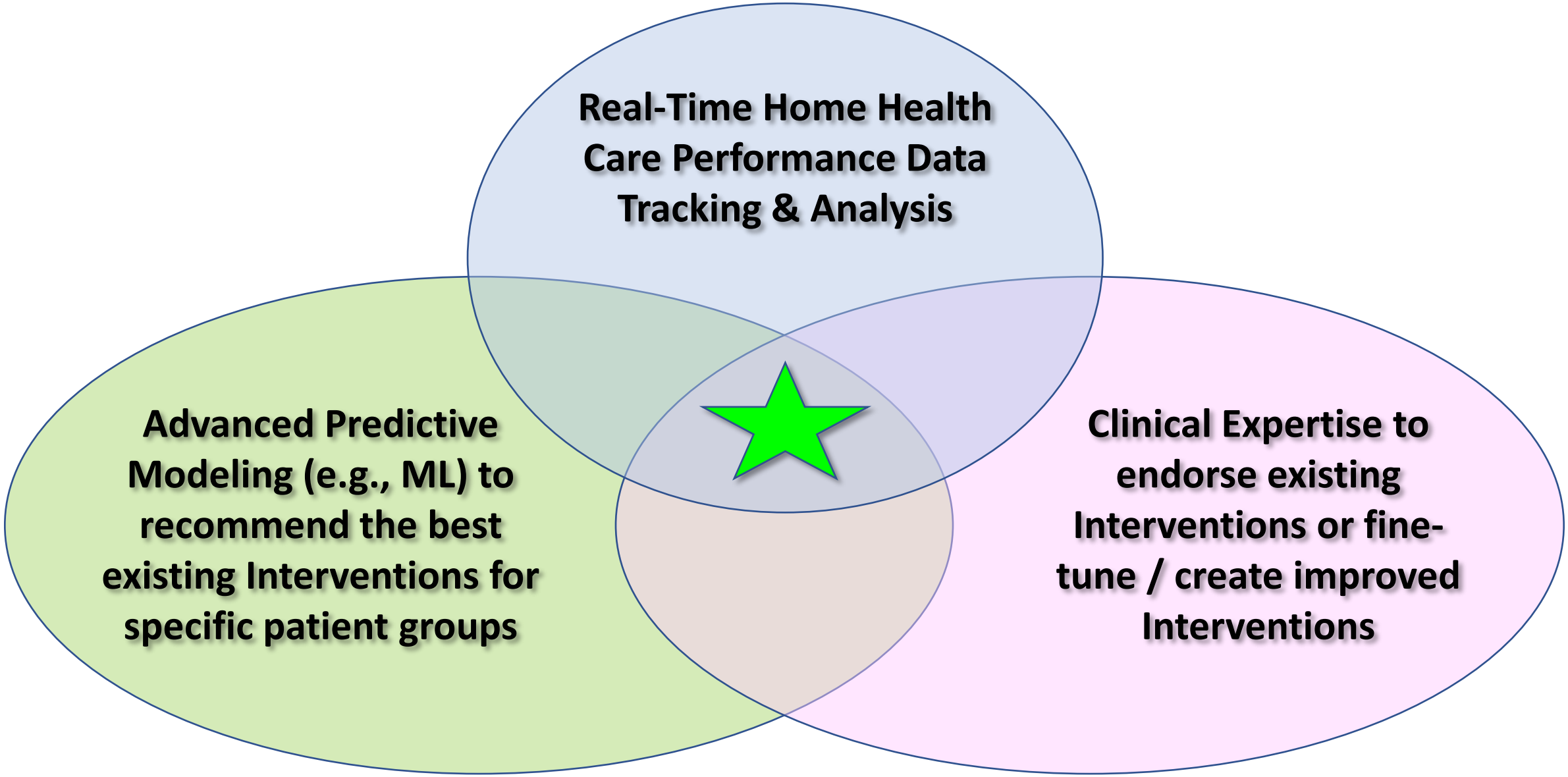
Technical Research & Analysis by:

David J Patrishkoff

Data Provided by:

Riversoft in Melbourne, Florida

3 Coordinated Activities can Revolutionize the Quality of Home Health & Chronic Care Services



RiverSoft

- Started in 1997 after Skip McCoy (founder and president) developed an early 1990s precursor system for Interim Healthcare which is the oldest and largest franchisor of home health in the USA.
- Designed to work for large agencies with thousands of patients in multiple locations with a wide variety of payer requirements
  - Medicare / Medicare Advantage / Medicaid / VA
  - Commercial insurance, HMOs, self-pay
- Complete office suite for home health and hospice agencies
  - Employee / staff management, payroll, and scheduling
  - Patient management including complete medical records
  - Complete configurable EMR
  - Complete billing management: invoices / claims / accounts receivable
  - Reporting
- Industrial strength software that is customized through configuration and software changes to meet unique demands of client agencies.
  - Quick handling of agency requested customized changes, preferences, and access.
  - Quick turnaround of ~20 updates to system each month
  - Unique “beta” testing of changes for small set of agencies that lead to updates for all.

Why use RiverSoft and its EMR data?

- EMR includes complete tracking system of clinical pathways with their own subsets of interventions.
  - All home health systems that are CMS compliant they track OASIS (Outcome and Assessment Information Set) outcomes although to our knowledge none incorporate relational database for tracking the clinical pathway interventions meant to drive improve in OASIS outcomes.
  - EMR continually checks compliance to schedules, authorizations, plan of care, physician orders.
  - All interventions are electronically tracked as to goals, progress toward goals, and provide voice to text narration of notes.
- RiverSoft has clients in over 20 states from which it can automatically gather and de-identify clinical performance data related to clinical pathway performance by intervention relative to OASIS outcomes (ADLs, risk of hospitalization, falls, ER visits, etc.).

Patient Driven Grouping Model (PDGM) – Key Factors

- Accurate Coding of Diagnoses
- Complete and Accurate OASIS assessment as part of “start of care”
- Complete documentation of interventions, referrals, nursing notes, progress tracking.
- Timely filing of OASIS
- Physician timely sign off and incorporation of verbal orders
- Technology to automate many systems
  - Some states require electronic visit verification system to be in use (ELVIS from RiverSoft)
  - Demonstration of whether interventions met or not met goals and associated narration.

Introduction to Activities of Daily Life Scores (ADLs)

(Lower Scores = More Independence)

Higher Quality of Life

- **M1800 Grooming:** **Current ability to tend safely to personal hygiene needs**
- **M1810 Dress Upper Body:** **Current Ability to Dress Upper Body safely** (with or without dressing aids)
- **M1820 Dress Lower Body:** **Current Ability to Dress Lower Body safely** (with or without dressing aids)
- **M1830 Bathing:** **Current ability to wash entire body safely**
- **M1840 Toilet Transferring:** **Current ability to get to and from the toilet or bedside commode safely** and transfer on and off toilet/commode
- **M1850 Transferring:** **Current ability to move safely from bed to chair**, or ability to turn and position self in bed if patient is bedfast
- **M1860 Ambulation/Locomotion:** **Current ability to walk safely, once in a standing position**, or use a wheelchair, once in a seated position, on a variety of surfaces

The Opportunity for Machine Learning (ML) Algorithm Applications in Home Health / Chronic Care

Historically: The Wild West

Individual instincts & impulse-driven decision-making

Instincts & Basic Statistics >

< Application of Local wisdom versus nationwide evidence-based recommendations

Clinical inputs to improve the list of endorsed interventions >

< ML & Predictive Modeling

Optimized Algorithms updated with monthly real-time performances for nationwide application of interventions

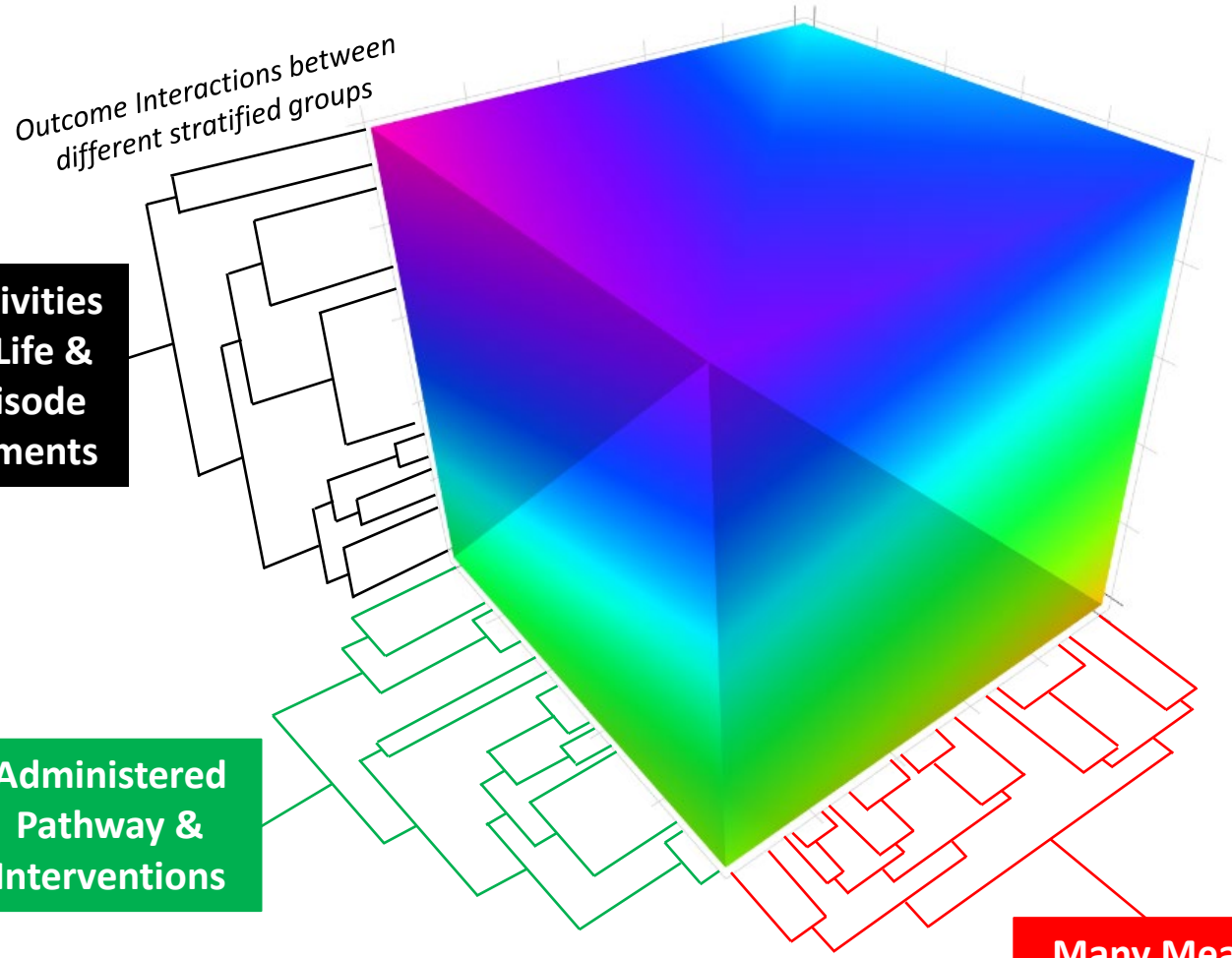
Self-Improving ML Algorithms track what interventions work well and which ones do not

We Have 5 Major Analytical and Clinical Research Questions Concerning Home Health Care

Major Research Question #1:

How do various factors and their interactions affect Home Health Care outcomes?

All shown clustering and stratified group interactions are purely hypothetical and will be updated when the actual results and known

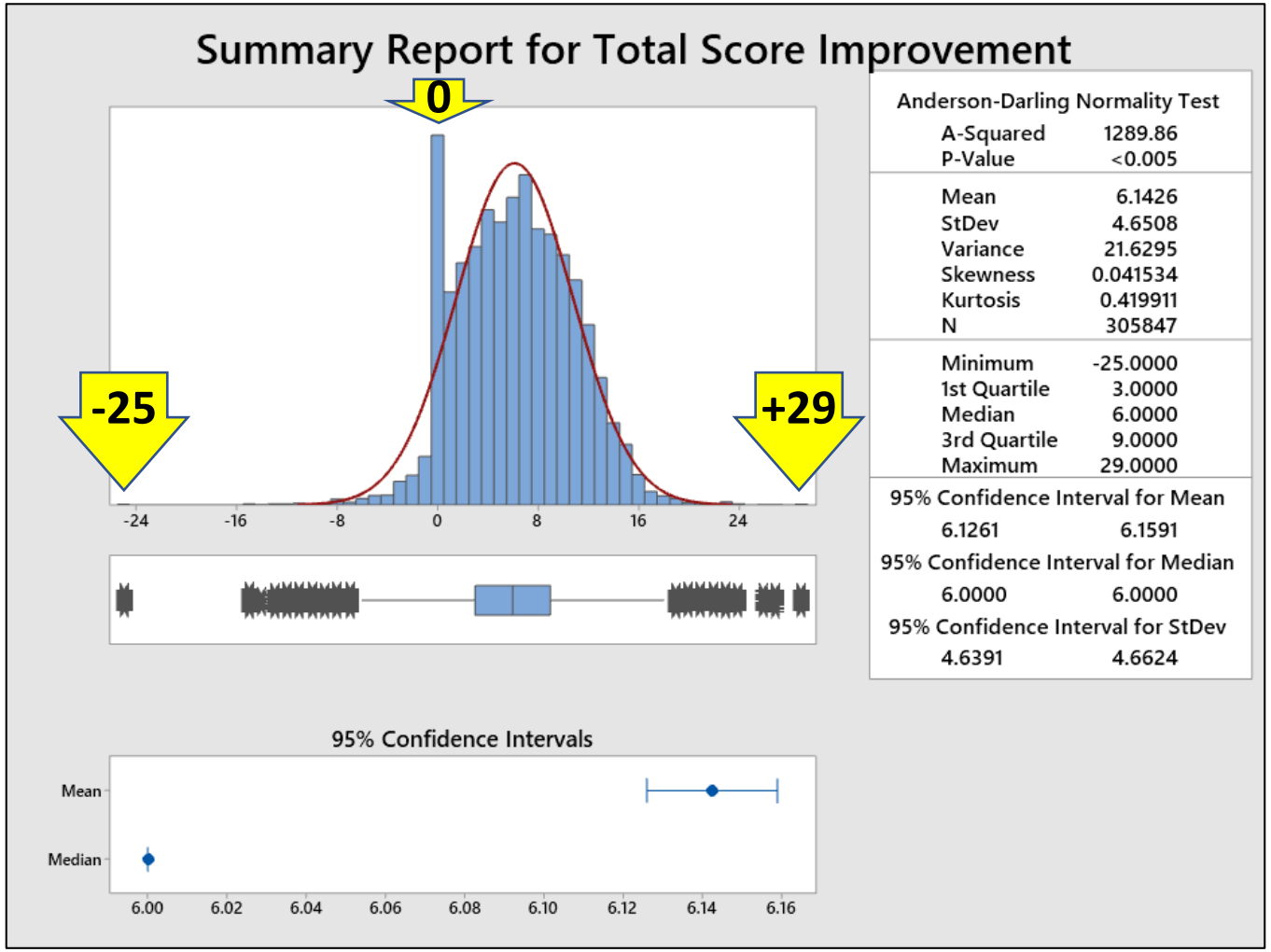


Analysis Factors (in alpha order):

1. Activities of Daily Life (ADL) Outcomes (total scores) - start, end & improvement scores
2. Age groups of patient
3. Home healthcare Agency code
4. Clinical groups (12), but R is absent until 2020
5. Comorbidity levels (3)
6. Days from Start of Care to Death (not much data)
7. Demographics and Race of patient
8. Diagnosis codes (43,000) - primary and secondary
9. # of Diagnosis codes (#) per patient
10. Employee Skill level
11. Fall Risks
12. Functional groups (L, M & H) as a starting assessment of patient limitation
13. Hospitalization Risks
14. Length of stay (end of intervention care minus start of care)
15. Medications used (from a list of 83,000 possibilities)
16. Medication interactions risks
17. Obese (Y/N)
18. Pathway and intervention code groups (1200+)
19. # of Pathway and interventions per patient
20. Prescribed Prescription drugs (83,000+)
21. Positive, Zero & Negative Improvement outcomes
22. Same employee at start and end of intervention (Y/N)
23. Skill Level of care-giver
24. Smoking + obesity group (Y/N)
25. Smoking (Y/N)
26. Sate where patient lives
27. Year for start of care (Pre-2017, 2017, 2018 & 2019)

Major Research Question #2:

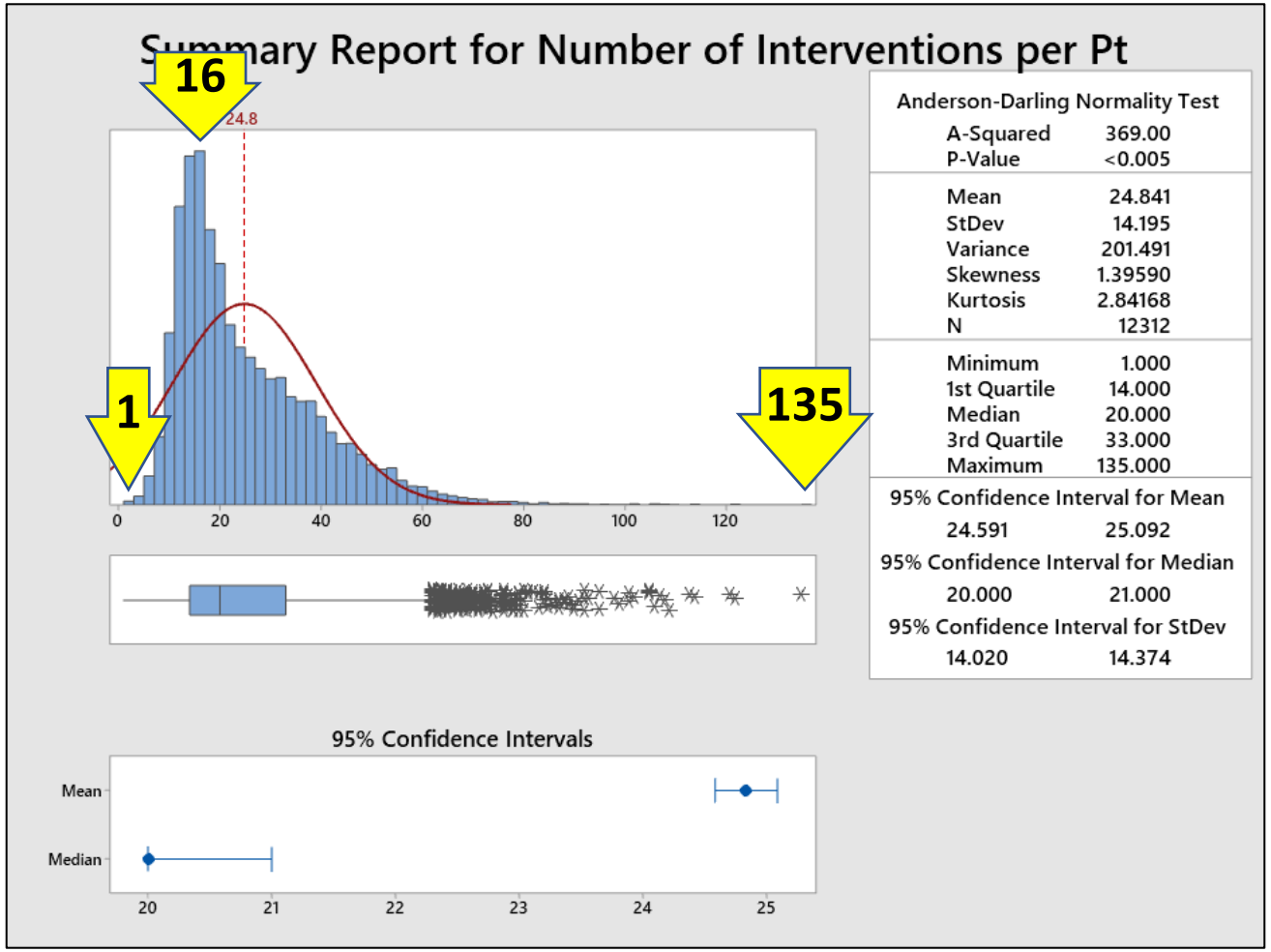
Why is there so much ***variation*** for Home Care ADL improvement outcomes?



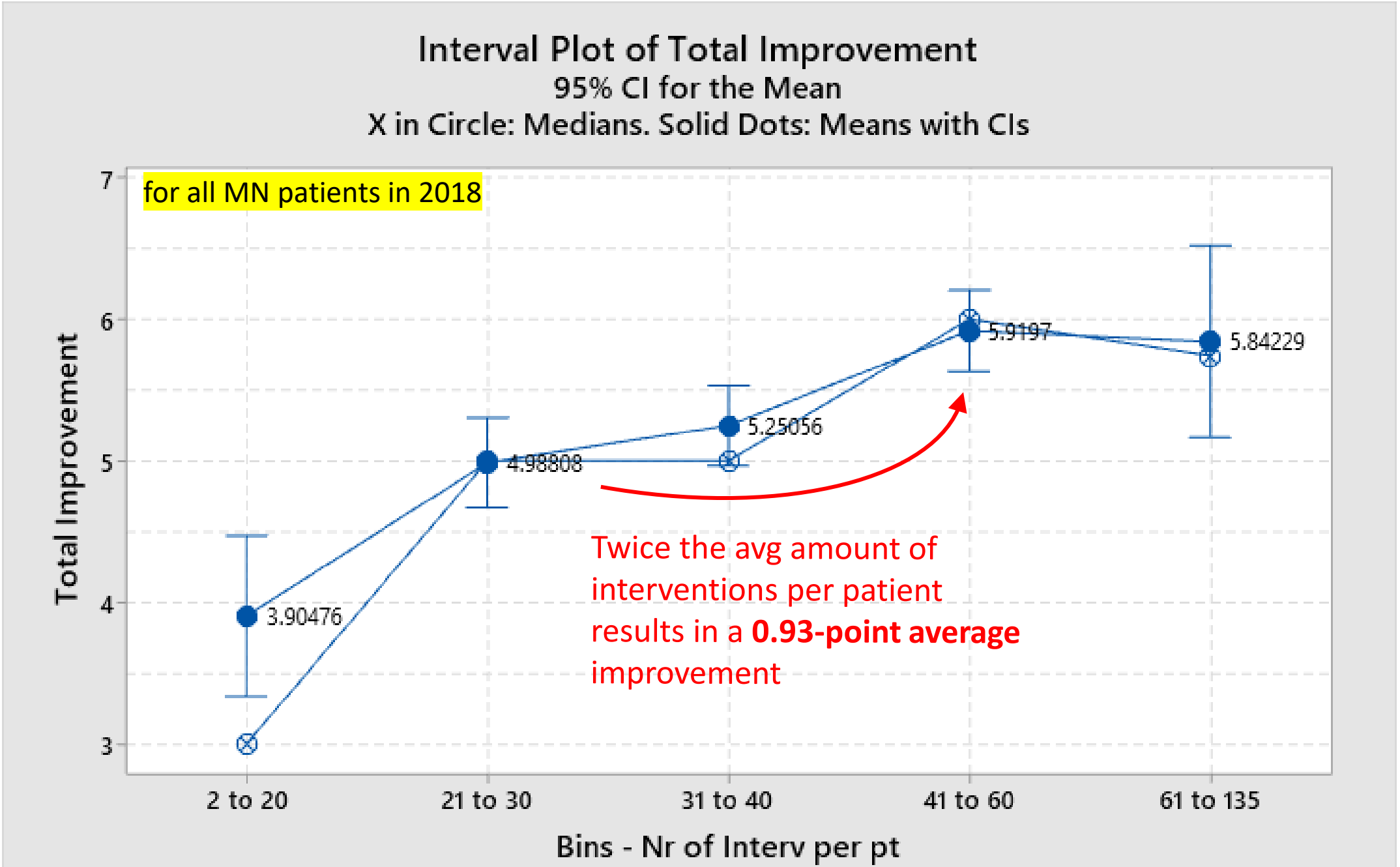
MN, NC & KS data: from 2017 to present-day (n = 305,849 interventions):
0 was the most common level of ADL improvement for home health care outcomes (n = 28,412 = 9.3%) with a range between **-25** & **+29**

Major Research Question #3:

How does the range of Interventions per patient affect the ADL Outcome Score?



MN, NC & KS data: from 2017 to present-day:
The # of interventions per patient range between **1** and **135**



Individual standard deviations are used to calculate the intervals.

Major Research Question #4:

What can we learn from the Best of the Best (BOBs), Worst of the Worst (WOWs) and Duds Interventions?

Example for MN:

- **1,259** different pathway & interventions combinations were applied over the past 3 years
  - **157** of them *always* attained an improvement of 5 or more (BOBs)
  - **90** of them *always* achieved an improvement between +1 and +4
  - **74** of them *always* attained a Zero improvement (Duds)
  - **14** of them *always* had a negative outcome (WOWs)

  - The remaining **924** interventions had very mixed results

Major Research Question #5:

How can clinical experts help us to:

1. Endorse continued use of strong existing interventions
2. Endorse the Deletion or replacement of weak interventions
3. Improve existing interventions
4. Create better interventions

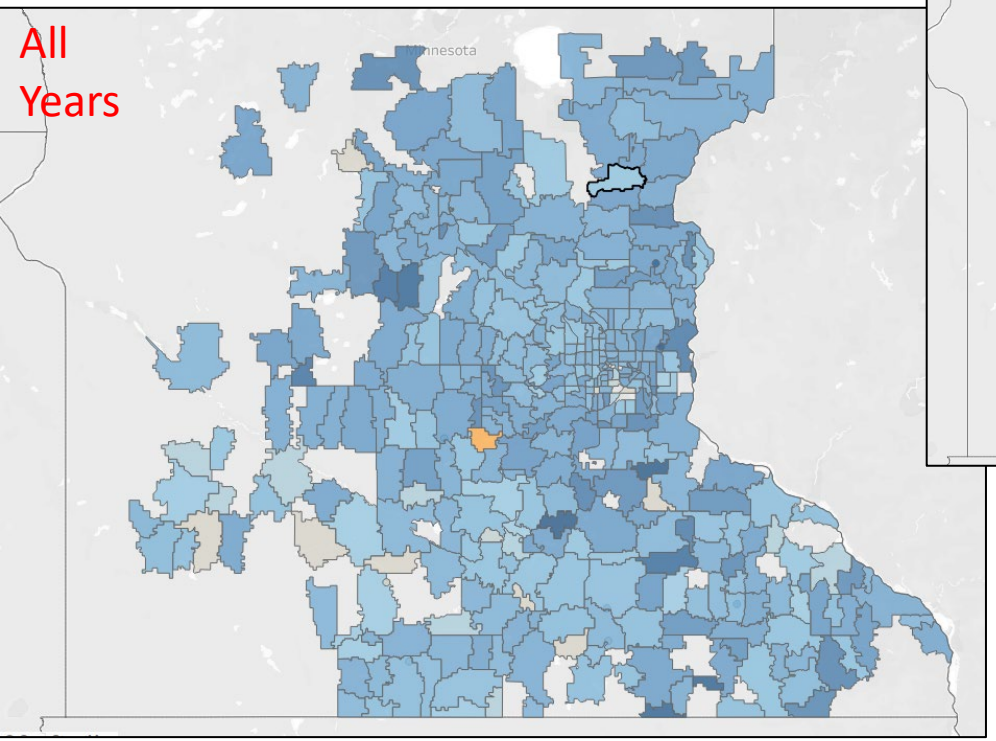
Unexplored Research Questions:

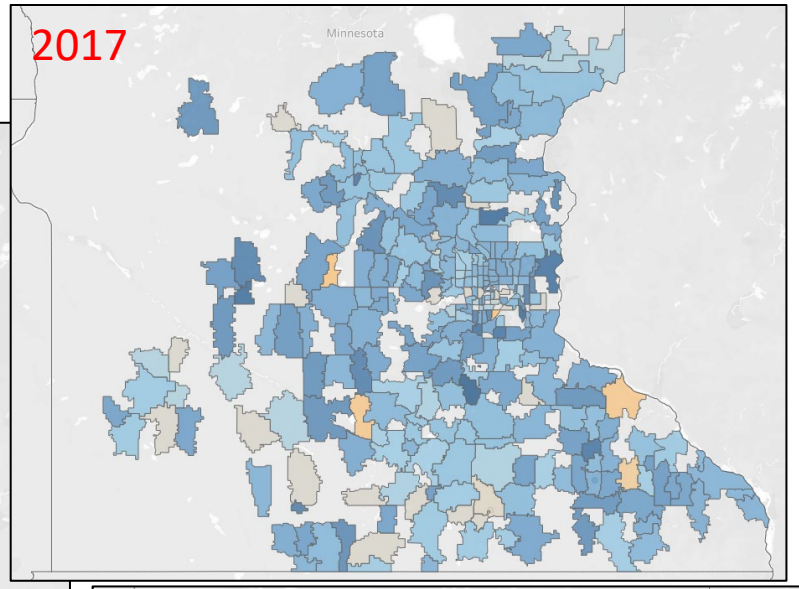
- What is the impact of Pharmaceuticals on:
    1. ADL Scores
    2. Fall Frequency
    3. ER Visits
    4. Unexpected complications

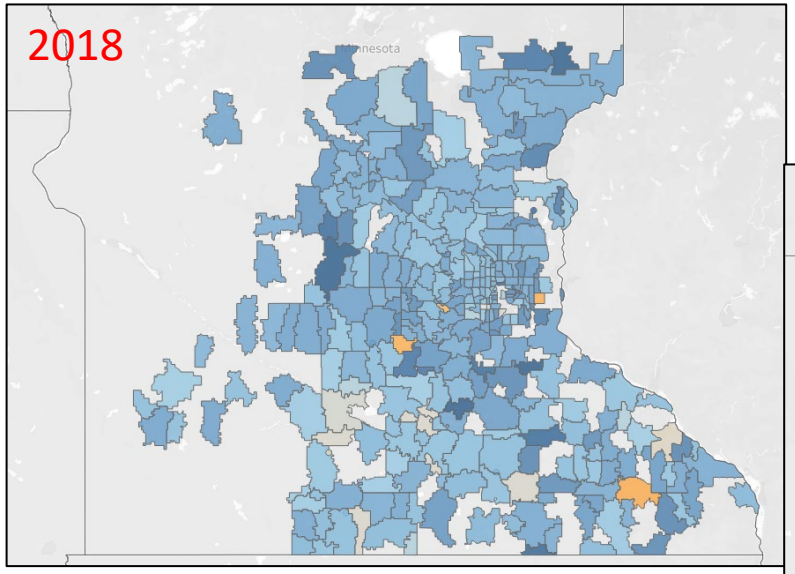
- Unrealized opportunities for tele-medicine and remote patient monitoring

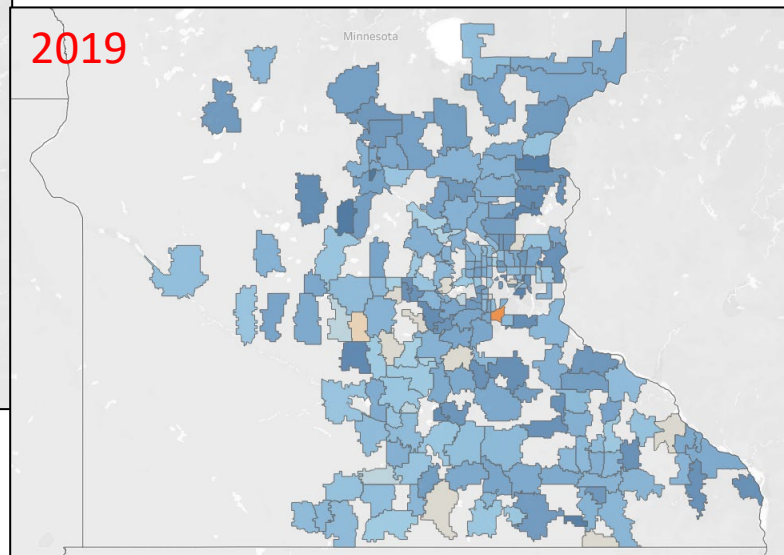
Geo-Spatial Analysis of MN data

Average Improvement scores – all years - darker is better. Orange is 0 or negative

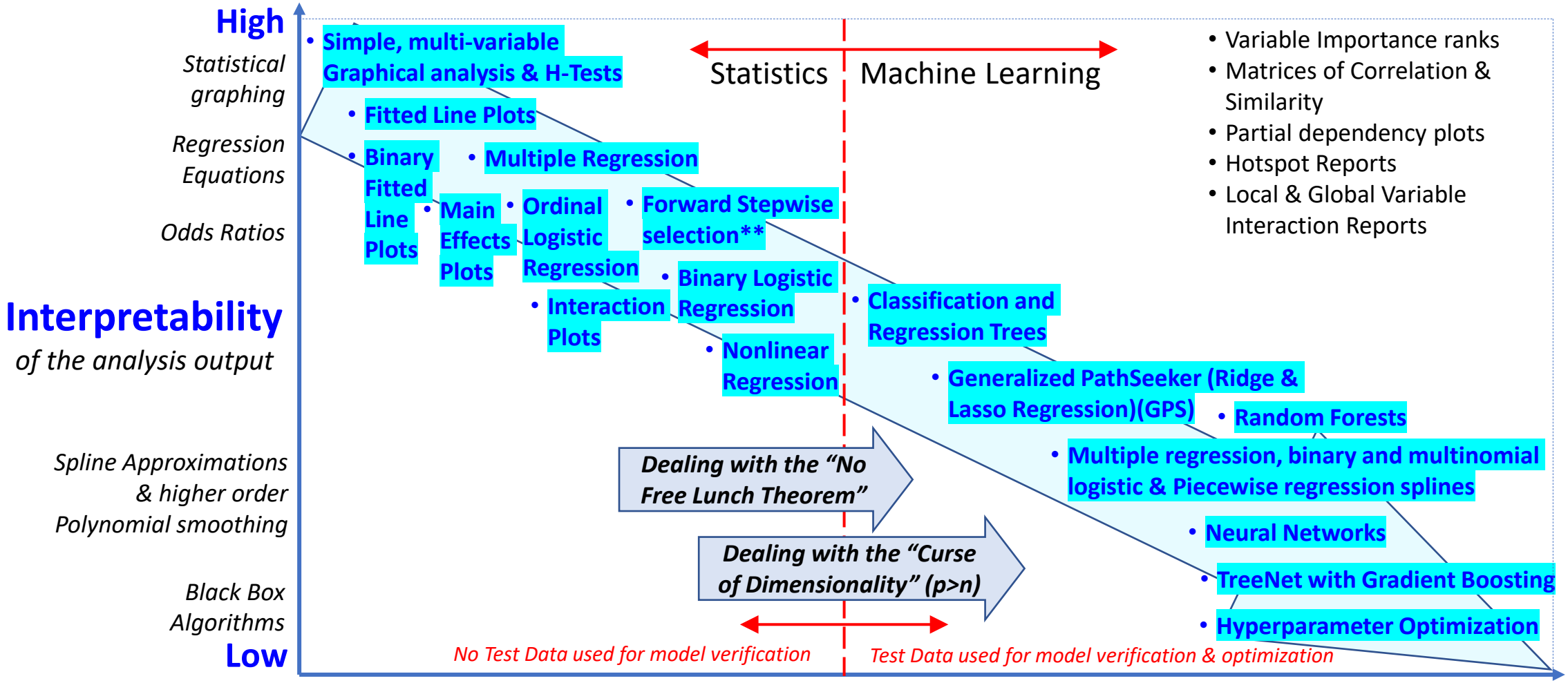








Various Methods of Analysis Applied



Main Analysis techniques are in ***Blue Highlights***
Supporting analysis features are in ***black font***

- James, G., et.l. (2013), If *n (# of observations)* is not much larger than *p* (# of predictors), then there can be a lot of variability in the least squares fit, resulting in overfitting and consequently poor predictions on future observations not used in model training. Page 204

** Unlike Best Subset reg that is limited to 40 p max, and backward stepwise cannot deal with p>n, forward stepwise can be used even when $n < p$, and so is the only viable subset method when *p* is very large. Pg 208

The Machine Learning *No-Free-Lunch* Theorem

No one learning algorithm dominates all others over all possible data sets. The best algorithms are customized, fine-tuned and optimized for each data set and its derivative.

Rank of 17 Home Health Care Factors

Factor	Mean Difference in Total Score Between High and Low Group	Known information up Front?
Pathway and Intervention code groups (1200+)	30	No, but we can predict the best P&I codes immediately based on the 10 known "Yes" factors
Patient Zip Codes for geo-mapping spatial analysis (379)	19	Yes, but use just for geo-mapping
Start of Care Total Score Groups (31)	8.06	Yes
Discharge Disposition	6.67	No
Clinical groups (12), but R is absent until 2020	4.48	Yes
Employee Skill Level (6)	3.81	No
Agencies in MN	3.28	Yes, but we will not include this factor
Functional groups (3)	2.27	Yes
# of Pathway and interventions Groups per patient	2.21	No
Year for start of care (2017, 2018 & 2019)	1.83	Yes, but we will only use the latest 2-18 months of data for recommendations
Age groups (6)	1.4	Yes
LOS (end of intervention care minus start of care)	1.14	No
Comorbidity levels (3)	0.97	Yes
Same employee at start and end of intervention (Y/N)	0.55	No
Smoking (Y/N)	0.54	Yes
Smoking + Obesity (Y/N)	0.46	Yes
Obesity (Y/N)	0.18	Yes

Benchmarking Various ML Methods to Determine the Variable Importance Ranks

		Predictor Importance Ranking Methods									
		Mean Diff between high and Low Group	**Median** Diff between high and Low Group	**SPM** Random Forest Regression Rank	**SPM** CART Regression Rank	**SPM** TreeNet Regression Rank	**SPM** Ridge & Lasso Regression Quasi-Rank	**JMP** Neural Predictive Modeling Rank	**SPM** TreeNet Classification Regression Rank	**SPM** CART Classification Rank	
Rank of Predictors in their ability to improve Total Outcome Scores >>				3.36	3.42	3.81	4	4.35			RMSE
				2.44	2.5	2.87	3.17				MAD
				11.3	11.7	14.5	16				MSE
									0.36	0.37	Balanced Error Rate
Predictors of Improvement available at the beginning of care:											
1	Start of Care Total Score Groups (31)	8.06	7	0.51	100	100	100	0.48	86.7	17.3	16-19 days is best and 0-4 days is worst
2	Clinical groups (12), but R is absent until 2020	4.48	6	0.43	44.9	92.6	47	0.36	100	100	MS Rehab is best & Complex Nursing is wors
3	Functional groups (3)	2.27	2	0.27	22	37	0	0.21	37	37.4	Group B (mid) is best & A (Low limitations) is
4	Age groups (6)	1.4	1	0.36	38.4	76.1	21	0.11	86.1	22.8	65-74 YO is best & 95+ YO is worst
5	Comorbidity levels (3)	0.97	2	0.24	30.9	51.3	0	0.01	58.7	20.9	Level 1 is best & 3 is worst
6	Smoking (Y/N)	0.54	1	0.06	10.7	28.8	23	0.02	18	4.2	Smokers improve more
7	Smoking + Obesity (Y/N)	0.46	0	0.01	6.73	25.7	21	0.02	16	2.9	Non-Obese + Non-smokers improve more
8	Obesity (Y/N)	0.18	0	0.13	12.2	34.4	0.4	0.12	33	10.2	Obese improve more

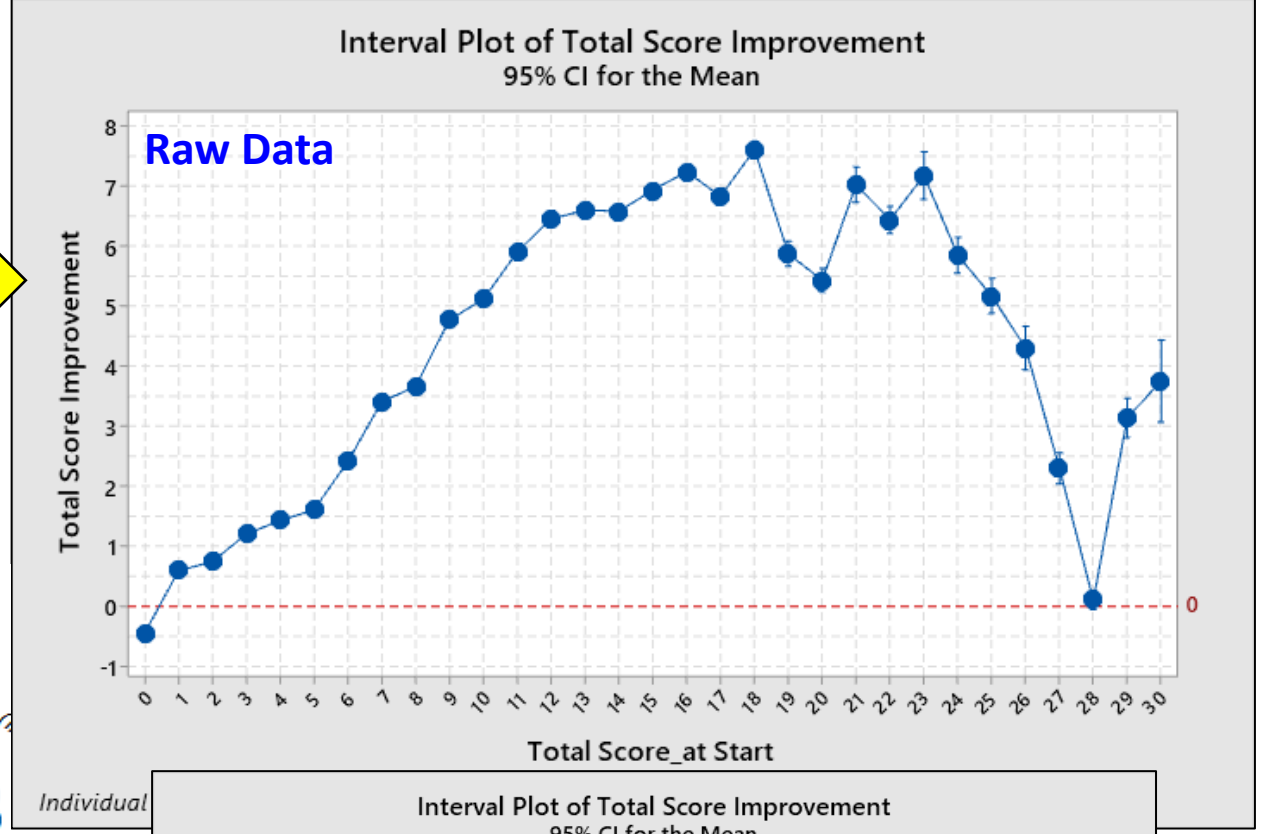
Some of the Top Predictors from which to Determine the Best P&I Codes to improve ADLs for Patients

Predictor	Mean Difference in Total Score Between High and Low Group	Known information up Front?
Start of Care Total Score Groups (31)	**8.06**	Yes
Clinical groups (12), but R is absent until 2020	**4.48**	Yes
Functional groups (3)	**2.27**	Yes
Age groups (6)	**1.4**	Yes
Comorbidity levels (3)	**0.97**	Yes
Smoking (Y/N)	**0.54**	Yes
Smoking + Obesity (Y/N)	**0.46**	Yes
Obesity (Y/N)	**0.18**	Yes

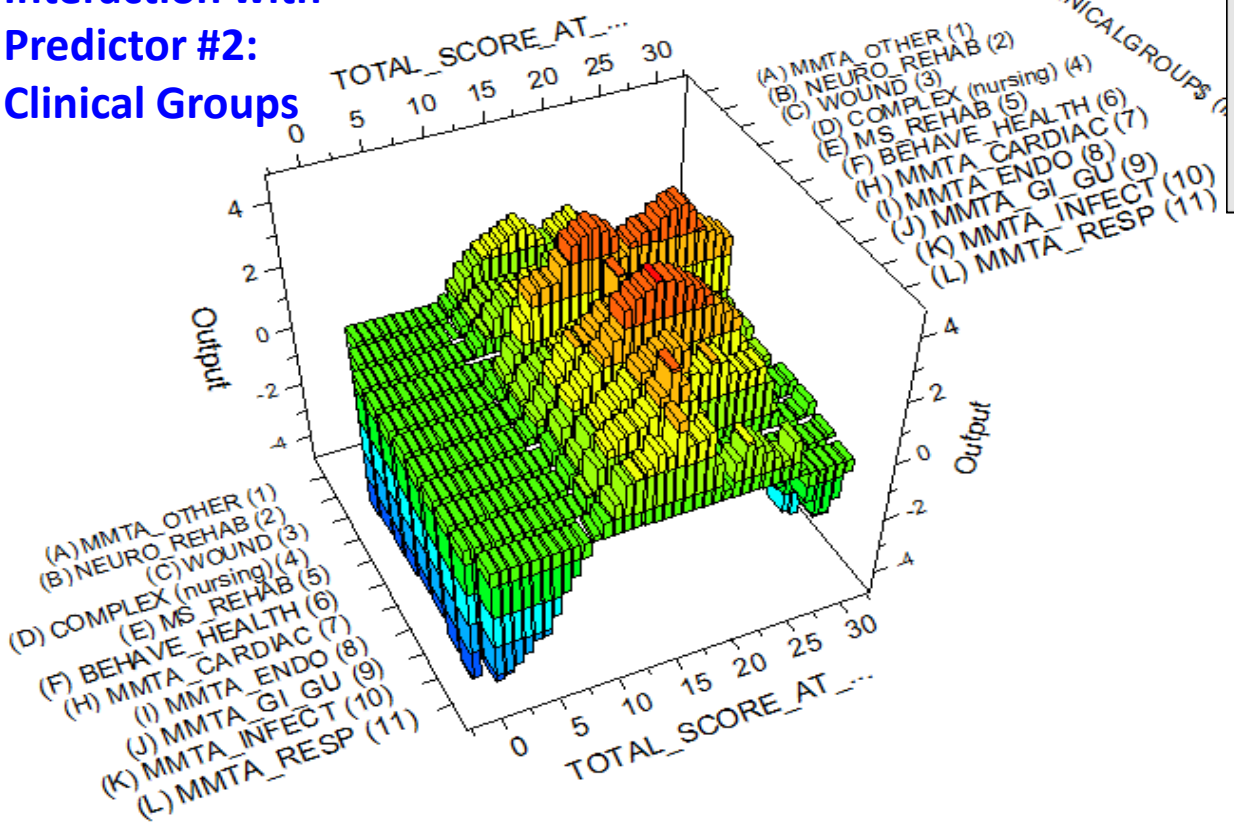
Predictor #1: Start of Care Score

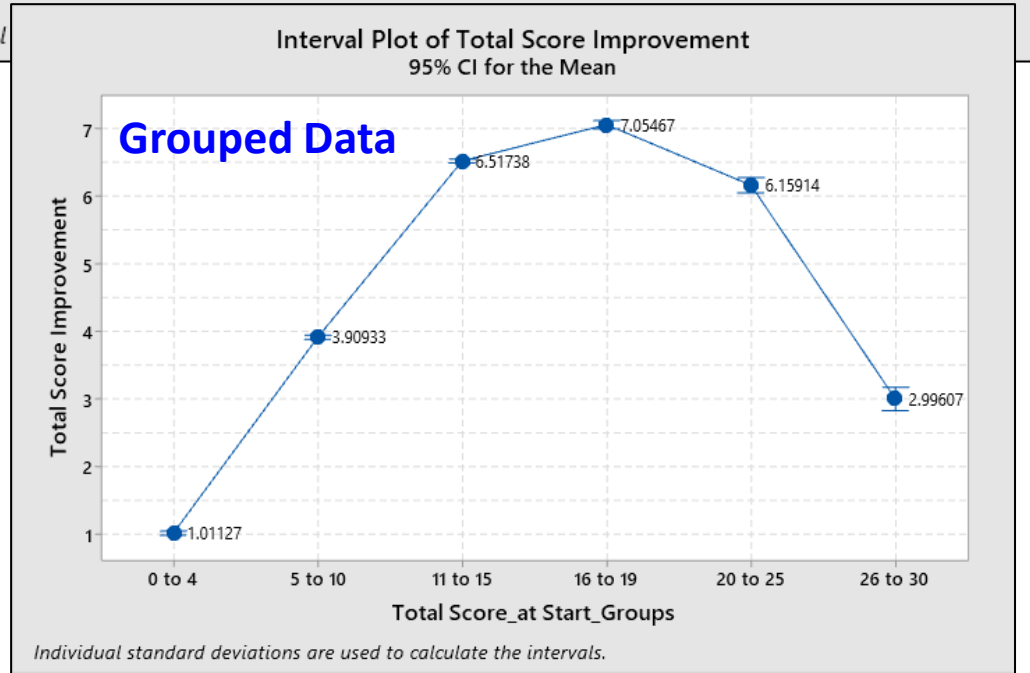
Predictor	Mean Difference in Total Score Between High and Low Group
Start of Care Total Score Groups (31)	8.06
Clinical groups (12), but R is absent until 2020	4.48
Functional groups (3)	2.27
Age groups (6)	1.4
Comorbidity levels (3)	0.97
Smoking (Y/N)	0.54
Smoking + Obesity (Y/N)	0.46
Obesity (Y/N)	0.18

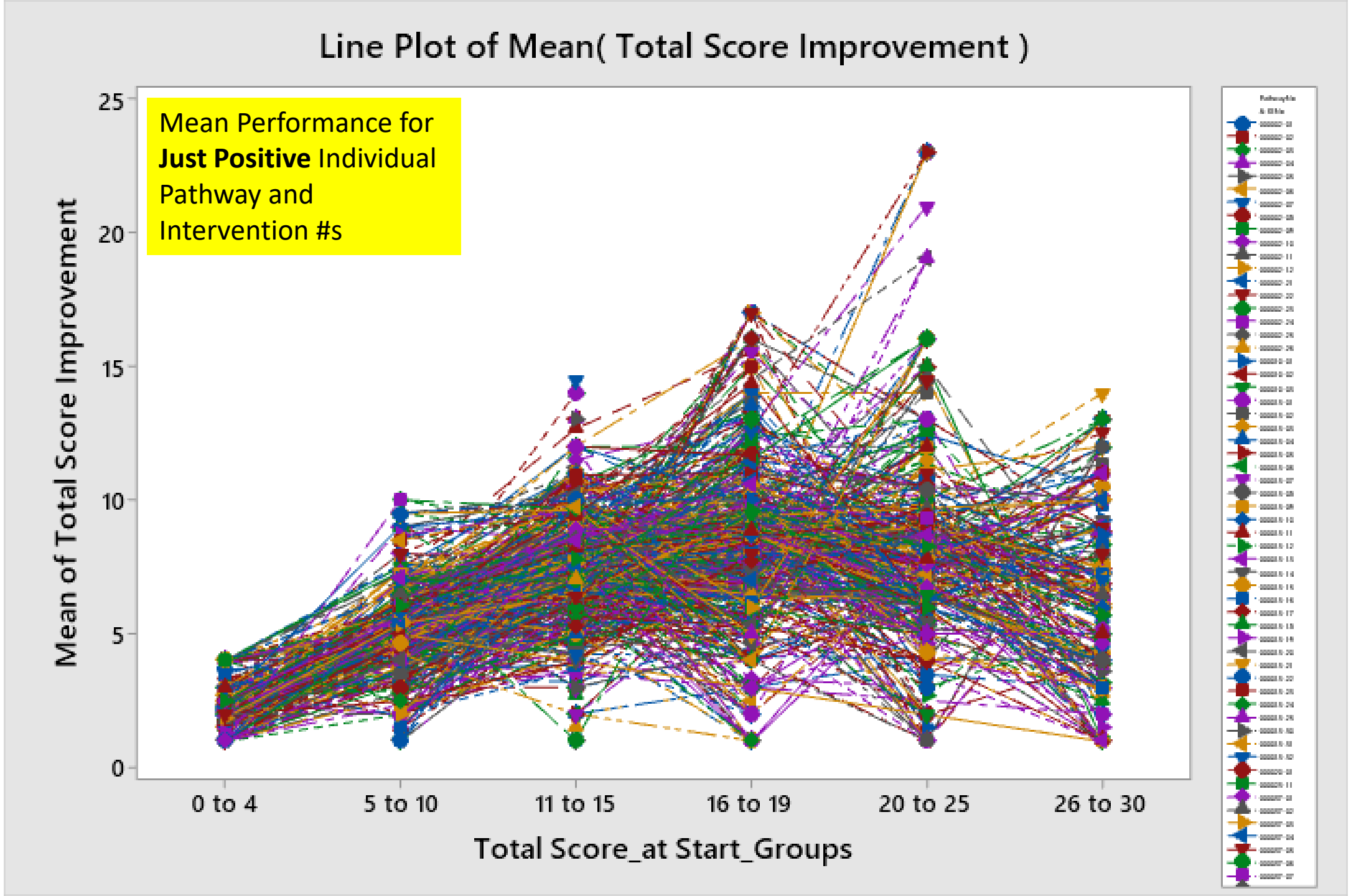
Predictor #1



Interaction with Predictor #2: Clinical Groups



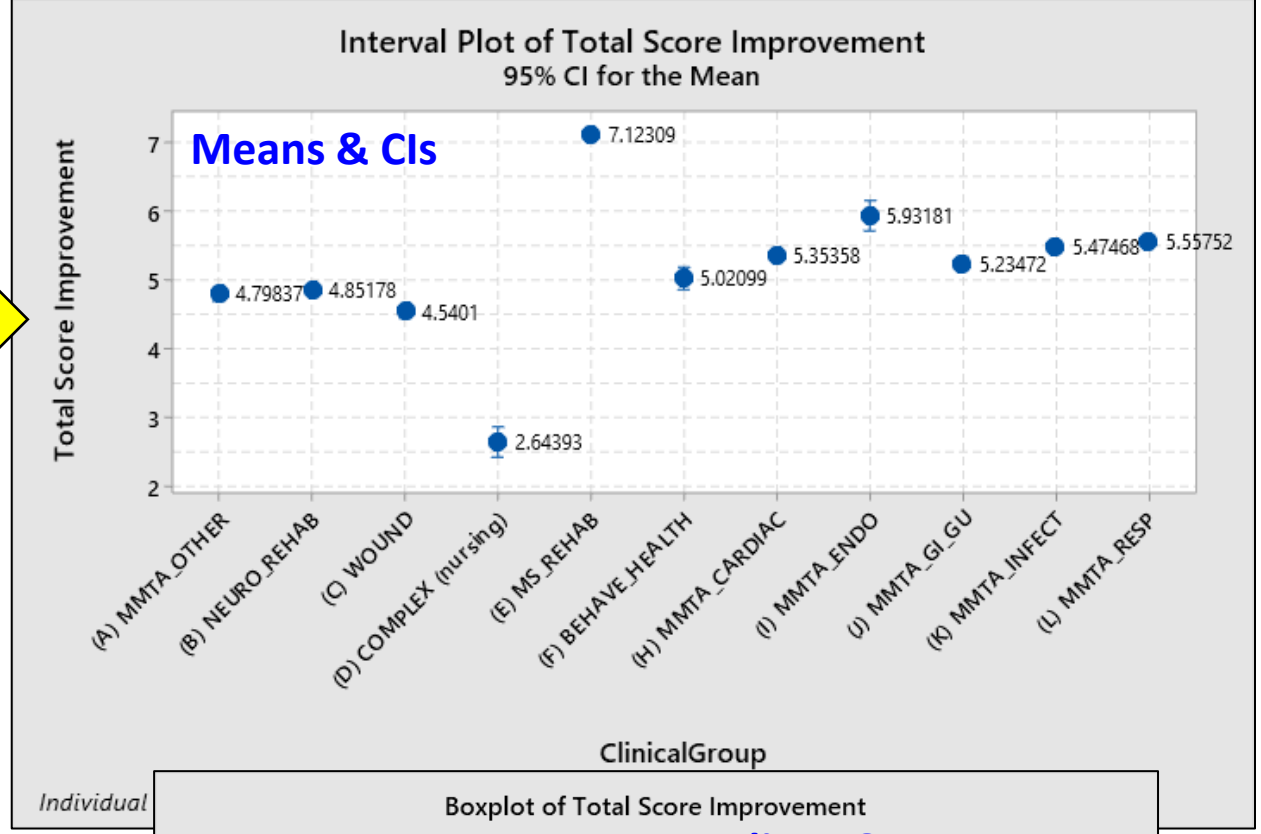



Line Plot of Mean(Total Score Improvement)
Mean Performance for **Just Positive** Individual Pathway and Intervention #s
Mean of Total Score Improvement
25
20
15
10
5
0
0 to 4
5 to 10
11 to 15
16 to 19
20 to 25
26 to 30
Total Score_at Start_Groups

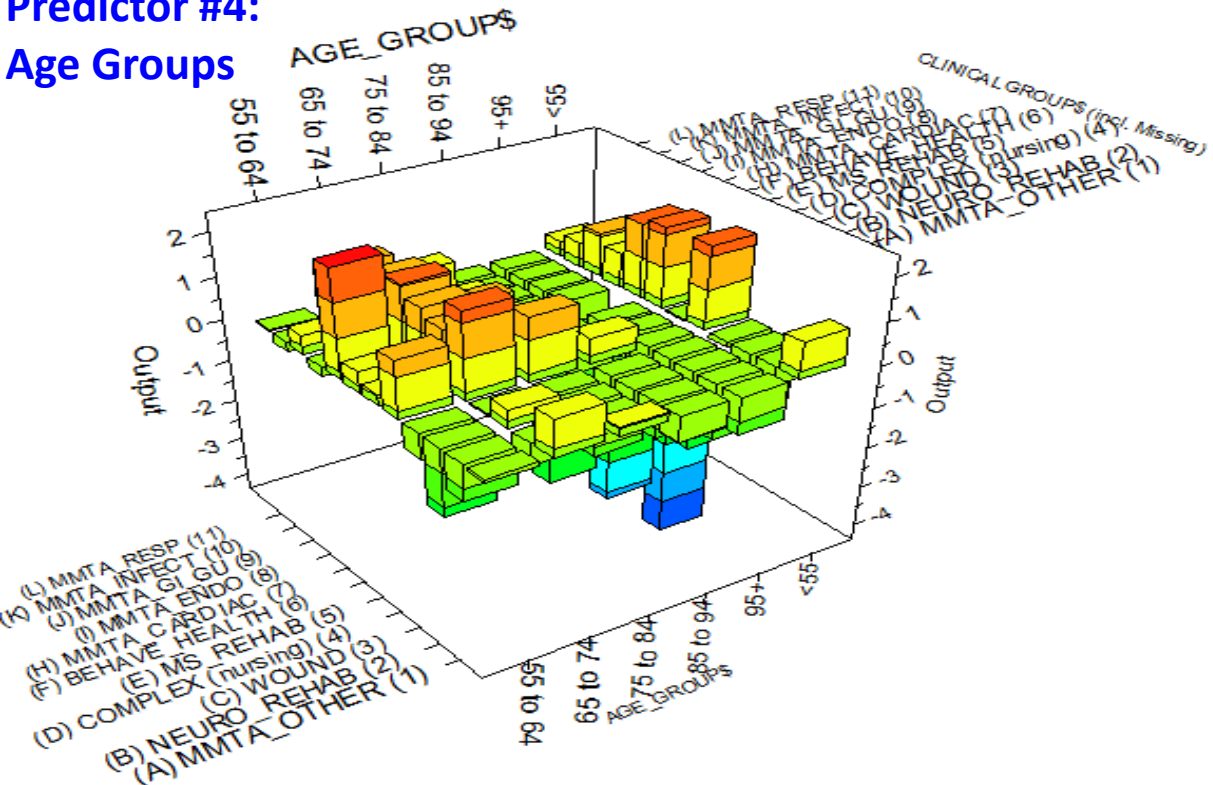
Predictor #2: Clinical Group

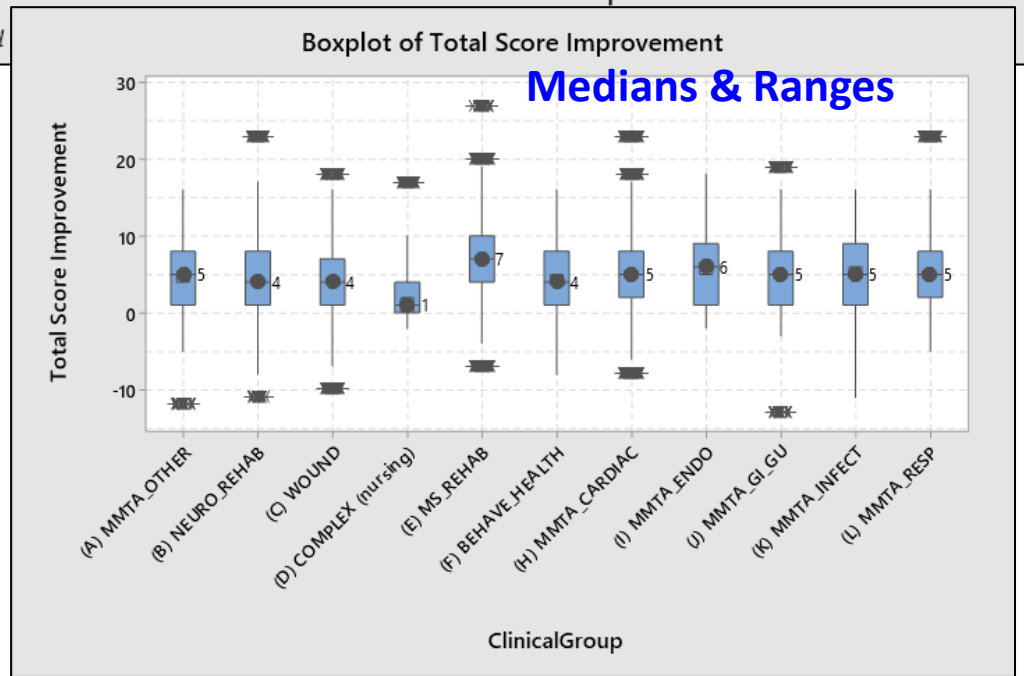
Predictor	Mean Difference in Total Score Between High and Low Group
Start of Care Total Score Groups (31)	8.06
Clinical groups (12), but R is absent until 2020	4.48
Functional groups (3)	2.27
Age groups (6)	1.4
Comorbidity levels (3)	0.97
Smoking (Y/N)	0.54
Smoking + Obesity (Y/N)	0.46
Obesity (Y/N)	0.18

Predictor #2

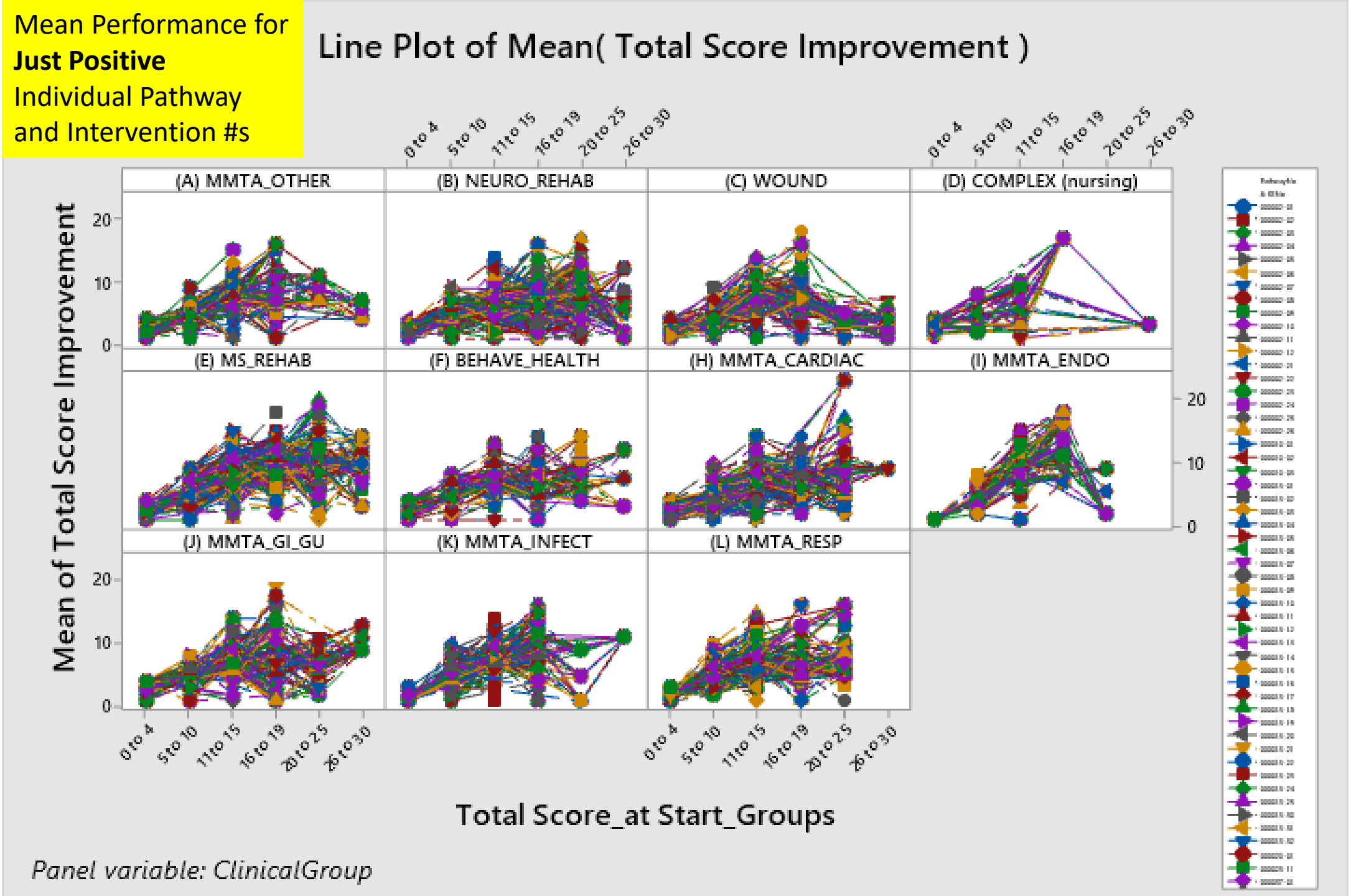


Interaction with Predictor #4: Age Groups





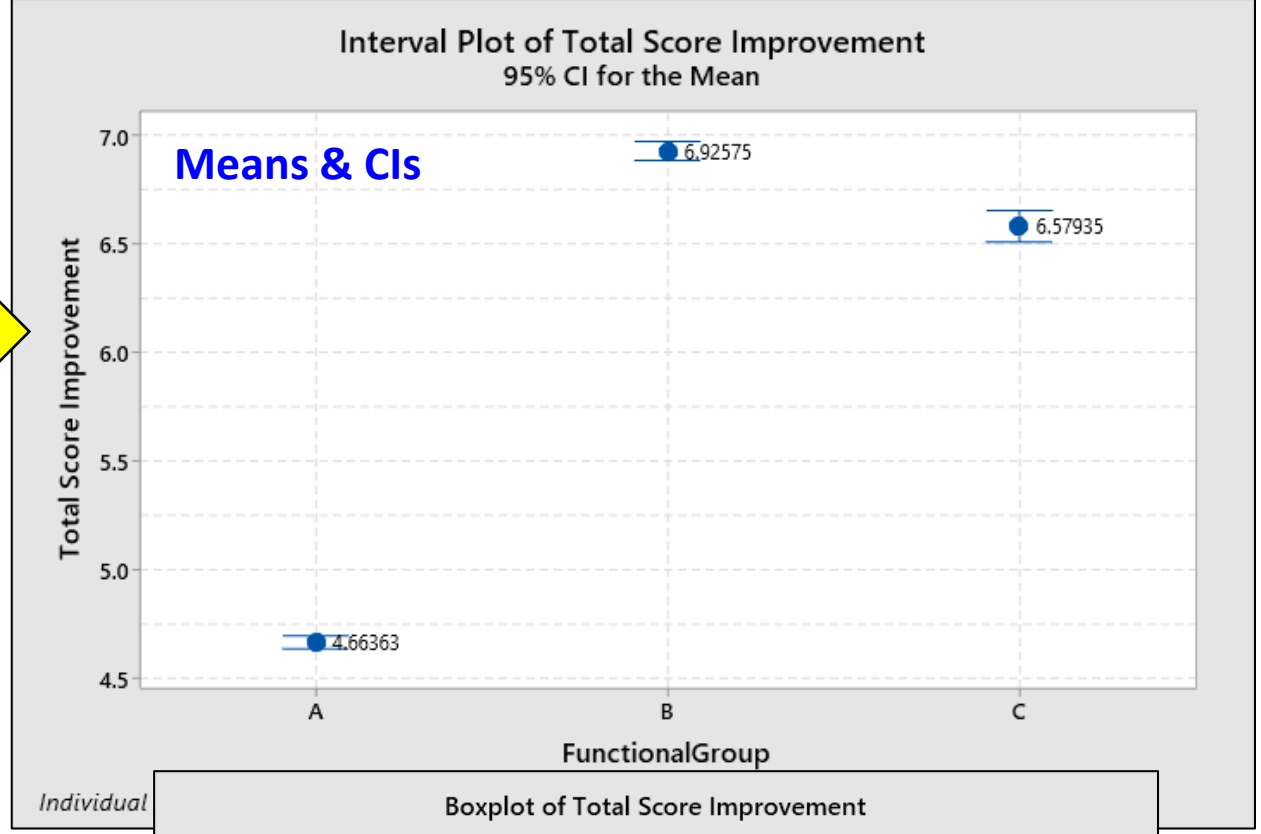
Mean Performance for **Just Positive** Individual Pathway and Intervention #s



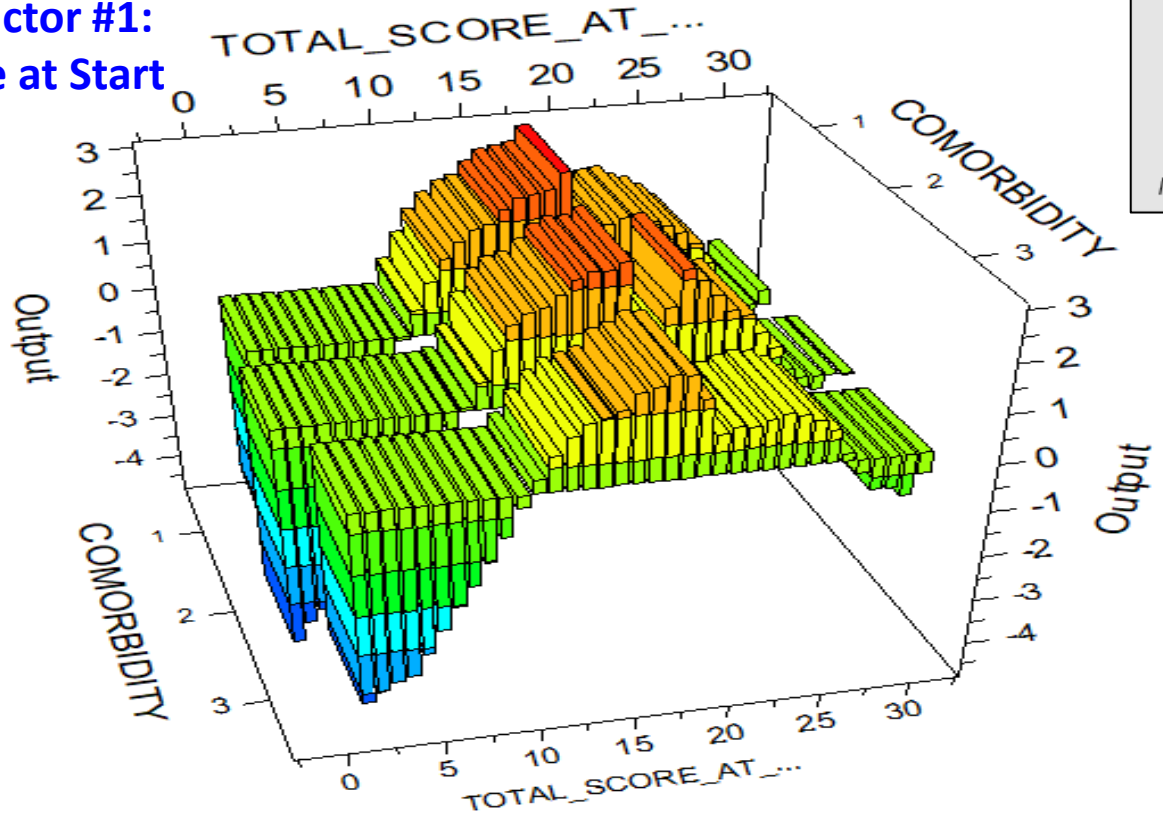
Predictor #3: Functional Group

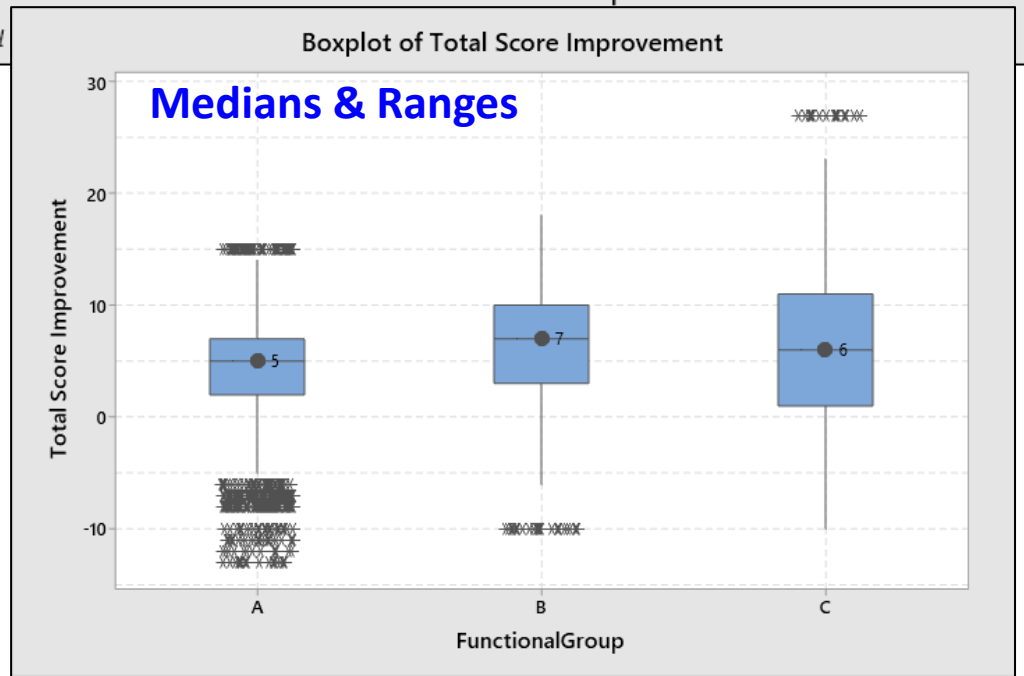
Predictor	Mean Difference in Total Score Between High and Low Group
Start of Care Total Score Groups (31)	8.06
Clinical groups (12), but R is absent until 2020	4.48
Functional groups (3)	2.27
Age groups (6)	1.4
Comorbidity levels (3)	0.97
Smoking (Y/N)	0.54
Smoking + Obesity (Y/N)	0.46
Obesity (Y/N)	0.18

Predictor #3



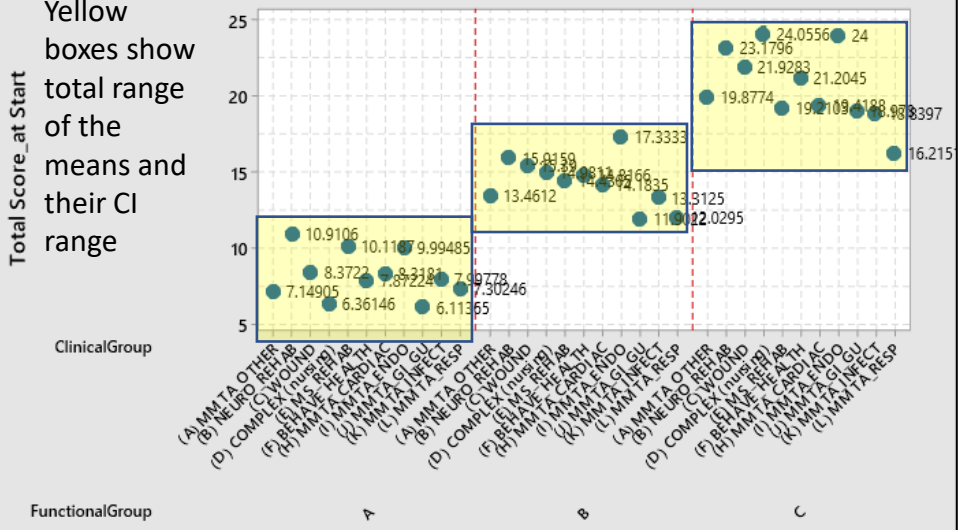
Interaction with Predictor #1: Score at Start





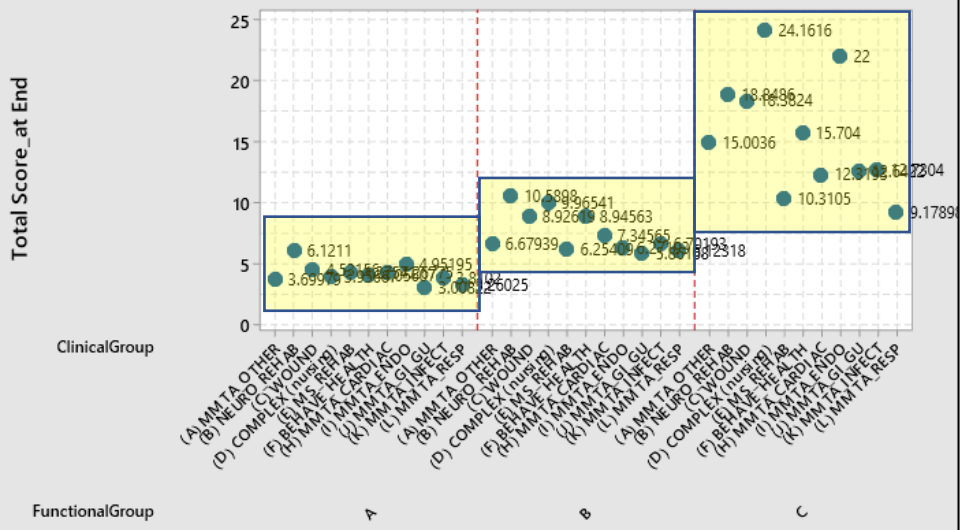
Start

Yellow boxes show total range of the means and their CI range



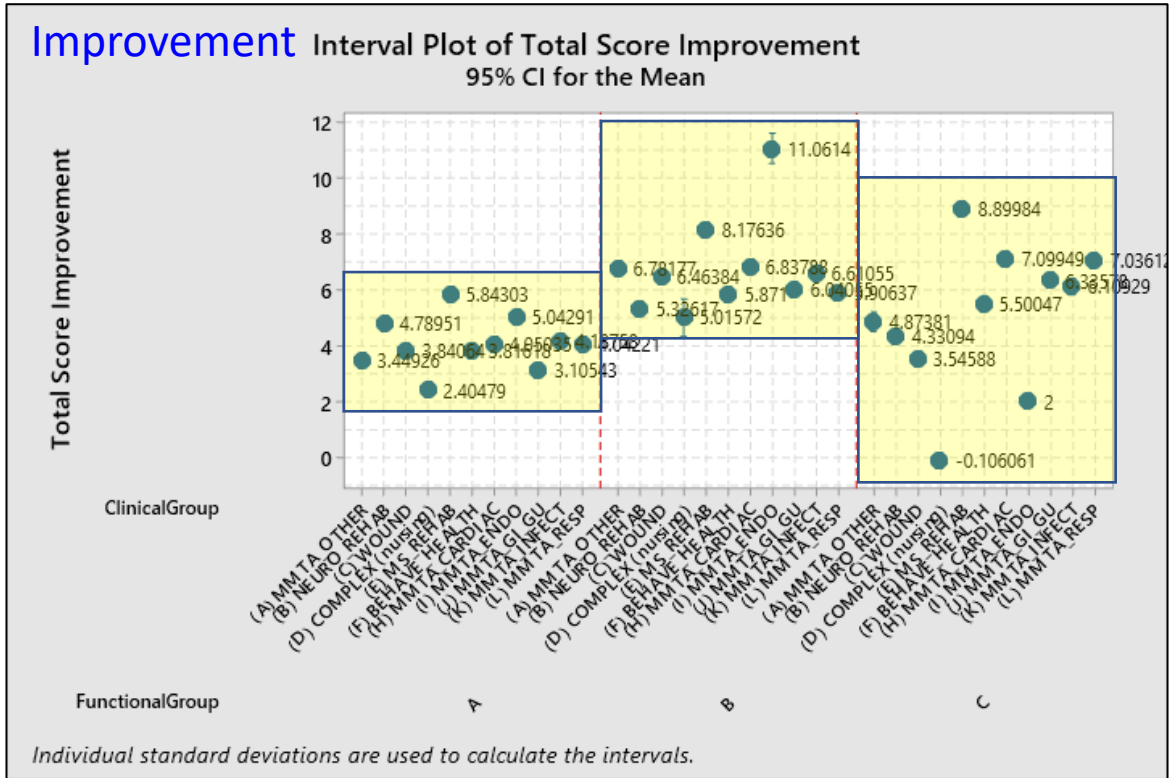
Individual standard deviations are used to calculate the intervals.

End



Individual standard deviations are used to calculate the intervals.

Improvement

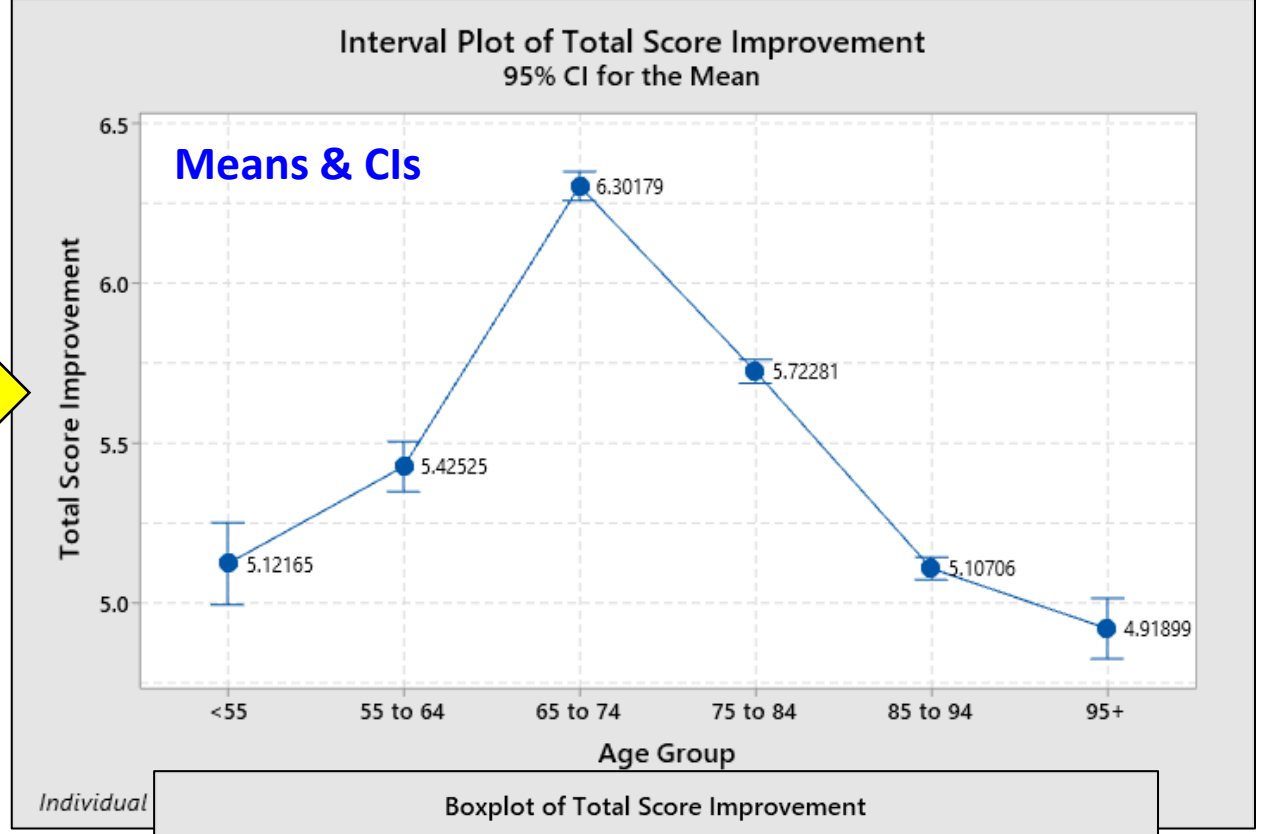


Individual standard deviations are used to calculate the intervals.

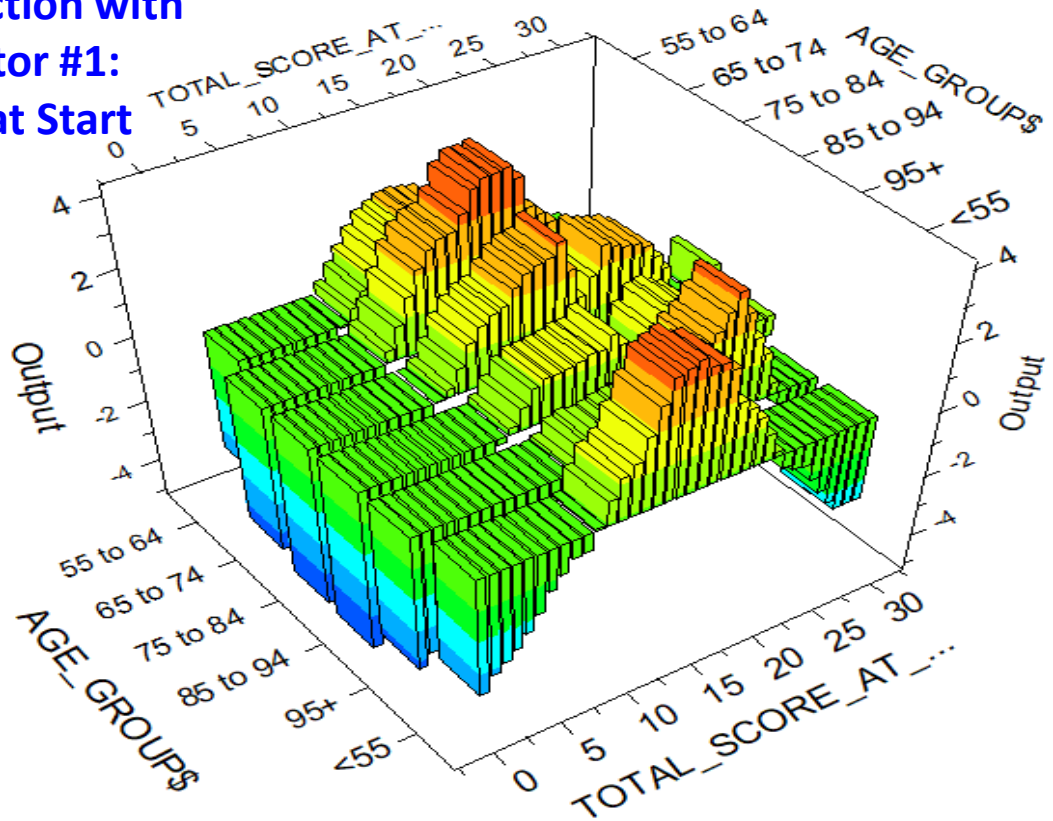
Predictor #4: Age Group

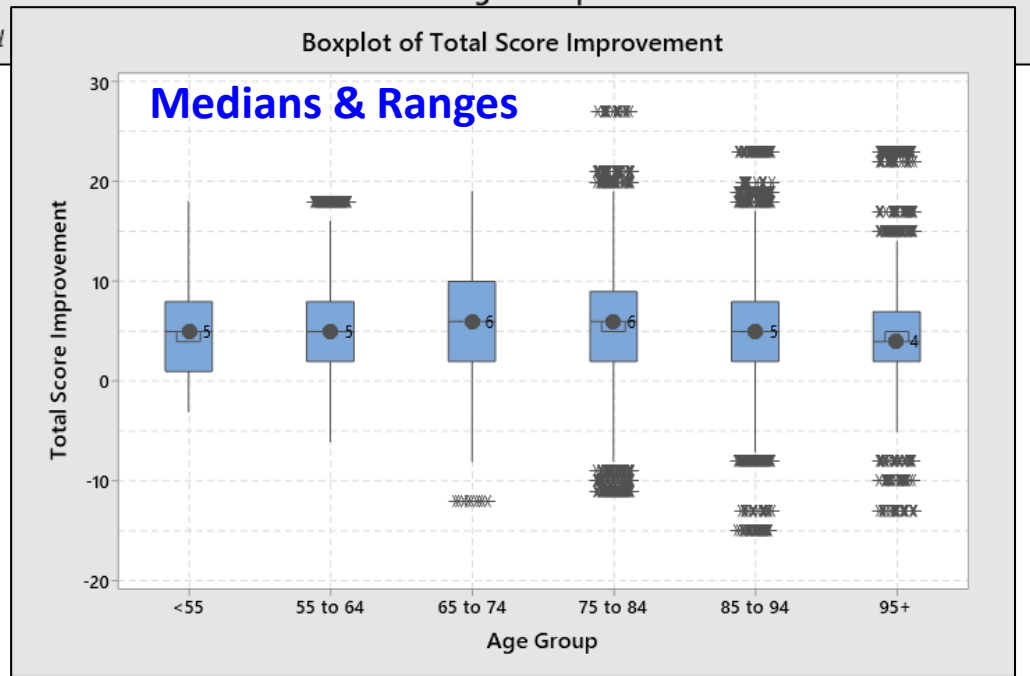
Predictor	Mean Difference in Total Score Between High and Low Group
Start of Care Total Score Groups (31)	8.06
Clinical groups (12), but R is absent until 2020	4.48
Functional groups (3)	2.27
Age groups (6)	1.4
Comorbidity levels (3)	0.97
Smoking (Y/N)	0.54
Smoking + Obesity (Y/N)	0.46
Obesity (Y/N)	0.18

Predictor #4

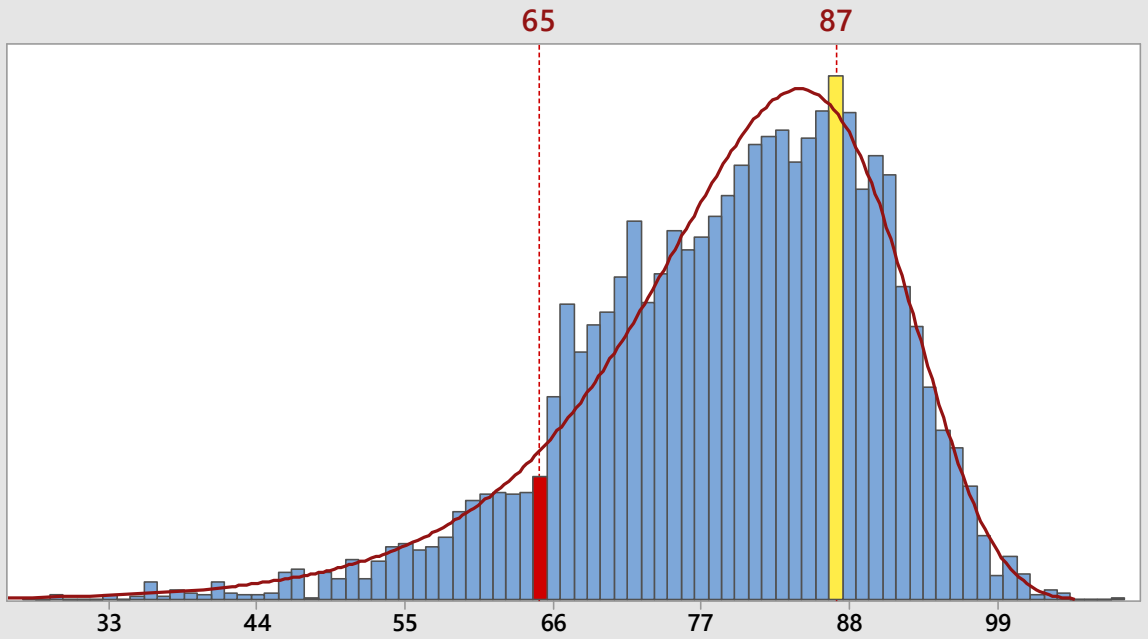


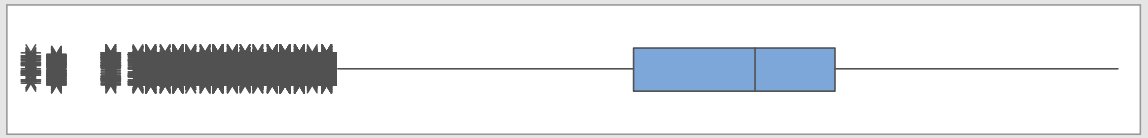
Interaction with Predictor #1: Score at Start

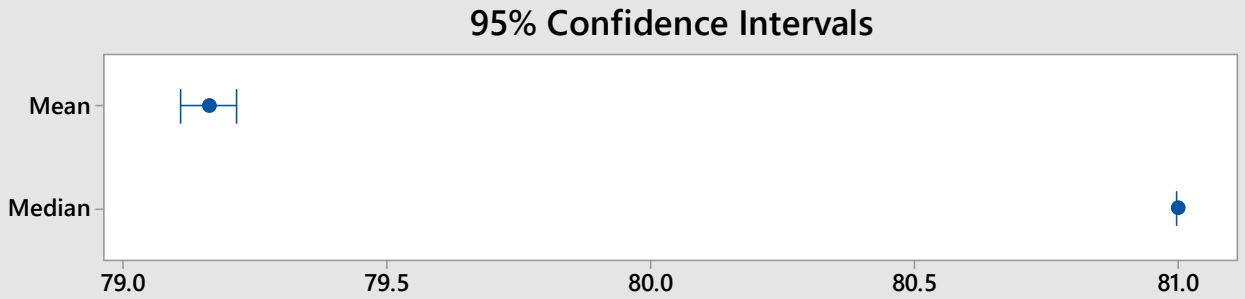




Summary Report for Age

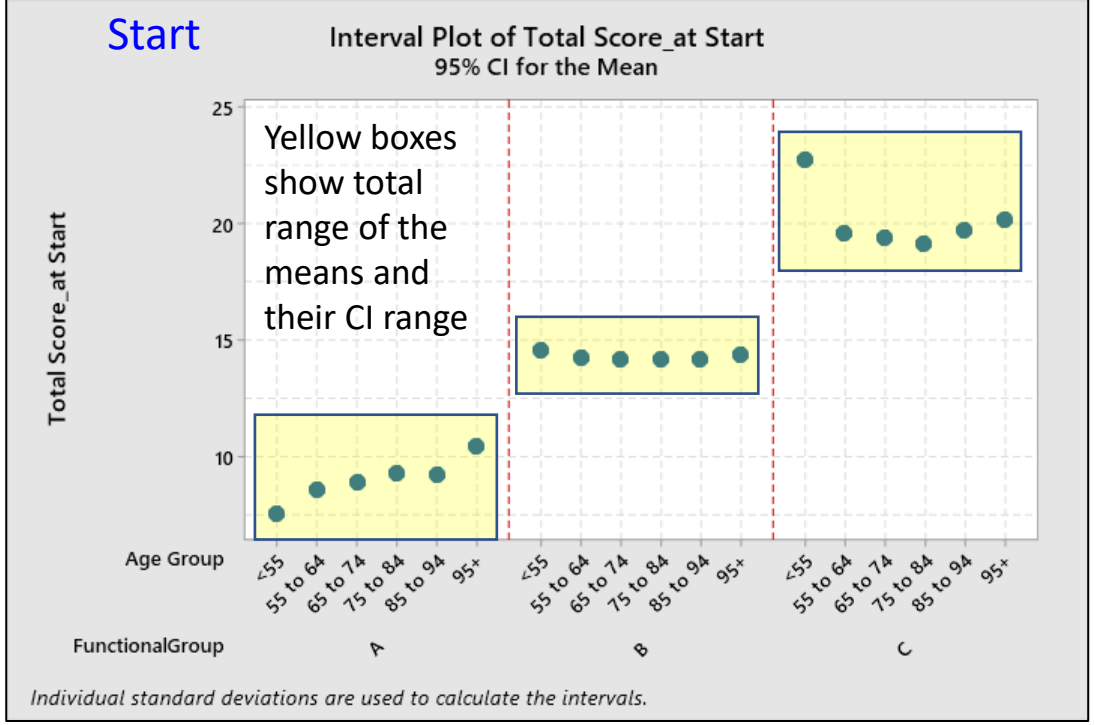




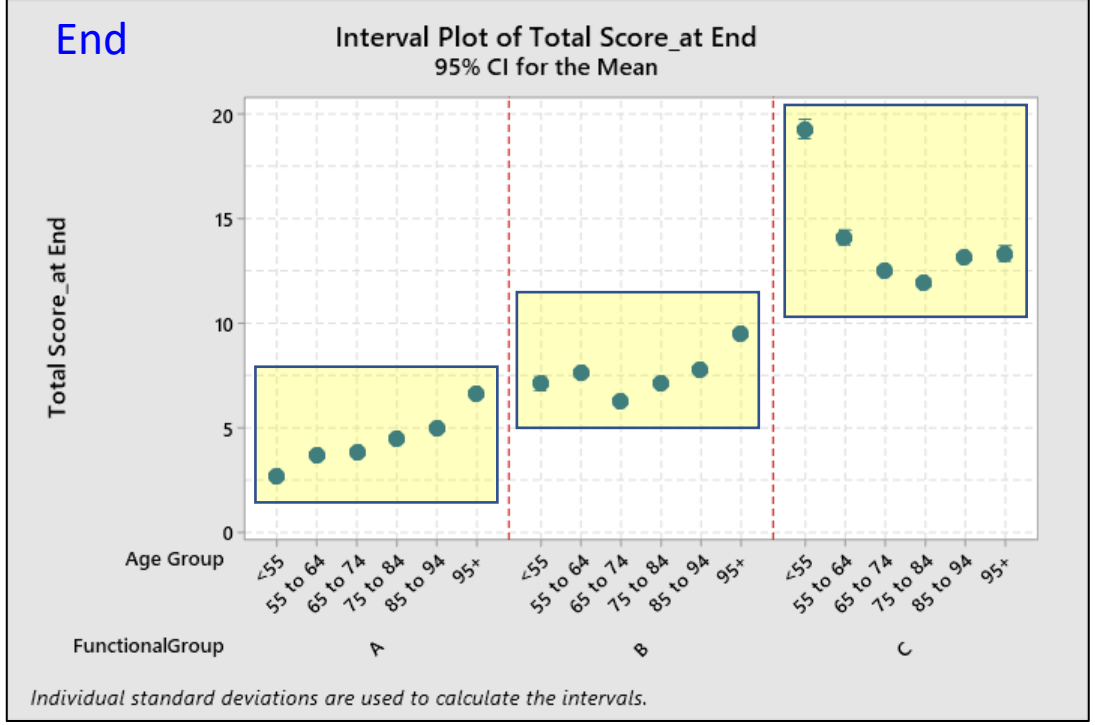


Anderson-Darling Normality Test	
A-Squared	1283.58
P-Value	<0.005
Mean	79.159
StDev	11.205
Variance	125.553
Skewness	-0.789424
Kurtosis	0.864548
N	166707
Minimum	27.000
1st Quartile	72.000
Median	81.000
3rd Quartile	87.000
Maximum	108.000
95% Confidence Interval for Mean	
79.106	79.213
95% Confidence Interval for Median	
81.000	81.000
95% Confidence Interval for StDev	
11.167	11.243

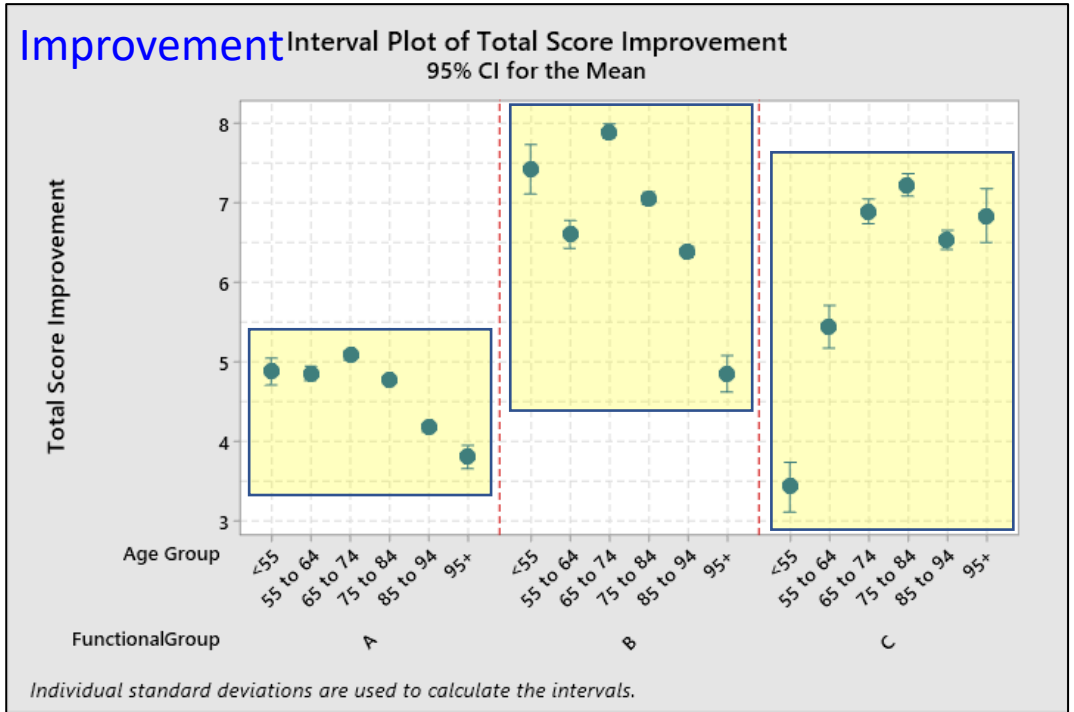
Start



End

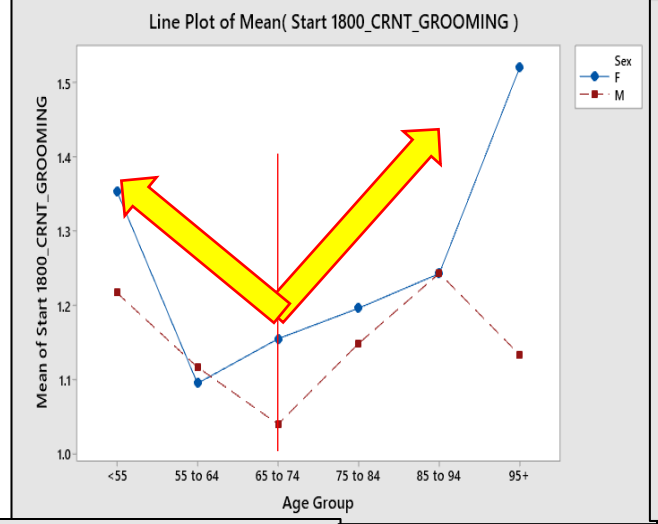


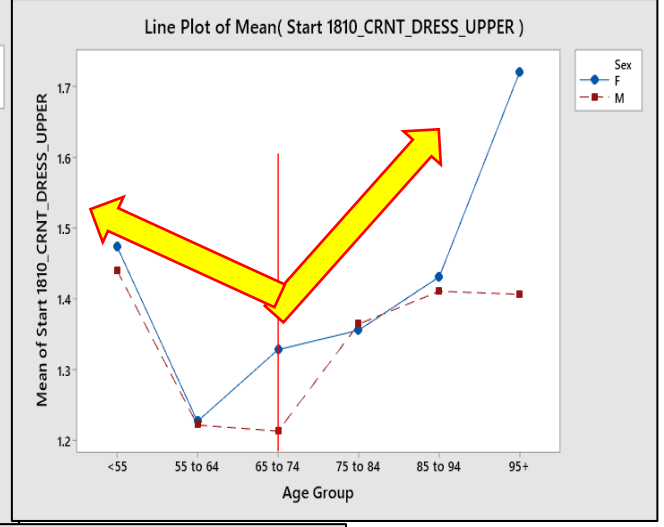
Improvement



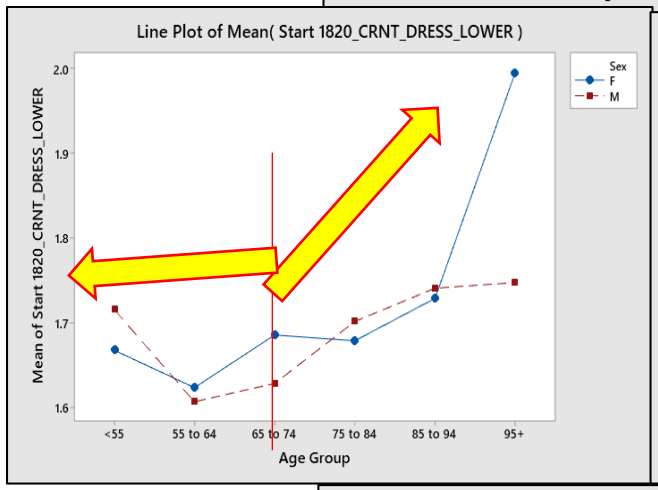
Starting Values

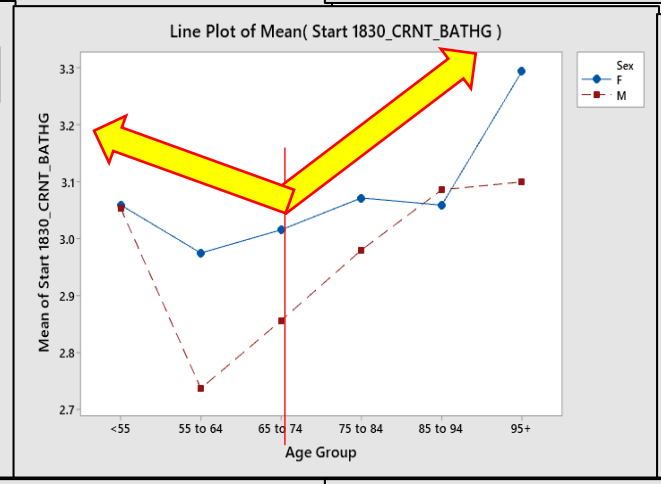
In general, All OASIS **Activities of Daily Living** (ADLs) **Start** ratings steadily decline after age group 65-74, while **males** maintain a slight advantage in the 96+ age group vs **females**

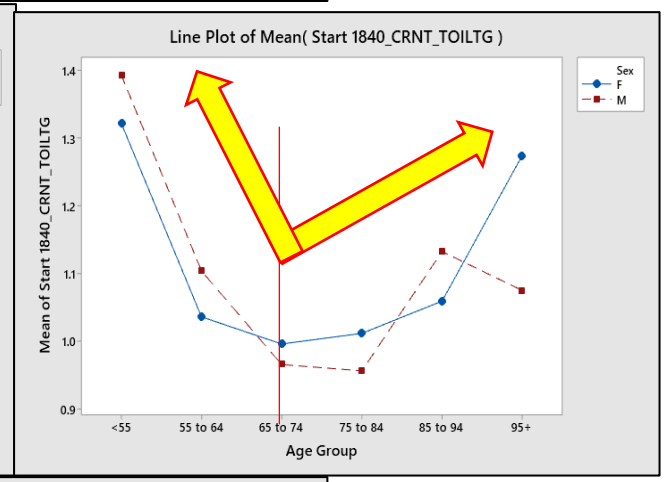




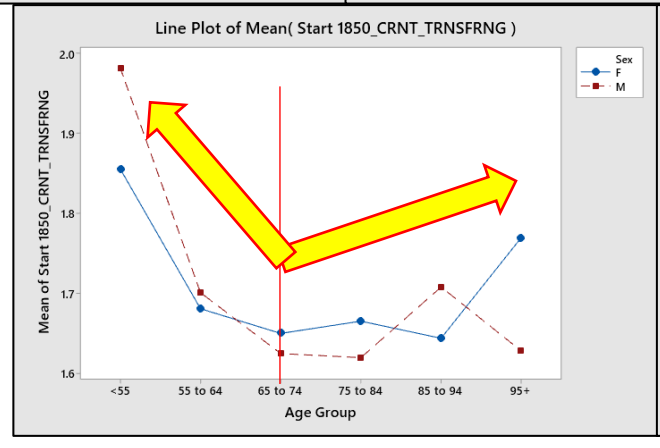
Beware of different Y-Axis scales between graphics

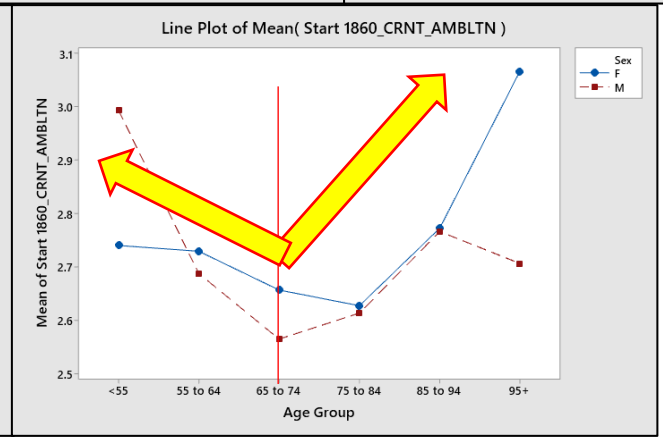






It's unusual that the 65-74 age group is often better then younger age groups

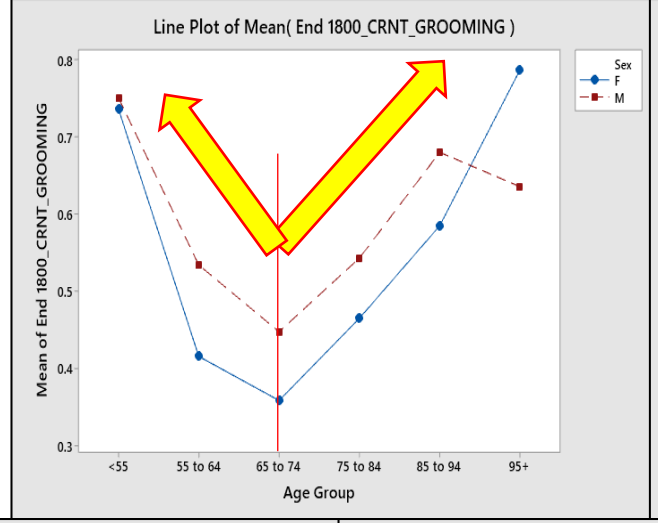


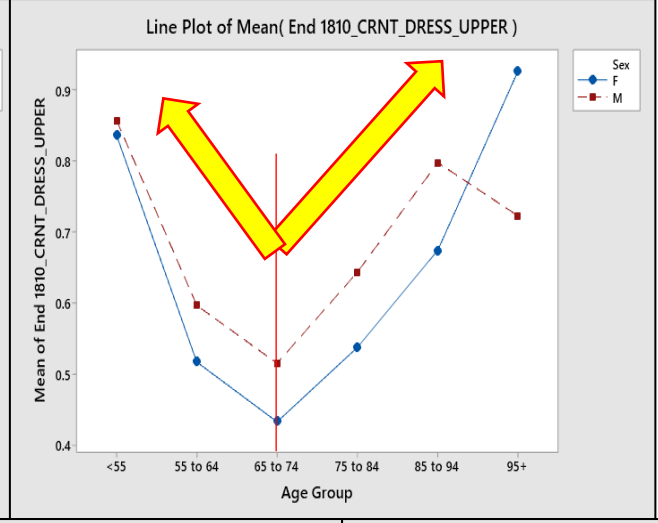


Lower scores are better

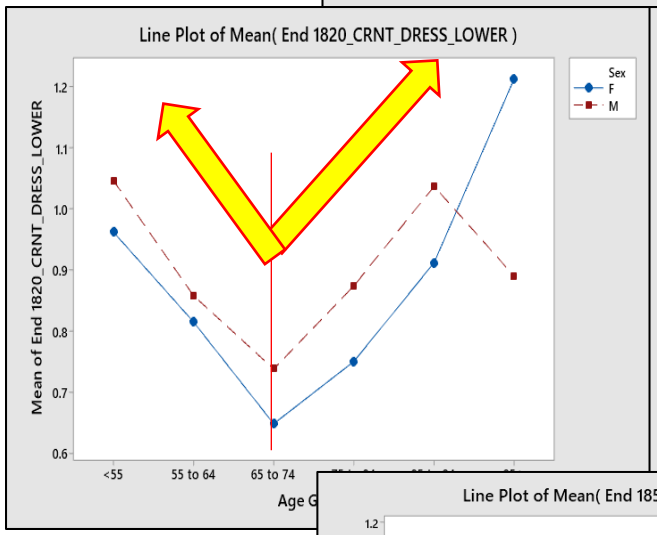
End values

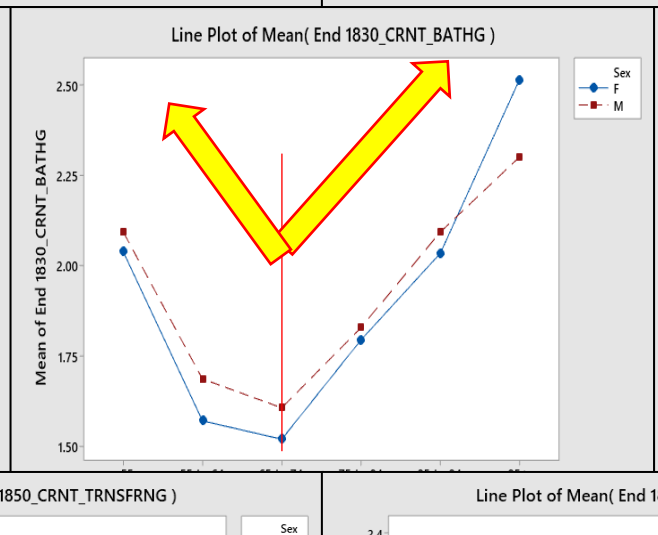
In general, All OASIS **Activities of Daily Living** (ADLs) **End** ratings steadily decline after age group 65-74, while **males** maintain a slight advantage in the 96+ age group vs **females**

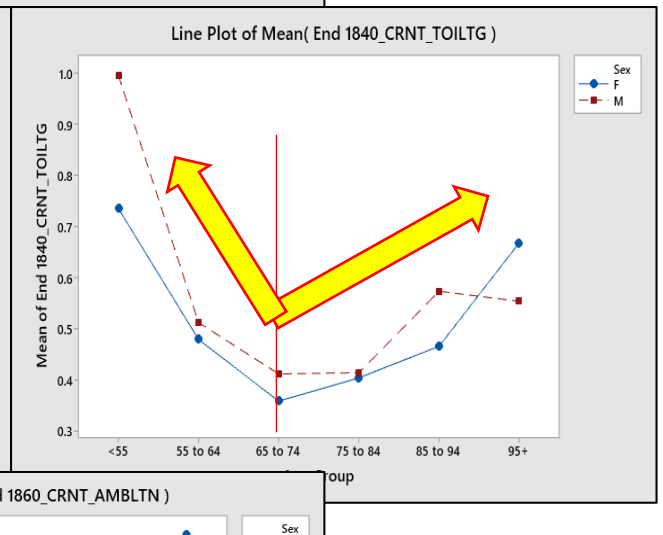




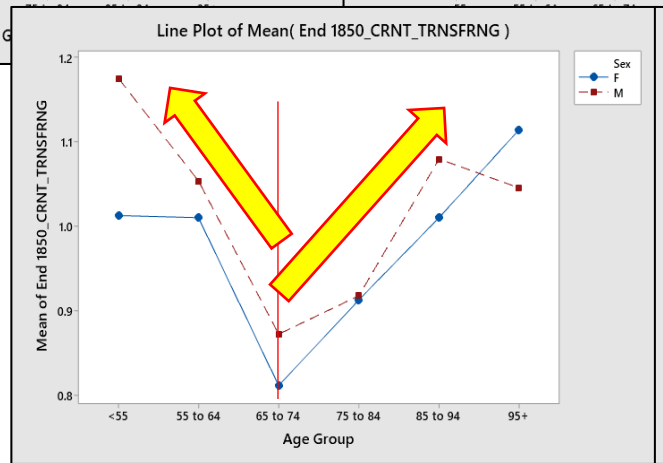
Beware of different Y-Axis scales between graphics

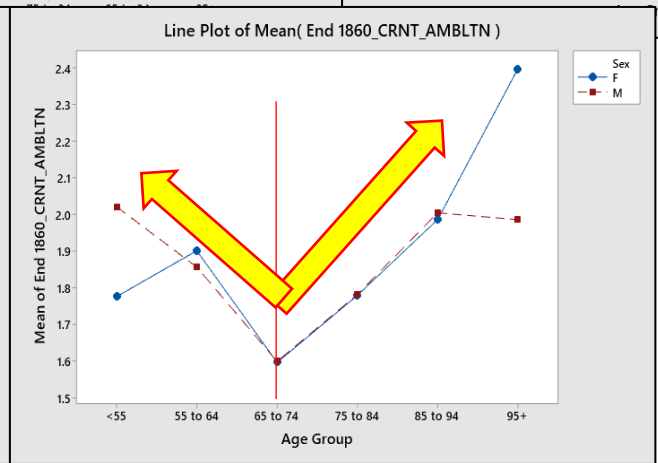






It's unusual that the 65-74 age group is always better then younger age groups

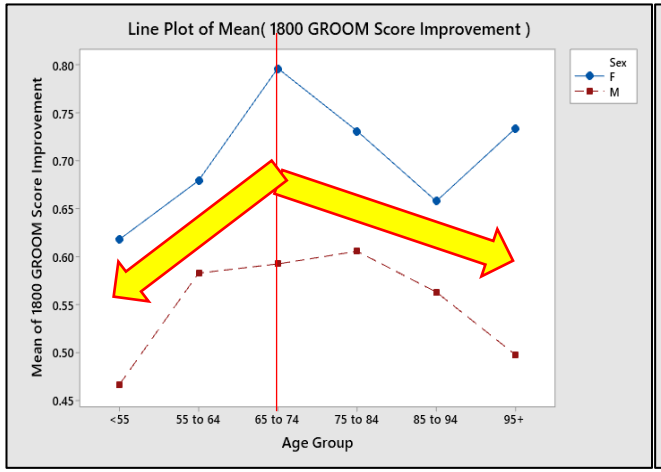




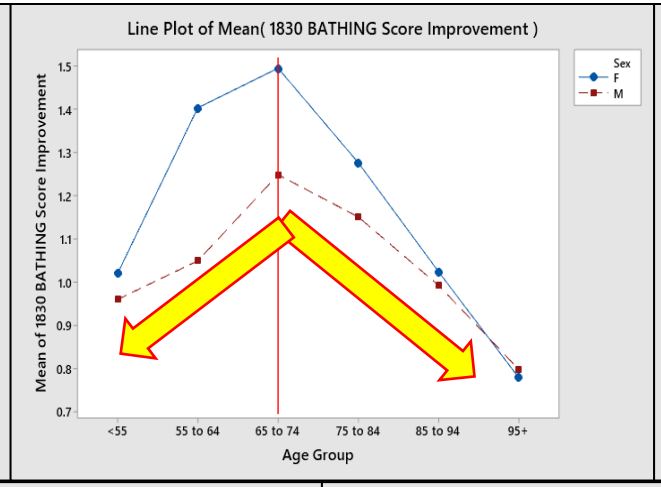
Lower scores are better

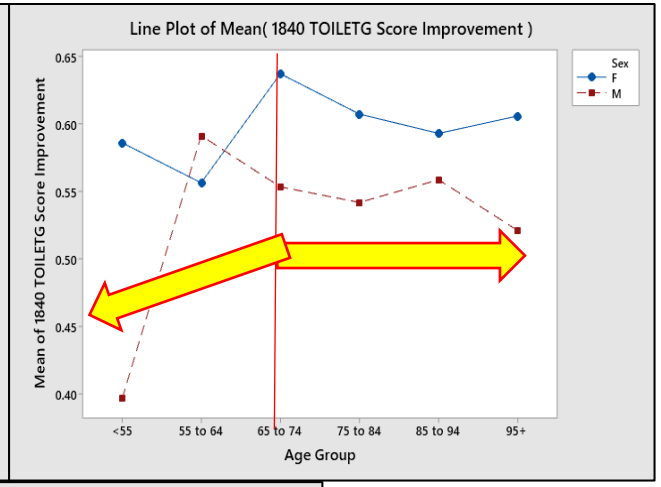
Improvement at end of intervention

In general, All but one OASIS **Activities of Daily Living** (ADLs) **Improvements** steadily declines after age group 65-74

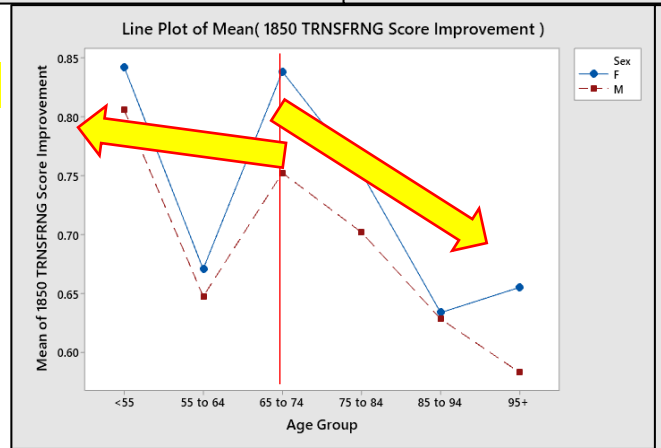


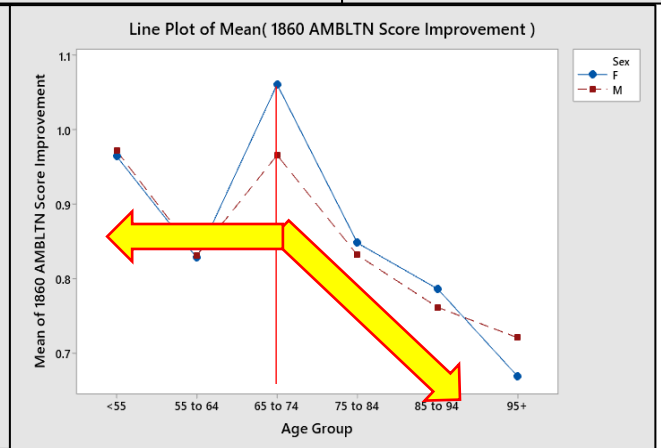
Beware of different Y-Axis scales between graphics





Improvements for 65-74 age groups are mixed, compared to younger age groups

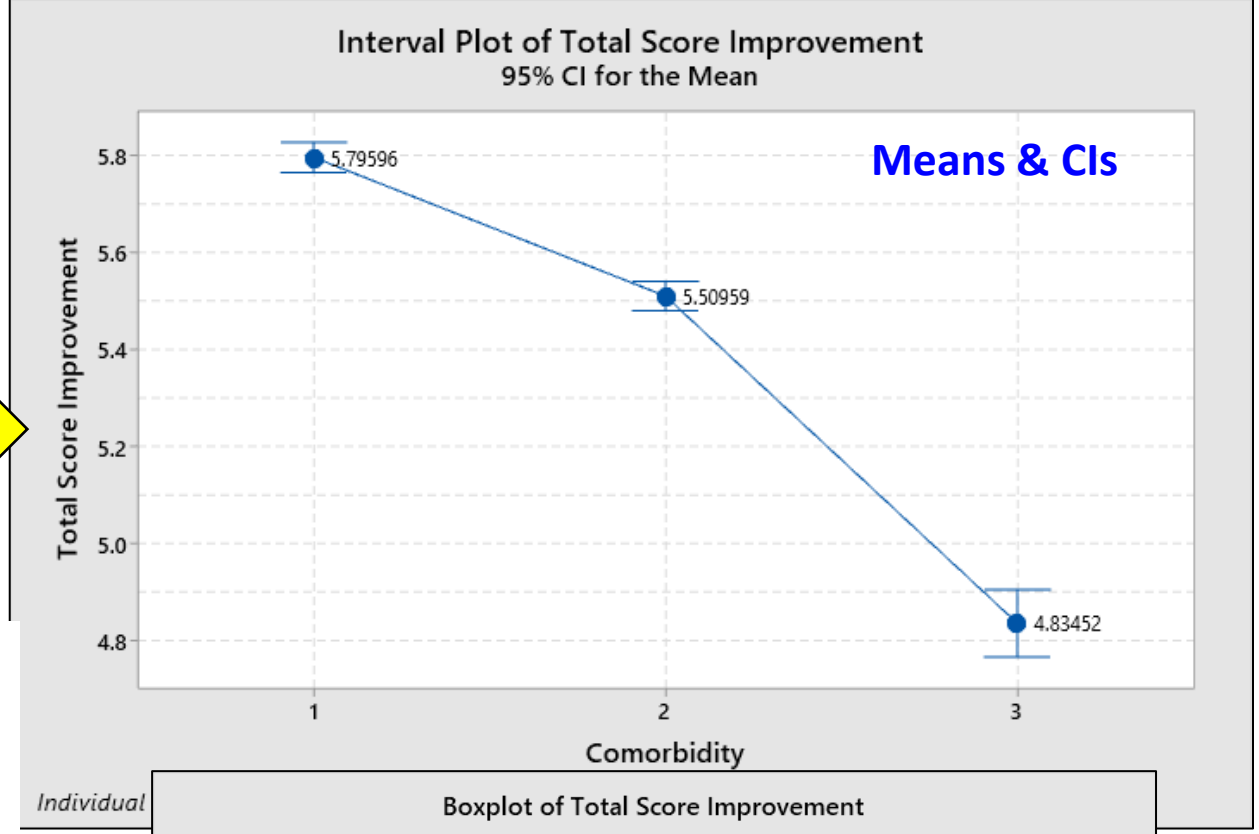




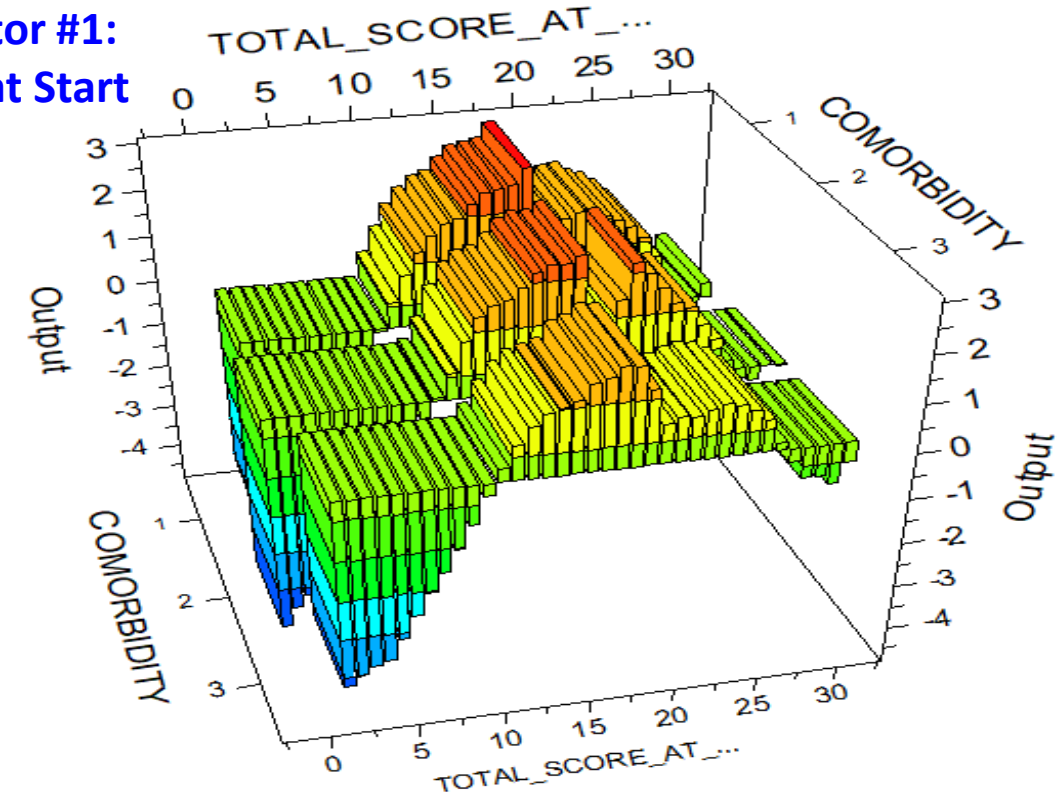
Predictor #5: Comorbidity

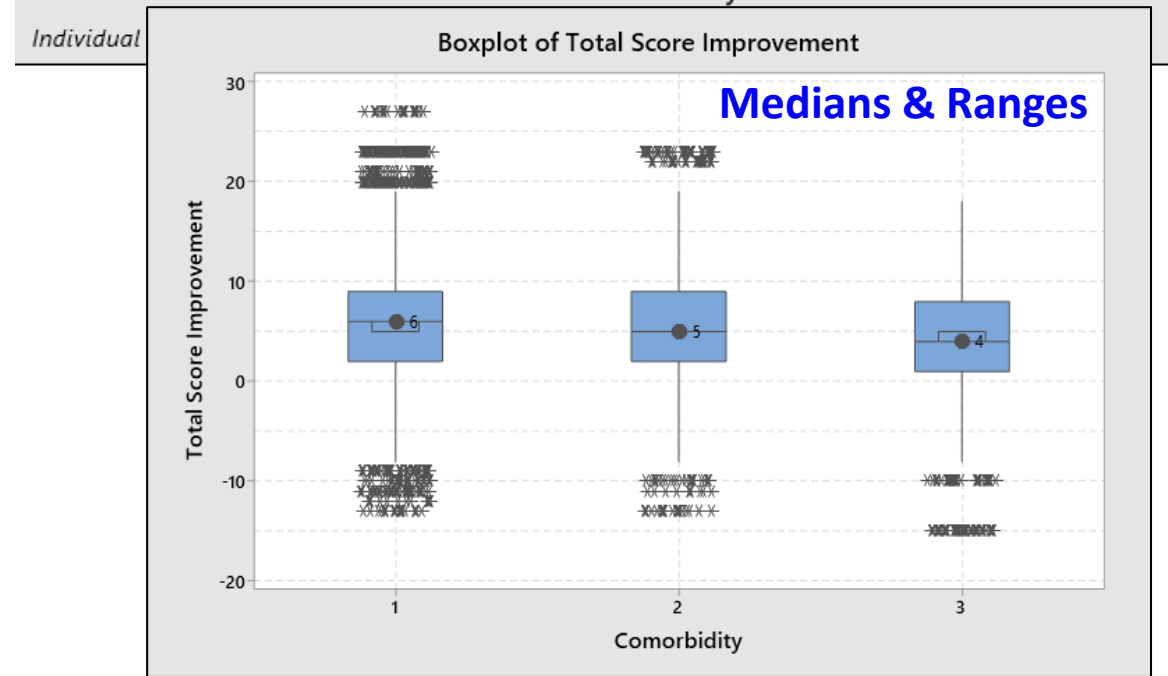
Predictor	Mean Difference in Total Score Between High and Low Group
Start of Care Total Score Groups (31)	8.06
Clinical groups (12), but R is absent until 2020	4.48
Functional groups (3)	2.27
Age groups (6)	1.4
Comorbidity levels (3)	0.97
Smoking (Y/N)	0.54
Smoking + Obesity (Y/N)	0.46
Obesity (Y/N)	0.18

Predictor #5

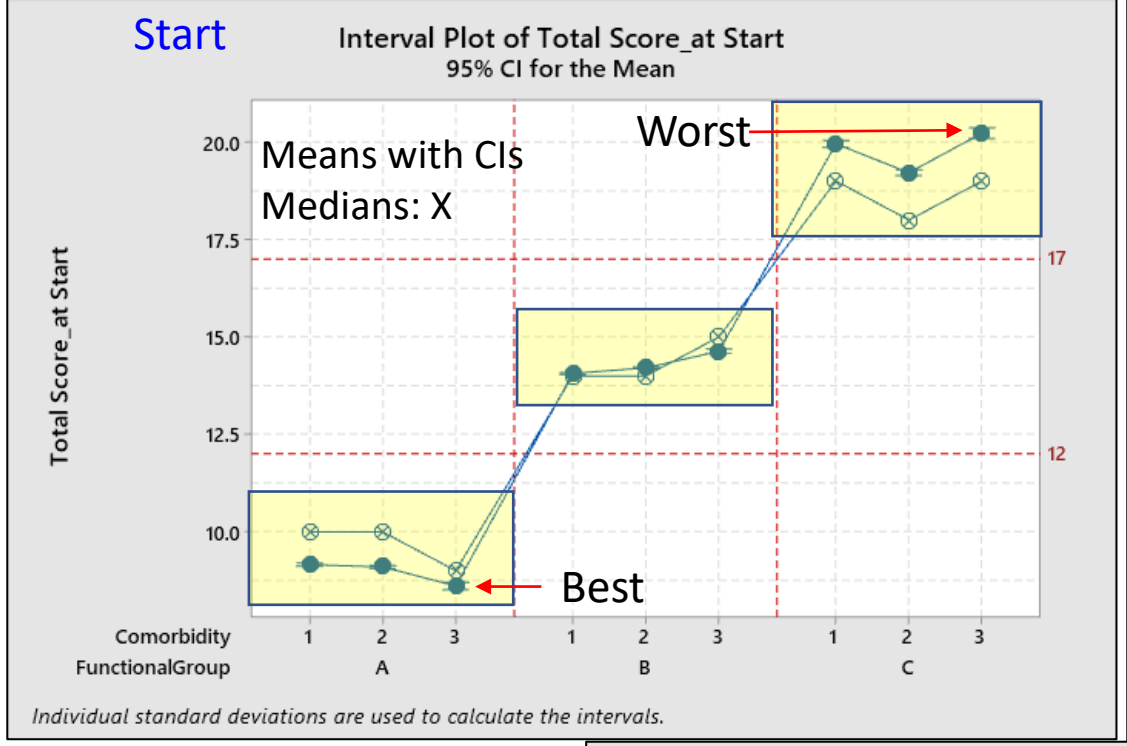


Interaction with Predictor #1: Score at Start

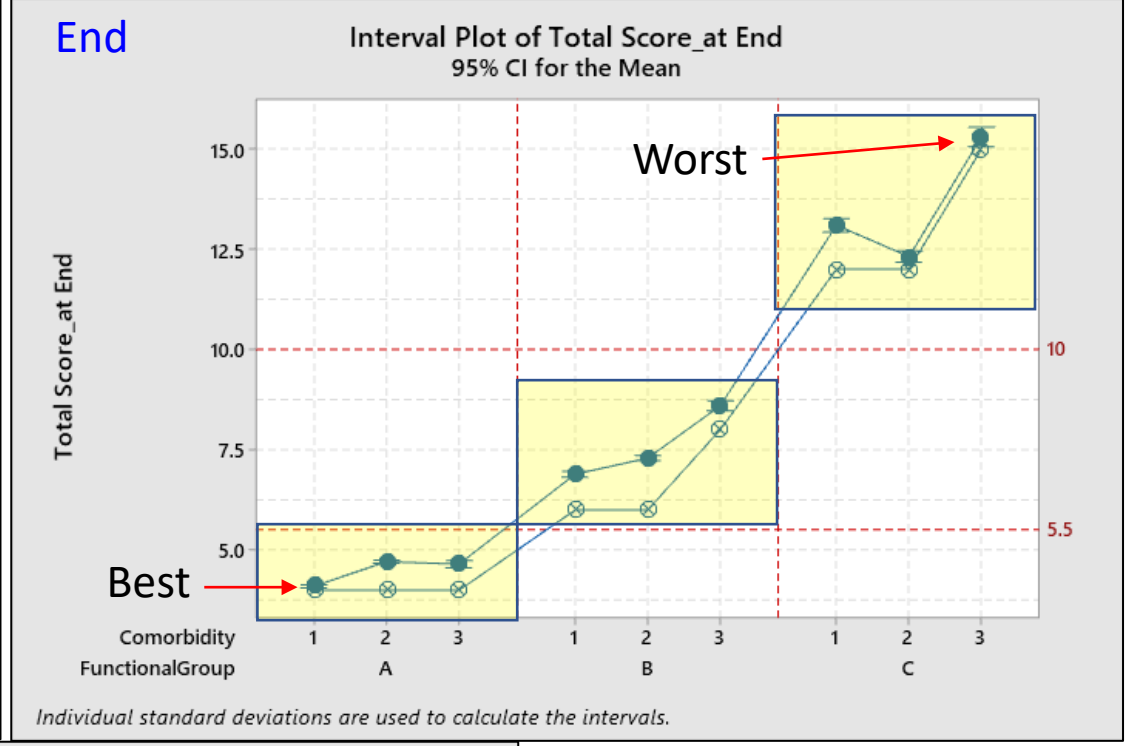




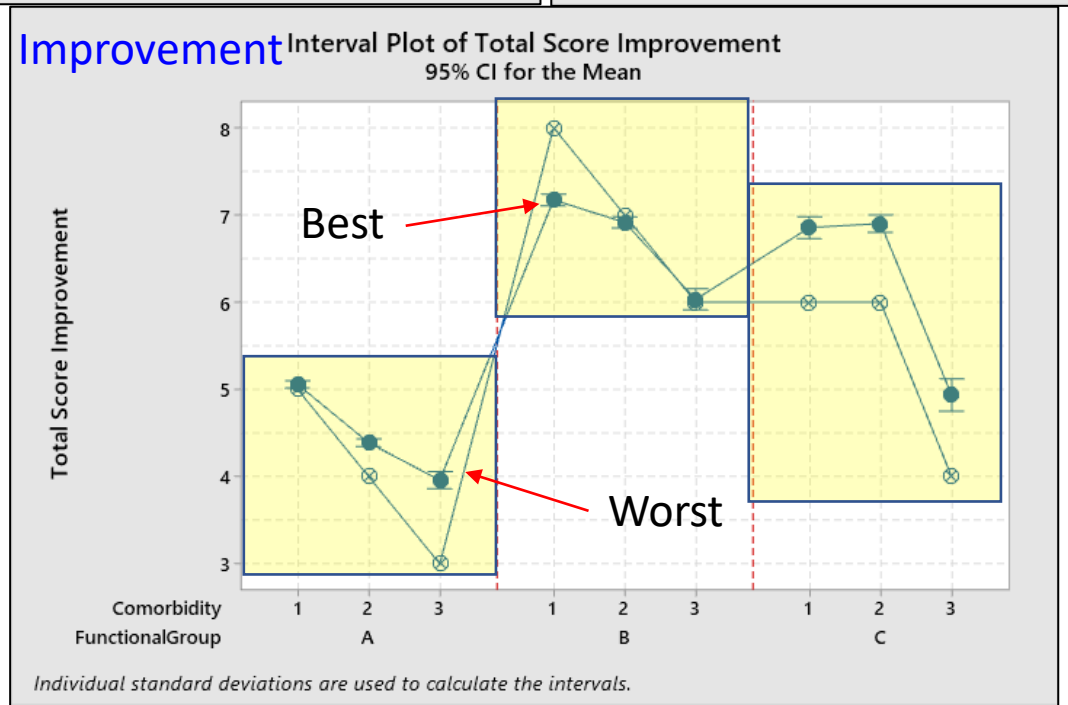
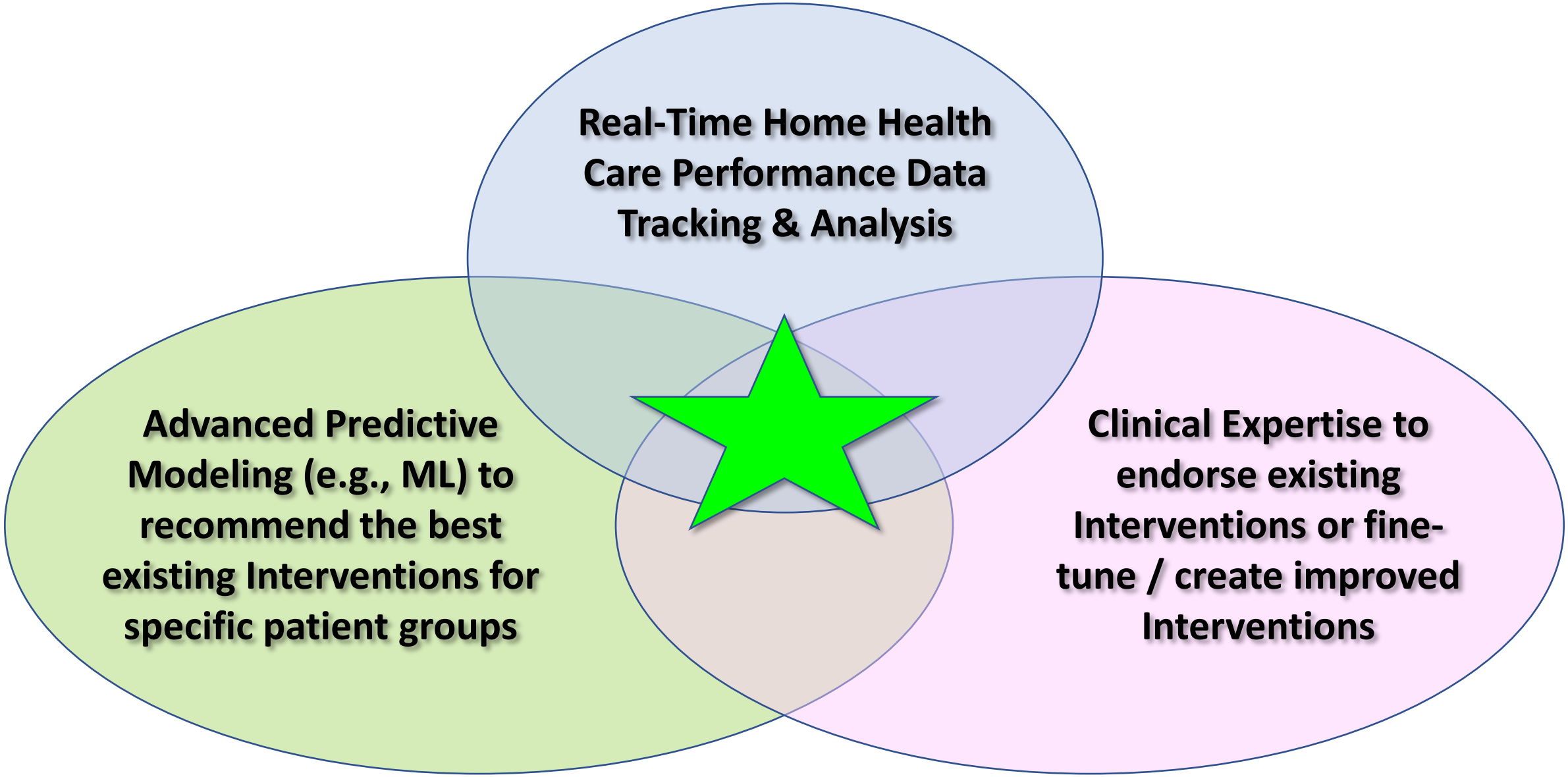
Start



End



Improvement



How Do We make this Star our Collective North Star?