Relative Efficacy of Virtual and In-Office Conservative Care for Musculoskeletal Conditions

Created By:

CWH⁺
ADVISORS

Date: 11/24/2021

Comparative Evaluation of Digital Health Options

As more and more patients seek out virtual care and an increasingly large number of providers seek to accommodate this demand, it is necessary to evaluate its effectiveness relative to traditional, in-office care. However, data with which to carry out this evaluation is scarce due to the short time frame during which virtual care options have been adopted. It is therefore necessary to locate and apply statistical techniques with which to carry out this analysis despite the limited nature of the data.

This white paper presents a technique for using propensity score matching to carry out this evaluation. The results of this analysis among Airrosti patients seeking conservative Musculoskeletal (MSK) care demonstrate that **virtual patients achieve similar positive outcomes to in-office patients**. Combined with previous Airrosti research showing strong positive outcomes for Airrosti in-office patients this work clearly shows that Airrosti solutions provide immense benefits to this patient population.

Overview of Study Design

The study compared a limited set of matched patient data to establish a comparative effectiveness of Airrosti Remote Recovery, Airrosti's virtual care model, as compared to in-office therapy. Our data is limited to cases where key variables are available. We controlled for confounding variables using propensity score matching matched to the same area of injury. This study evaluates the effectiveness of virtual conservative MSK care for upper and lower body injuries on pain improvement, surgical avoidance, injury resolution, and visit completion.

Patient Waterfall

Patients were identified from within Airrosti's internal database of all patients seen in January 2019 through March 2021. This group was then limited to those whose listed insurer was Blue Cross Blue Shield of Texas, in order to ensure access to insurance claims data for later analysis. Within those patients whose injury was a weight-bearing location, we then eliminated patients whose records did not contain useable BMI data. The patient counts are identified in Figure 1 below.

Figure 1 - Patient Episodes

Figure 1 - Patient Episodes
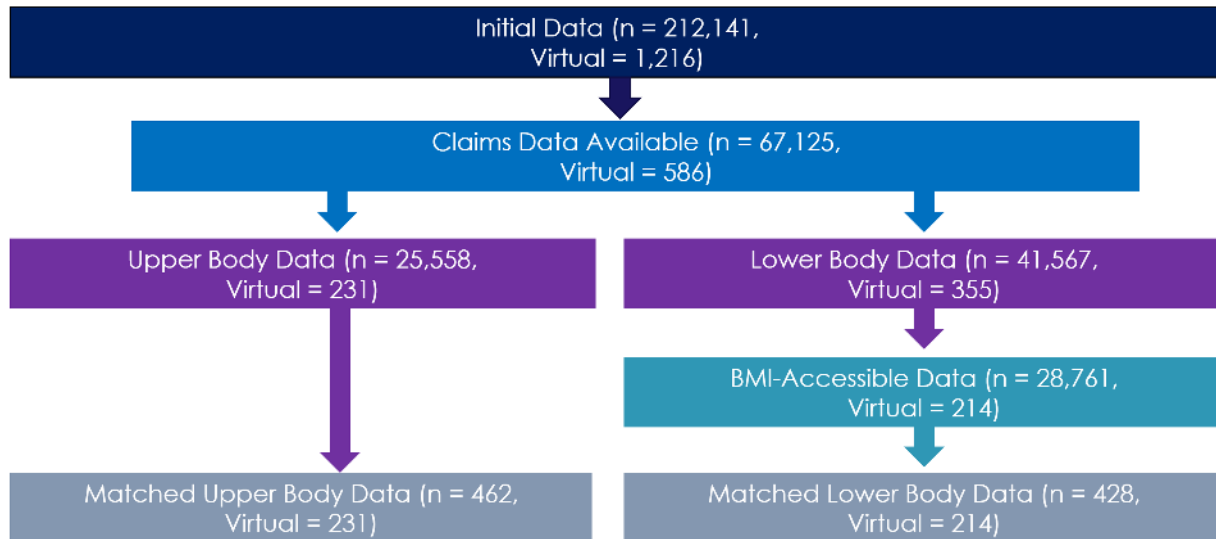
Data

Data was obtained from Airrosti's Raintree EMR system. It contained patient diagnosis, treatment dates, comorbidities, limited data on social determinants of health, and the results of patient surveys before and after receiving Airrosti care.

Categorizing Injury Location

Due to a preponderance of evidence that injuries in different locations have different outcomes[1], virtual subjects were only permitted to match with in-office subjects whose injury was in the same location. Upper body and lower body (I.E, weight-bearing) injuries were analyzed separately, due to the highly disparate effects of BMI on outcomes for those two groups. Within these groups, injuries were categorized into Hip, Lumbar/Sacral, Knee, Ankle/Foot, Thigh, Lower Leg (Lower Body) and Neck, Head, Shoulder, Upper Arm, Thoracic, Elbow, Hand/Wrist (Upper Body).

Outcome Variables

Patient outcomes were evaluated using variables selected from prior reviews of the subject.[7] These outcomes were:
- o Pain Improvement: The difference between initial and final reported pain on a 5-point scale

---

[1] We performed descriptive analyses on each injury location and performed the propensity matching with and without this criterion.

- o Visit Completion: The absolute number of Airrosti visits completed
- o Surgery Avoidance: Whether the patient reported avoiding a considered or scheduled surgery based on their Airrosti results
- o Injury Fixed: Whether the patient reported their injury as 'fixed' in post-therapy surveys.

Methodology

Different groups of people may choose virtual conservative MSK care than in-office. This presents a potential confounder to analysis of the relative effectiveness of the two treatment options.

Propensity score matching addresses these unknown confounders by mimicking the effects of a randomized control trial, matching patients with patients who are similar in variables deemed likely to affect outcomes.[1] Figure 2 below provides an overview of our modeling approach.

Figure 2 - Propensity Model Diagram
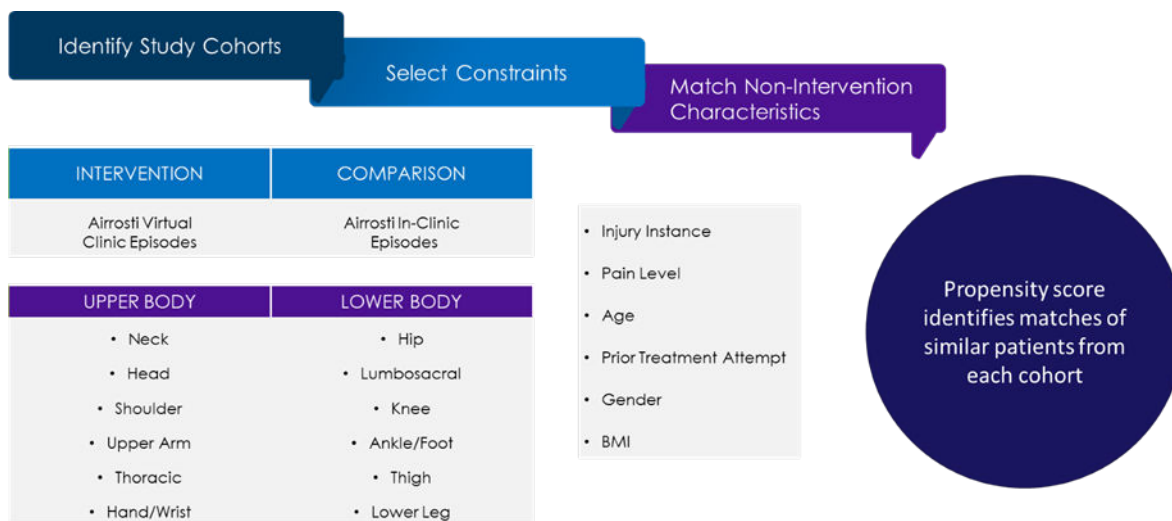
We matched subjects on: Injury location, instance of injury, initial reported pain level (1-5), age, presence or absence of prior treatment attempts, gender, and BMI (for lower-body, weight-bearing injuries only). Matching was accomplished via the optmatch R package implementing the RELAX-IV algorithim.[2] These values were selected based on available data and characteristics controlled for in prior studies of musculoskeletal injury.[3]

Results

Single parametric tests (t-tests) are typically used to compare the degree and significance of differences between two datasets.[5] In this study we want to examine the level of equivalence and answer the question "Are these two treatment modes similar or dissimilar?". We therefore applied a Two One-sided T-Test (TOST) to the matched data, which determines the maximum possible degree and direction of the difference in outcome between two conditions.[6]

Without having access to all conceivable data, it is impossible to prove that the effect of two conditions is exactly equivalent. We therefore defined "equivalence" in this case as 95% confidence that the difference between the mean result for the virtual and in-office conditions was less than one half of one standard deviation.

At this effect size and level of significance (α = .05), there is no statistically meaningful difference between treatment modes (see Figure 3).[8]

Figure 3 - TOST Results

| Variable	Upper Body Result	Lower Body Result
Pain Improvement	Virtual is Equivalent or Superior	Virtual is Equivalent or Inferior
Visit Completion	Virtual is Equivalent or Superior	Virtual is Equivalent or Superior
Surgery Avoidance	Virtual is Equivalent or Superior	Virtual is Equivalent or Superior
Injury Fixed	Virtual is Equivalent	Virtual is Equivalent

TOST Results

A Two One Sided T Test applies a one-sided parametric T-test to each side of the mean, establishing possible overall variance. In Figure 4, negative values represent better relative performance for virtual care.

Figure 4 - Upper Body TOST Results

Pain Improvement   Surgery Avoidance

Visit Completion   Injury Fixed
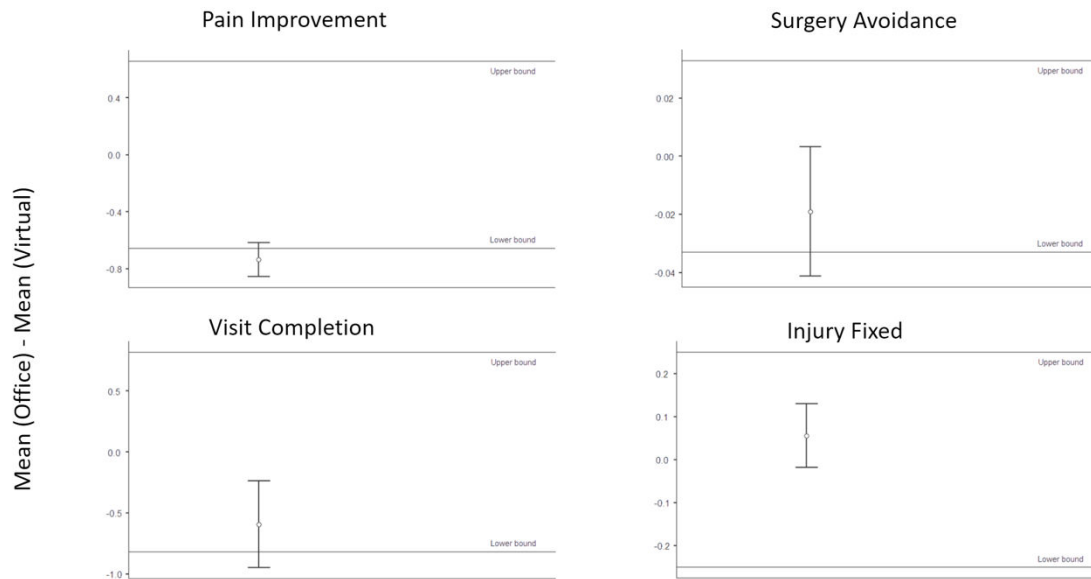
We see improved performance for Virtual in visit completion and pain improvement. The improvement in visit completion is an expected and validating result, since Virtual significantly improves access to care. It is likely that this increased number of visits is the driving force behind the improved performance in pain improvement.

Figure 5 - Lower Body TOST Results

Pain Improvement   Surgery Avoidance

Visit Completion   Injury Fixed
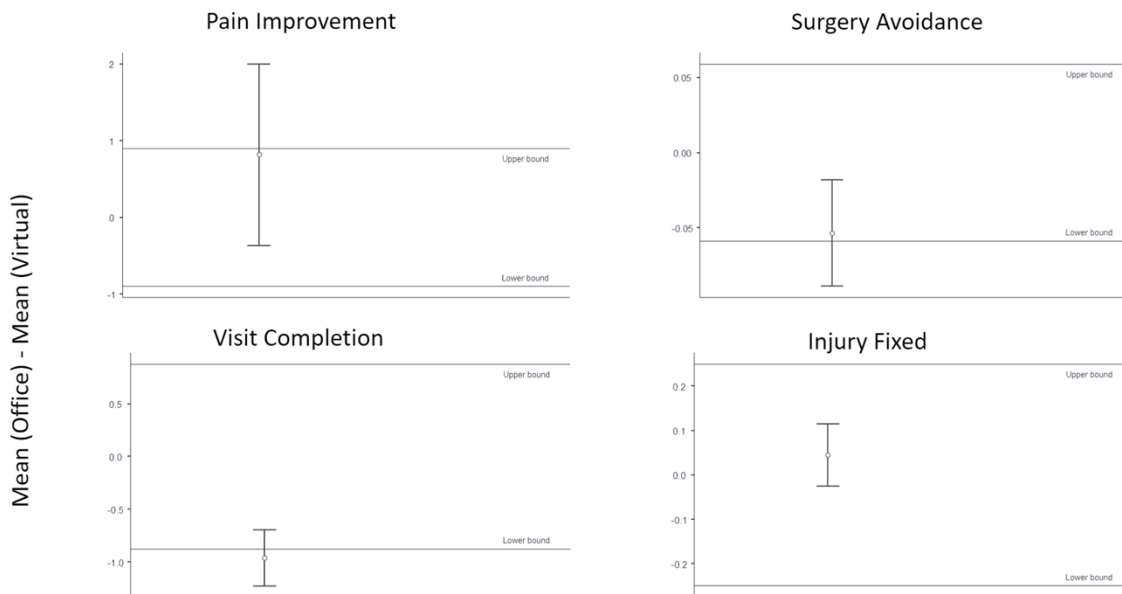
The lower body category (as shown in Figure 5) shows very distinct behavior from upper body with respect to pain improvement. The confidence range crosses the upper bound, indicating a possibility that Virtual performs poorly, but the confidence range is drastically wider than that

of other variables. This suggests that lower body injuries may be split into those that will and will not be apt to respond to virtual conservative MSK care, and that further research could identify which subgroups are well-suited for particular treatment options.

Summary

Independent third-party claims studies from Milliman MedInsight, and Koan Health present the high-level outcomes of Airrosti in-clinic patients.[9] These studies reviewed up to a billion claims and over 2 million episodes of care using Blue Health Intelligence Data. The lack of any clear difference between the virtual and in-office treatment groups indicates that the following metrics can be applied to virtual treatment with an extremely high degree of confidence.

43%
Reduction in Total
Cost of Care

80%
Reduction in
Surgical Utilization

67%
Reduction in High-Tech
Imaging Utilization

55%
Reduction in
Episode Length

The Airrosti Difference

1M+
Patient Cases

99.6%
Patient Satisfaction

88%
Injury Resolution

91%
Prevented Further Medical Services

References

1. Williamson EJ, Forbes A. Introduction to propensity scores. Respirology. 2014 Jul;19(5):625-35. doi: 10.1111/resp.12312. Epub 2014 May 29. PMID: 24889820.
2. Bertsekas, D., and Tseng, P. 1994. ''RELAX-IV: A faster version of the RELAX code for solving minimum cost flow problems.'' Completion Report under NSF Grant CCR-9103804, Department of Electrical Engineering and Computer Science, MIT, Cambridge, Mass
3. Peat G, McCarney R, Croft P. Knee pain and osteoarthritis in older adults: a review of community burden and current use of primary healthcare. Ann Rheum Dis 2001 Feb; 60(2):91-97
4. Andrade C. Mean Difference, Standardized Mean Difference (SMD), and Their Use in Meta-Analysis: As Simple as It Gets. J Clin Psychiatry. 2020 Sep 22;81(5):20f13681. doi: 10.4088/JCP.20f13681. PMID: 32965803.
5. Jankowski KRB, Flannelly KJ, Flannelly LT. The t-test: An Influential Inferential Tool in Chaplaincy and Other Healthcare Research. J Health Care Chaplain. 2018 Jan-Mar;24(1):30-39. doi: 10.1080/08854726.2017.1335050. Epub 2017 Jun 16. PMID: 28622103.
6. Shieh G. Exact Power and Sample Size Calculations for the Two One-Sided Tests of Equivalence. PLoS One. 2016 Sep 6;11(9):e0162093. doi: 10.1371 PMID: 27598468
7. Peat G, McCarney R, Croft P. Knee pain and osteoarthritis in older adults: a review of community burden and current use of primary healthcare. Ann Rheum Dis 2001 Feb; 60(2):91-97
8. Shieh G. Exact Power and Sample Size Calculations for the Two One-Sided Tests of Equivalence. PLoS One. 2016 Sep 6;11(9):e0162093. doi: 10.1371 PMID: 27598468
9. Study completed by Koan Health utilizing Milliman MedInsight developed methodology and BHI claims from BCBSTX paid between Feb. 2012 and Dec. 2020.