Evaluationsbericht (gemäß Nr. 14.1 ANBest-IF)

Konsortialführung:	Charité - Universitätsmedizin Berlin
Förderkennzeichen:	01NVF20014
Akronym:	CFS_CARE
Projekttitel:	Versorgungskonzept für Patienten mit Myalgischer Enzephalomyelitis/Chronischem Fatigue Syndrom (ME/CFS)
Autorinnen und Autoren:	PD Dr. phil. Ulrike Grittner, Lukas Mödl
Förderzeitraum:	1. August 2021 – 31. März 2025
Ansprechpartner:	PD Dr. phil. Ulrike Grittner Institut für Biometrie und Klinische Epidemiologie Charité - Universitätsmedizin Berlin

Das dieser Veröffentlichung zugrundliegende Projekt CFS_CARE wurde mit Mitteln des Innovationsausschusses beim Gemeinsamen Bundesausschuss unter dem Förderkennzeichen 01NVF20014 gefördert. Die Darstellungen im Evaluationsbericht sind das Ergebnis der unabhängigen Evaluation zur neuen Versorgungsform.

Gefördert durch:
GBA Innovationsausschuss

Zusammenfassung

Hintergrund: Myalgische Enzephalomyelitis/Chronisches Fatigue Syndrom (ME/CFS) ist eine schwere, chronische neuroimmunologische Erkrankung, die durch eine erhebliche Einschränkung der Lebensqualität und Funktionsfähigkeit gekennzeichnet ist. Trotz der Schwere und Häufigkeit der Erkrankung existiert bislang keine zugelassene Therapie, die ME/CFS wirksam behandeln kann. Ziel dieser neuen Versorgungsform war es zu prüfen, ob eine innovative, sektorenübergreifende Versorgung die Krankheitsschwere von ME/CFS verbessern kann.

Methodik: Es handelte sich um eine prospektive, monozentrische, nicht randomisierte kontrollierte zweiarmige Interventionsstudie. Patient:innen im Interventionsarm erhielten eine ambulante Betreuung in unseren Spezialambulanzen und eine spezialisierte stationäre Reha, die von einer umfassenden Diagnostik (u.a. schlafmedizinisch, sportmedizinisch, kardiologisch) begleitet wurde. Die Reha fokussierte sich auf das Erlernen von Pacing und Entspannungstechniken, symptomatische Behandlung sowie die Analyse und Planung einer möglichen beruflichen Wiedereingliederung. Die ambulante Betreuung erfolgte im Rahmen von 3-monatlichen Vorstellungen in den Spezialambulanzen. Die Versorgung von Patient:innen im Kontrollarm erfolgte nach einmaliger Vorstellung in unseren Spezialambulanzen über die Hausärzte. Nach 12 Monaten wurde die selbstberichtete körperliche Funktionsfähigkeit (SF-36) in der Primäranalyse sowie subjektive Einschätzungen der Patient:innen (Einschränkungen, Symptome, Fatigue, Schlafstörungen, körperliche Aktivität, Lebensqualität, psychische Gesundheit), objektive Messparameter (Handgriffkraft, Puls, körperliche Aktivität, Laborwerte) und Krankenkassendaten (Kosten, Fehltage) in Sekundäranalysen zwischen Interventions- und Kontrollgruppe verglichen.

Ergebnisse: Bei Studieneinschluss und nach 12 Monaten zeigte sich im SF-36 körperliche Funktionsfähigkeit (primärer Endpunkt) ein Mittelwert von 40,4 bzw. 36,0 Punkten (Standardabweichung: 18,5 und 20,6 Punkte) in der Interventionsgruppe und ein Mittelwert von 40,3 bzw. 41,8 Punkten (Standardabweichung: 19,2 bzw. 25,9 Punkte) in der Kontrollgruppe. Die adjustierte Mittelwertdifferenz von −3,8 an Monat 12 (95-% Konfidenzintervall: [−9,9; 2,2]) entspricht einer etwas schlechteren mittleren körperlichen Funktionsfähigkeit in der Interventionsgruppe im Vergleich zur Kontrollgruppe und war auf dem zweiseitigen 5 %-Signifikanzniveau nicht statistisch signifikant (P-Wert = 0,211). Sensitivitätsanalysen bestätigten die Robustheit dieses Ergebnisses. Auch in den explorativ ausgewerteten sekundären Endpunkten fand sich kein relevanter Unterschied zwischen Interventions- und Kontrollgruppe. Zwar berichteten 76 % in der Reha hilfreiche Konzepte zur Symptomlinderung erhalten zu haben, jedoch war die Bell-Funktionsskala nach der Reha bei 44,7 % niedriger und nur bei 13,8 % höher als vor der Reha.

Diskussion: Es zeigte sich kein signifikanter Unterschied in der SF-36 Subskala körperliche Funktionsfähigkeit zwischen Intervention und Kontrollarm zu Monat 12. Eine symptomorientierte Behandlung und Rehabilitation konnte die Krankheitsschwere bei ME/CFS nicht verbessern. Obwohl das Rehabilitationskonzept darauf abzielte, individuelle Belastungsgrenzen zu berücksichtigen und Pacing umzusetzen, kam es bei fast der Hälfte der Patienten nach der Rehabilitation zu einer klinischen Verschlechterung.

Schlagworte: Versorgungsform, Myalgische Enzephalomyelitis/Chronisches Fatigue Syndrom, ME/CFS, Post-akutes Infektions Syndrom, PAIS, Rehabilitation, SF36.

Gefördert durch: GBA Innovationsausschuss

Inhaltsverzeichnis

I Abkürzungsverzeichnis 6

II Abbildungsverzeichnis 7

III Tabellenverzeichnis 7

1 Ziele der Evaluation 9

1.1 Fragestellung 9

1.2 Darstellung des primären Endpunktes 9

1.3 Darstellung der sekundären Endpunkte 9

1.4 Hypothesen und Endpunkte 9

1.4.1 Primäre Hypothese 9

1.4.2 Primärer Endpunkt 9

1.4.3 Sekundäre Hypothesen 9

1.4.4 Sekundäre Endpunkte 10

2 Darstellung des Evaluationsdesigns 11

2.1 Studiendesign 11

2.2 Einschluss- und Ausschlusskriterien 11

2.3 Allokation der Patient:innen zu den Behandlungsgruppen 11

2.4 Studienablauf 12

2.4.1 Interventionsarm 12

2.4.2 Kontrollarm 13

2.5 Darstellung und Begründung der Ziel- und Vergleichspopulation 14

2.6 Patient:innenperspektive 14

2.7 Drittvariablen und Konfounder 14

2.8 Begründung der Fallzahl/Stichprobengrößen (Power-Berechnungen) 14

2.9 Rekrutierungsstrategie 15

2.10 Subgruppen und Bedeutung der Subgruppenanalysen 15

2.11 Risiko des Drop-Outs der Kontrollgruppe 16

2.12 Darstellung der Datenerfassung und des Archivierungskonzeptes 16

2.13 Erhebungsinstrumente 16

2.14 Studienpopulationen 17

2.15 Statistische Analyse der Hypothesen 17

2.15.1 Haupthypothese 17

2.15.2 Sekundären Hypothesen 18

2.16 Darstellung der Ergebnisse 18

2.17 Statistische Analysen der Subgruppen 19

Gefördert durch: GBA Innovationsausschuss

2.18 Gesundheitsökonomische Auswertung 19
2.19 Prozessevaluation 20
2.20 Umgang mit fehlenden Werten 21
2.21 Transformationen 21

3 Ergebnisse der Evaluation 21
3.1 Studienpopulation 21
3.1.1 CONSORT 21
3.1.2 Baseline-Charakteristika der Patient:innen 24
3.1.3 Baseline-Charakteristika der ausgeschlossenen Patient:innen 30
3.2 Primärer Endpunkt: Körperliche Funktionsfähigkeit (SF-36) 30
3.2.1 Sensitivitätsanalysen 31
3.2.2 Subgruppenanalysen 32
3.3 Sekundäre Endpunkte 33
3.3.1 Bell Disability Scale 34
3.3.2 Chalder Fatigue Scale 34
3.3.3 CCC-Kriterien 35
3.3.4 COMPASS-31 Score 36
3.3.5 SF-36 weitere Skalen 38
3.3.6 Handgriffkraft 41
3.3.7 Schellong-Test (Anlehn-Stehtest) 43
3.3.8 Durchschnittliche Schrittzahl pro Tag 45
3.3.9 IPAQ 46
3.3.10 Blutparameter 47
3.3.11 PHQ-9 50
3.3.12 ESS 51
3.3.13 EQ-5D-5L Indexwert 52
3.4 Analyse der Krankenkassendaten 52
3.4.1 EQ-5D-5L Indexwerte und Qualität-adjustierte Lebensjahre (QALY) 52
3.4.2 Gesundheitskosten im Zeitraum 3 Monate bis 12 Monate 54
3.4.3 Kosteneffizienzanalyse 55
3.4.4 Krankentage im Zeitraum 3 Monate bis 12 Monate 58
3.5 Ergebnisse der Prozessevaluation 58

4 Diskussion der Projektergebnisse 60
4.1 Effektevaluation 60
4.1.1 Primärer Endpunkt: Körperliche Funktionsfähigkeit (SF-36) 60
4.1.2 Sekundäre Endpunkte 62
4.2 Gesundheitsökonomische Evaluation 64
4.3 Limitationen der Studie 64

Gefördert durch: GBA Innovationsausschuss

5 Schlussfolgerungen und Empfehlungen des Evaluators 65

IV Literaturverzeichnis 66

V Anlagen 67

Gefördert durch:
GBA Innovationsausschuss

I Abkürzungsverzeichnis

AM	Arithmetischer Mittelwert
ANCOVA	Analysis of Covariance
BKK	Betriebskrankenkasse
BMI	Body Mass Index
CCC	Kanadische Konsensus Kriterien
COMPASS-31	Composite Autonomic Symptom Score 31
DRV	Deutsche Rentenversicherung
eCRF	elektronische Case-Report-Form
EQ-5D-5L	European Quality of Life 5 Dimensions 5 Level Version
ESS	Epworth Sleepiness Scale
G-BA	Gemeinsamer Bundesausschuss
GM	Geometrischer Mittelwert
IA	Innovationsausschuss
ICER	Incremental cost-effectiveness ratio
IPAQ	International Physical Activity Questionnaire
IQR	Interquartile range (Interquartilsspannweite)
ITT	Intention-to-Treat
ISCED	International Standard Classification of Education
KI	Konfidenzintervall
MAR	Missing at random
MCAR	Missing completely at random
MCID	Minimal Clinically Important Difference
ME/CFS	Myalgische Enzephalomyelitis/Chronisches Fatigue Syndrom
MET	Metabolisches Äquivalent
PEM	Post-exertional Malaise
PMM	Predictive Mean Matching
PHQ-9	Patient Health Questionnaire 9
REDCap	Research Electronic Data Capture
QALY	Quality adjusted life years
SD	Standardabweichung
SF-36	Short-Form 36

Gefördert durch: GBA Innovationsausschuss

II Abbildungsverzeichnis

Abbildung 1: Visualisierung des Studienablaufs 13
Abbildung 2: CONSORT-Diagramm. 23
Abbildung 3: Kosten-Effektivitäts-Ebene mit 95-% Konfidenzellipse. 57
Abbildung 4: Selbsteinschätzung des Rehabilitationserfolgs der Patient:innen. 59
Abbildung 5: Differenz zwischen dem Bell-Score vor und nach der Rehabilitation. 59

III Tabellenverzeichnis

Tabelle 1: Übersicht primärer Endpunkt und sekundäre Endpunkte. 16
Tabelle 2: Baseline-Charakteristika von Patient:innen in Auswertepopulation B und C. 25
Tabelle 3: Ergebnisse der Analyse der körperlichen Funktionsfähigkeit (SF-36). 31
Tabelle 4: Ergebnisse der Sensitivitätsanalysen der körperlichen Funktionsfähigkeit (SF-36). 31
Tabelle 5: Ergebnisse der Subgruppenanalyse nach Geschlecht und Schwere der Erkrankung der körperlichen Funktionsfähigkeit (SF-36). 33
Tabelle 6: Ergebnisse der Analyse der Bell Disability Scale. 34
Tabelle 7: Ergebnisse der Analyse der Chalder Fatigue Scale (mental, körperlich und Gesamtscore). 35
Tabelle 8: Ergebnisse der Analyse der Symptome Muskelschmerz, Kopfschmerzen, nach den Kanadischen Konsensus Kriterien. 36
Tabelle 9: Ergebnisse der Analyse des Composite Autonomic Symptom Score 31 (Blase, Gastrointest-inaltrakt, Orthostase, Pupillomotorik, Sekretomotorik und Vasomotorik und Gesamtscore). 37
Tabelle 10: Ergebnisse der Analyse des Short Form-36 (körperliche Schmerzen, emotionales Wohlbefinden, allgemeine Gesundheitswahrnehmung, Änderung des Gesundheitszustandes, emotionale Rollenfunktion, körperliche Rollenfunktion, soziale Funktion, Vitalität). 39
Tabelle 11: Ergebnisse der Analyse der Handgriffkraft Messungen (Fmean, Fmax, Fatigue Ratio, Recovery Ratio). 41
Tabelle 12: Ergebnisse der Analyse der Schellong-Tests (nach 0, 2, 5 und 10 Minuten Stehen). 44
Tabelle 13: Ergebnisse der Analyse der durchschnittlichen Schrittzahl pro Tag (mittels Troiano-Algorithmus ermittelt). 45
Tabelle 14: Ergebnisse der Analyse des Physical Activity Questionnaire (Gehen/leichte körperliche Aktivität, moderate körperliche Aktivität, anstrengende körperliche Aktivität). 46
Tabelle 15: Ergebnisse der Blutparameter Albumin, anorganisches Phosphat, Calcium, Creatinkinase, Ferritin, Interleukin-8 nach Erylyse, Laktatdehydrogenase und NT-pro-BNP. 48
Tabelle 16: Ergebnisse der Analyse des Patient Health Questionnaires. 51

Gefördert durch: GBA Innovationsausschuss

Tabelle 17: Ergebnisse der Analyse der Epworth Sleepiness Scale. 51
Tabelle 18: Ergebnisse der Analyse des European Quality of Life 5 Dimensions 5 Level Version Indexwerts. 52
Tabelle 19: Ergebnisse der Analyse der EQ-5D-5L-Indexwerte und quality-adjusted life-years (QALYs). 53
Tabelle 20: Ergebnisse der Analyse der monatlichen Gesundheitskosten. 55
Tabelle 21: Ergebnisse der Kosteneffizienzanalyse. 56
Tabelle 22: Krankentage im Zeitraum 3 Monate bis 12 Monate nach Studienbeginn. 58

Gefördert durch: GBA Innovationsausschuss

1 Ziele der Evaluation

Ziel dieser neuen Versorgungsform war es zu prüfen, ob eine innovative, sektorenübergreifende Versorgung die Krankheitsschwere und berufliche Teilhabe von Patienten mit Myalgischer Enzephalomyelitis/Chronischem Fatigue Syndrom (ME/CFS) verbessern kann. Durch ein innovatives sektorenübergreifendes Versorgungskonzept sollte unter Einbeziehung einer Rehabilitation über die Deutschen Rentenversicherung (DRV) diese Versorgungsform entwickelt und evaluiert werden.

1.1 Fragestellung

Führt die neue Versorgungsform zu einer verbesserten körperlichen Funktionsfähigkeit, Symptomschwere und Teilhabe (gemessen mit der körperlichen Funktionsfähigkeit des SF-36) in der Interventionsgruppe im Vergleich zur Kontrollgruppe 12 Monate nach Behandlungsbeginn.

1.2 Darstellung des primären Endpunktes

Der SF-36 körperliche Funktionsfähigkeit (PF) ist der am häufigsten verwendete Endpunkt in klinischen Studien zu ME/CFS. Er erfasst die physische Belastbarkeit im Alltag, die bei Patient:innen mit ME/CFS eingeschränkt ist. Eine ausgeprägte Belastungsintoleranz mit Symptomverschlechterung bereits nach Alltagsbelastungen, die sogenannte post-exertionelle Malaise (PEM), ist das Kern- bzw. Leitsymptom von ME/CFS.

1.3 Darstellung der sekundären Endpunkte

Als sekundäre Parameter wurden Symptome mittels Fragebögen, objektive Messparameter (Handkraft, Puls, Aktivität, Laborwerte) und Arbeitstage und Krankenkassendaten erfasst, wie unten ausgeführt. Von den Betriebskrankenkassen (BKK) der Bahn, der Verkehrsbauunion sowie Siemens wurden die Daten über Krankheitstage (ambulante und stationäre), Arznei- und Hilfsmittel, sowie Kosten in den 12 Monaten vor und während der Studienteilnahme zur Verfügung gestellt. Für Patient:innen der anderen gesetzlichen Krankenkassen standen keine Krankenkassendaten zur Evaluation zur Verfügung. Diese Daten wurden somit nur für die Patient:innen der BKK ausgewertet. Dies betraf jedoch nur drei von 16 sekundären Studienzielen, so dass davon ausgegangen wird, dass die Gesamtaussagekraft der Studie damit nicht beeinträchtigt wurde.

1.4 Hypothesen und Endpunkte

1.4.1 Primäre Hypothese

Patient:innen in der Interventionsgruppe haben 12 Monate nach Studieneinschluss eine bessere körperlichen Funktionsfähigkeit als Patient:innen in der Kontrollgruppe.

1.4.2 Primärer Endpunkt

- Körperliche Funktionsfähigkeit des SF-36 (zu Baseline und Monat 12). Wertebereich: 0-100; 100 ist bester Wert.

1.4.3 Sekundäre Hypothesen

Patient:innen in der Interventionsgruppe haben im Vergleich zur Kontrollgruppe:

Gefördert durch: GBA Innovationsausschuss

- Eine Reduktion des Grades der Behinderung durch ME/CFS (gemessen mit der Bell Disability Scale) nach 12 Monaten.
- Eine Reduktion der Schwere der Fatigue (gemessen mit der Chalder Fatigue Scale) nach 12 Monaten.
- Weniger Symptomschwere (Muskelschmerz, Kopfschmerzen, Konzentrationsprobleme) gemäß der quantifizierten (1-10) Kanadischen Konsensus Kriterien (CCC) für ME/CFS nach 12 Monaten.
- Eine bessere autonome Funktion (gemessen mit der COMPASS-31 Scale) nach 12 Monaten.
- Eine Verbesserung des funktionalen Gesundheitszustands (gemessen mit dem SF-36) nach 12 Monaten.
- Eine höhere Handgriffkraft (Fmean und Fmax in kg) nach 12 Monaten.
- Eine verbesserte Kreislaufsituation (Schellong-Test) nach 12 Monaten.
- Eine Steigerung der durchschnittlichen täglichen Schrittzahl (gemessen mittels Akzelerometer) nach 6 und 12 Monaten.
- Eine Steigerung der körperlichen Aktivität (erfasst mittels IPAQ) nach 12 Monaten.
- Eine Verbesserung bestimmter Blutparameter (Creatinkinase, Laktatdehydrogenase, NT-pro BNP, Albumin, Ferritin, Transferrin-Sättigung, Calcium, Phosphat, Interleukin-8) nach 12 Monaten.
- Weniger depressive Symptome (erfasst mittels PHQ-9) nach 12 Monaten.
- Weniger Schlafstörungen (erfasst mittels ESS) nach 12 Monaten.
- Eine bessere Lebensqualität (erfasst mittels EQ-5D-5L) nach 12 Monaten.
- Weniger Krankenkassenkosten im Zeitraum 3 Monate bis 12 Monate nach Studieneinschluss (nur BKK).
- Weniger Krankentage im Zeitraum 3 Monate bis 12 Monate nach Studieneinschluss (nur BKK).

1.4.4 Sekundäre Endpunkte

- Bell Disability Scale (Bell, 1995).
- Chalder Fatigue Scale (Chalder, et al., 1993).
- CCC (Carruthers, et al., 2003).
- COMPASS-31 Scale (Sletten, Suarez, Low, Mandrekar, & Singer, 2012).
- SF-36 (Ware Jr & Sherbourne, 1992).
- Handgriffkraft (Jäkel, 2021).
- Schellong-Test.
- Schrittzahl.
- IPAQ-Fragebogen (Craig C. L., 2003).
- Blutparameter.
- PHQ-9 (Kroenke, Spitzer, & Williams, 2001).

Gefördert durch: GBA Innovationsausschuss

- ESS (Johns, 1991).
- EQ-5D-5L (EuroQol Group, 1990).
- Nur BKK: Krankenkassenkosten (von 3 bis 12 Monaten nach Studieneinschluss).
- Nur BKK: Krankentage (von 3 bis 12 Monaten nach Studieneinschluss).

2 Darstellung des Evaluationsdesigns

2.1 Studiendesign

Bei dieser Studie handelte es sich um eine prospektive, monozentrische, nicht randomisierte kontrollierte zweiarmige Interventionsstudie.

2.2 Einschluss- und Ausschlusskriterien

Einschlusskriterien waren:

- ME/CFS mit einer Erkrankungsdauer von mindestens 6 Monaten und höchstens 5 Jahren.
- Gesetzliche Versicherung.
- Einwilligungsfähigkeit.
- 18-65 Jahre.
- CCC erfüllt.
- Leistungsfähigkeit nach Bell 30-70.

Ausschlusskriterien waren:

- Andere Erkrankungen, die mit chronischer Fatigue einhergehen z.B. Fatigue nach Krebserkrankung.

2.3 Allokation der Patient:innen zu den Behandlungsgruppen

Die Patient:innen wurden nicht randomisiert der Interventionsgruppe und der Kontrollgruppe zugeordnet. Eine Randomisation war nicht möglich, da nur Patient:innen, die alle Einschlusskriterien erfüllten und gesetzlich und zusätzlich über DRV-Bund versichert waren, für die Interventionsgruppe mit stationärer Rehabilitation in Frage kamen. Patient:innen, die die Einschlusskriterien erfüllten und gesetzlich aber nicht über DRV-Bund versichert waren, wurden in die Kontrollgruppe eingeschlossen. Die Teilnahme am Intervenstionsarm war nur für Patient:innen möglich, die über die DRV-Bund versichert sind, da die Kosten der Reha im Rahmen dieser Studie von der DRV-Bund getragen wurden.

Es wurde jedoch davon ausgegangen, dass diese Einschränkung nicht zu relevanten Verzerrungen wichtiger Basischarakteristika zwischen der Interventions- und Kontrollgruppe führt, da die DRV-Bund seit 2005 allen Versicherungsnehmer:innen offensteht, etwa die Hälfte der BKK Versicherten über die DRV-Bund versichert sind und das Haupterkrankungsalter an ME/CFS zwischen 20 und 50 liegt. Das bestätigte sich in der Auswertung, die in den beiden Gruppen zu Monat 0 einen vergleichbaren SF36 PF zeigte.

Bei der Zuordnung der Gruppen war die von dem Patient:innen angegebene Information ausschlaggebend, ob dieser in der Rentenversicherung Bund (DRV-Bund) oder nicht versichert war/ist. Neben den oben genannten Untersuchungen musste die im Rahmen des CFS_CARE-Projektes konzipierte Reha von dem DRV-Bund genehmigt werden. 35 der ursprünglich in die Interventionsgruppe eingeschlossene Patient:innen mussten deswegen wieder ausgeschlossen werden.

Gefördert durch: GBA Innovationsausschuss

2.4 Studienablauf

Die Patient:innen wurden von niedergelassenen Ärzten überwiesen. Da nach bisherigen Erfahrungen auch viele Patient:innen mit chronischer Fatigue im Zusammenhang mit anderen Erkrankungen überwiesen wurden, erfolgte vor Terminvergabe zunächst eine Erfassung der Symptome und Begleiterkrankungen mittels Fragebogen. Patient:innen bei denen anhand dieser Kriterien der Verdacht auf ME/CFS bestand, erhielten einen Termin. Die anderen Patient:innen erhielten eine Information für den/die überweisenden Arzt/Ärztin. Patient:innen mit Verdacht auf ME/CFS stellten sich in der Spezialambulanz für ME/CFS im Institut für Medizinische Immunologie oder der Ambulanz der Neuroimmunologie zur Diagnostik durch ein spezialisiertes interdisziplinäres Team vor (Klinische Immunologie, Hämatologie, Rheumatologie, Neurologie). Patient:innen, die eine weitergehende differentialdiagnostische Abklärung und Therapieempfehlung benötigten, wurden zeitnah durch die Kooperationspartner gesehen. ME/CFS-Patient:innen erhielten die Empfehlung für ein evidenz-basiertes Therapiekonzept (basierend auf den Empfehlungen des europäischen Netzwerkes EUROMENE) (Nacul, 2021) und eine Bestimmung der körperlichen Leistungsfähigkeit in der Sportmedizin. Patient:innen, die eine chronische Fatigue im Zusammenhang mit einer anderen Erkrankung hatten, erhielten ebenfalls eine Empfehlung für eine Therapie und die Weiterbetreuung durch den/die überweisende Arzt/Ärztin, wurden aber im Rahmen dieser Studie nicht weiterbetreut.

Der (ursprünglich) geplante Ablauf ist in Abbildung 1 dargestellt.

2.4.1 Interventionsarm

ME/CFS-Patient:innen, die über die DRV-Bund versichert sind, wurden für eine Reha in einer spezialisierten Rehaklinik in Kreischa aufgenommen. Die Antragstellung für die Reha erfolgte mit Unterstützung von Sozialarbeiter:innen. Im Rahmen der ersten Vorstellung in den Hochschulambulanzen und vor der Aufnahme in der Rehaklinik erfolgte an der Charité folgende weitergehende Untersuchungen und entsprechende Beratungen:

- Eine schlafmedizinische Untersuchung und Beratung.
- Die Bestimmung der körperlichen Leistungsfähigkeit in der Sportmedizin und eine Beratung bezüglich Aktivitätsgrenzen.
- Die Bestimmung der Endothelfunktion und Herzratenvariabilität in der Kardiologie.
- Eine physikalisch medizinische Untersuchung und Beratung und Therapieempfehlung
- Bei Bedarf eine psychosomatische Vorstellung
- Bei Bedarf eine neurologische Vorstellung

Die Patient:innen erhielten in der Sportmedizin einen Activity Tracker Fitbit Sense®, der neben der Schrittzahl, aktiven Minuten und Herzfrequenz auch die Herzratenvariabilität, Atemfrequenz, Sauerstoffsättigung, elektrodermale Aktivität und Hauttemperatur misst. Dieses Gerät stand den Patient:innen über die ganze Studiendauer zur Verfügung. Ein weiteres Gerät, das Akzelerometer (ActiGraph wgt3x-BT) sollte zu Baseline, Monat 6 und Monat 12 je 9 Tage getragen werden. Die Übergabe an Monat 0 und 12 erfolgte persönlich; an Monat 6 wurde der Tracker den Patient:innen zugeschickt.

Inhalte und Themen der stationären Reha waren:

- Erlernen von Aktivitätsmanagement (Pacing).
- Erlernen von Entspannungstechniken (Mind-Body Therapie).
- Symptomatische Behandlung (Schlaf, Schmerztherapie, Physiotherapie, Lymphdrainage, Kreislaufunterstützung, Mobilitätshilfen).
- Ernährungsberatung.
- Hilfe beim Management sekundärer sozialer und emotionaler Probleme.

Gefördert durch: GBA Innovationsausschuss

- Schulung von Angehörigen.

Während der Reha erfolgte auch die Analyse der Arbeitsfähigkeit und des Arbeitsplatzes und Planung von Wiedereingliederungsmaßnahmen, um die (Teil)-Arbeitsfähigkeit der Patient:innen zu ermöglichen. Wurde eine Erwerbsunfähigkeit festgestellt, so erhielt der/die Patient:in während der Reha und über die Sozialarbeiter:innen eine Beratung und Unterstützung bei der Antragstellung. Die ambulante Betreuung erfolgte für Patient:innen über die Spezialambulanz (Vorstellungen zu Monat 6, 9 und 12). Zudem erfolgte eine Verlaufsbeurteilung mittels Fragebogen zu Monat 3.

2.4.2 Kontrollarm

Die Betreuung im Kontrollarm entsprach der bisherigen Versorgung. Nach Diagnosestellung erfolgte die Weiterbetreuung der Patient:innen, die nicht über die DRV-Bund versichert waren über den/die überweisenden Arzt oder Ärztin. Patient:innen erhielten einen:

- Ausführlicher Arztbericht und entsprechende Therapieempfehlung.

Ärzte und Ärztinnen erhielten:

- Das Angebot einer Fortbildung.
- Telefonische Beratung.

Eine Verlaufsbeurteilung mittels Fragebögen erfolgte zu Monat 3 und 6. Zu Monat 12 erfolgte die Wiedervorstellung in der Spezialambulanz der Charité.

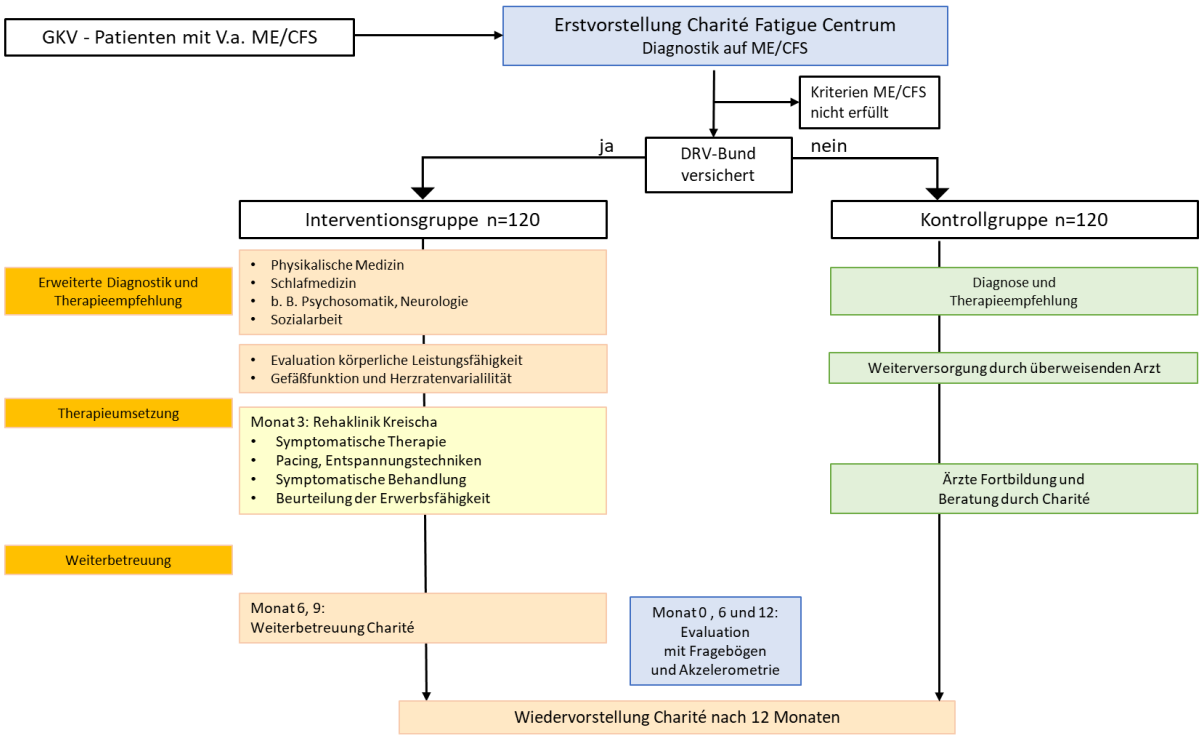


Abbildung 1: Visualisierung des Studienablaufs

Gefördert durch: GBA Innovationsausschuss

2.5 Darstellung und Begründung der Ziel- und Vergleichspopulation

Die Ziel- und Vergleichspopulation bestand aus einwilligungsfähigen, erwachsenen Patient:innen (18-65 Jahre) mit ME/CFS, die gesetzlich versichert sind. Die Diagnose ME/CFS sollte seit mindestens 6 Monaten bestehen und nicht länger als 5 Jahre zurückliegen. In die Interventionsgruppe wurden Patient:innen, die zusätzlich über die DRV-Bund versichert sind, eingeschlossen. In die Kontrollgruppe wurden Patient:innen eingeschlossen, die nicht über die DRV-Bund versichert sind. Der Fokus lag auf Patient:innen, die bei Studieneinschluss noch keine umfangreiche Diagnostik und Rehabilitation erhalten hatten.

2.6 Patient:innenperspektive

Nach Abschluss der Reha erhielten die Patient:innen in der Interventionsgruppe einen Befragungsbogen zur Evaluation der Rehabilitation. Dieser Fragebogen wurde gemeinsam mit der Selbsthilfegruppe Fatigatio und der Deutschen Gesellschaft für ME/CFS erstellt. Informationen über die Studie wurden auch über diese beiden Gruppen an die Mitglieder kommuniziert.

2.7 Drittvariablen und Konfounder

In dieser nicht randomisierten kontrollierten Studie wurden die folgenden wichtige Konfounder erfasst und alle Analysen dafür adjustiert:

- Alter.
- Geschlecht (männlich, weiblich).
- BMI (kg/m^2).
- Krankheitsdauer (Monate).
- Schweregrad der Erkrankung (Bell).
- Bildungsstand (hoch, mittel/gering; nach dem Internationalen Standardklassifikation für das Bildungswesen, ISCED) (UNESCO, 2011).

Ursprünglich war die Berücksichtigung weiterer Kovariaten, wie der Auslöser der Erkrankung und der Beschäftigungsstatus, geplant. Letztlich wurden diese jedoch nicht in der Auswertung berücksichtigt, da jeweils eine extrem ungleiche Kategorienbesetzung vorlag, die zu einer starken Multikollinearität in den Modellen führte.

2.8 Begründung der Fallzahl/Stichprobengrößen (Power-Berechnungen)

Es gibt bislang kaum publizierte Daten zur Evaluation strukturierter spezialärztlicher Versorgung bei Patient:innen mit ME/CFS. Als Grundlage für die Kalkulation dieser Studie dienten daher zwei kleine Studien von Jason, die eine verhaltenstherapeutische Intervention durchführten (Brown, Khorana, & Jason, 2011; Jason, Benton, Torres-Harding, & Muldowney, 2009). Brown et al. verwendete die „Energy Envelope Theory“, wonach Patient:innen, die im Rahmen ihrer individuellen Belastungsgrenzen aktiv sind (Pacing), eher funktionale Verbesserungen erreichten, als Patient:innen, die ihre individuelle Belastungsgrenze überschreiten. In der kleinen Studie mit 17 Patient:innen, die ihre Aktivitäten zu Baseline nicht über ihre Belastungsgrenzen hinaus ausdehnten und 27 Patient:innen, die ihre individuelle Belastungsgrenzen zu Baseline überschritten, zeigten sich Unterschiede in der physischen Gesundheit nach 12 Monaten: in der Untergruppe derer mit weniger Aktivität zu Baseline von einer Effektstärke von 0,49 (Cohen’s d, mittlere Verbesserung: 16,7 vs. 3,6 Punkte, gemeinsame Standardabweichung: 26,6), in der Untergruppe mit höherer Aktivität zu

Gefördert durch: GBA Innovationsausschuss

Baseline von einer Effektstärke von 0,68 (Cohen's d, mittlere Verbesserung 19,3 vs. 1,2, gemeinsame Standardabweichung: 26,6). In dieser Studie wurden etwas geringere Effekte von unter 0,4 (Cohen's d) erwartet. Mit 106 Patient:innen pro Gruppe, lässt sich mittels eines t-Tests für unverbundene Stichproben mit einer Power von 80 % ein Unterschied mit einer Effektstärke von 0,387 oder größer zu einem zweiseitigen Signifikanzniveau von $\alpha = 0,05$ zeigen. Da eine Drop-out Rate von etwa 10 % erwartet wurde, wurden 120 Patient:innen pro Gruppe in die Studie eingeschlossen.

Für die Berechnung der Fallzahl wurde ein zweiseitiger t-Test verwendet, obwohl die Auswertung mittels ANCOVA-Modell erfolgte. Dieses Vorgehen stellt jedoch eine konservative Vorgehensweise für die Schätzung der Fallzahl von ANCOVA-Modellen dar, da ein ANCOVA-Modell mit $(1 - 2\rho) \times n$ Beobachtungen mindestens die gleiche Power wie ein t-Test mit n Beobachtungen hat, wobei ρ der Varianzinflationsfaktor (Korrelation zwischen Baseline-Messung und Follow-up-Messung) ist. Im schlechtesten Fall von $\rho = 0$ führt dies zu einer benötigten Fallzahl gleich dem t-Test (Borm, Fransen, & Lemmens, 2007).

2.9 Rekrutierungsstrategie

Die Patient:innen wurden über die ME/CFS Ambulanz des Instituts für Medizinische Immunologie der Charité rekrutiert. Über die Studie wurden die BKK-Versicherten direkt von der BKK informiert. Auch auf den Seiten des Instituts für Medizinische Immunologie, des Fatigue Centrums und über die Patient:innenorganisationen (Fatigatio, Lost-Voices-Stiftung und Deutsche Gesellschaft für ME/CFS) gab es Informationen zu dieser Studie. Es wurden die für die Studie in Frage kommenden Patienten aus Berlin, Brandenburg und anliegenden Regionen rekrutiert. Patient:innen, bei denen bei der Anmeldung der Verdacht auf ME/CFS bestand (CCC erfüllt), die weniger als 5 Jahre krank waren, Leistungsfähigkeit nach Bell 30-70) erhielten einen Termin in der Ambulanz der Charité zur Diagnostik. Bei ca. 65 % der erwachsenen Patient:innen, die sich an der Charité in den Jahren 2018/19 vorgestellt hatten, wurde die Diagnose ME/CFS gestellt. Bei den übrigen Patient:innen werden nach Ausschlussdiagnostik andere Erkrankungen diagnostiziert oder die Ursache der chronischen Fatigue kann nicht eindeutig geklärt werden. Vor der Erstvorstellung wurde daher zusätzlich mittels eines Fragebogens auch relevante Ausschlusserkrankungen erfasst, gegebenenfalls erfolgte eine weitergehende telefonische Anamnese über die Studienärzte und Kontakt mit dem Hausarzt/der Hausärztin. Es wurde deshalb davon ausgegangen, dass bei ca. 80 % der sich vorstellenden Patient:innen die Diagnose ME/CFS gestellt werden konnte, so dass für die anvisierte Fallzahl von 240 Patient:innen, etwa 300 Patient:innen rekrutiert werden mussten.

2.10 Subgruppen und Bedeutung der Subgruppenanalysen

Die Prävalenz von ME/CFS ist unter Frauen sehr viel höher als unter Männern. Möglicherweise sprechen Frauen daher anders auf die Intervention an als Männer. Ebenso könnte die Schwere der Erkrankung (gemessen anhand der Bell Disability Scale) Ursache für differentielle Interventionseffekte sein. Dementsprechend wurden die folgenden Subgruppen bezüglich möglicher differentieller Interventionseffekte analysiert:

- Geschlecht (Frauen vs. Männer).
- Schweregrad zu Baseline (Bell ≤ 30 vs. Bell ≥ 40).

Weitere Subgruppenanalysen nach Schlafapnoe, Depression, Halswirbelsäuleninstabilität und Ruhetachykardie waren ursprünglich zwar geplant, wurden letztendlich aber nicht realisiert. Die Gründe hierfür waren zum einen, dass die notwendigen Informationen nicht vollständig erhoben wurden bzw. in der Datenbank nicht vorlagen. Weitere Subgruppenanalyse nach Auslöser der Erkrankung (Infekt vs. kein Infekt) waren nicht sinnvoll, da nur bei neun der Auslöser keine Infektion war.

Gefördert durch: GBA Innovationsausschuss

2.11 Risiko des Drop-Outs der Kontrollgruppe

Das Risiko für Drop Outs in der Kontrollgruppe wurde als relativ gering eingeschätzt, da Patient:innen in der Kontrollgruppe eine Weiterbehandlung durch den Hausarzt erhalten und zu Monat 12 eine weitere Evaluation und Beratung an der Charité erhalten sollten.

2.12 Darstellung der Datenerfassung und des Archivierungskonzeptes

Die Daten wurden mittels elektronischem Case-Report-Form (eCRF) im REDCap Datenerfassungssystem erfasst. Mitarbeiter:innen der Clinical Trial Organisation der Charité erstellten den eCRF und monitorierten die Dateneingabe. In den eCRF wurden klinische Daten und anamnestische Daten sowie Fragebogenantworten pseudonymisiert eingegeben. Zusätzlich wurden durch die BKK Routinedaten zur Verfügung gestellt und in der zentralen Studiendatenbank mit den anderen Daten über die Patient:innenpseudonyme verknüpft.

2.13 Erhebungsinstrumente

In

Tabelle 1 ist der primäre Endpunkt und die sekundären Endpunkte sowie die entsprechenden Erhebungszeitpunkte aufgeführt.

Tabelle 1: Übersicht primärer Endpunkt und sekundäre Endpunkte.

Variable	Erhebungs-instrument	Erhebungszeitpunkt				
		Baseline	3 Monate	6 Monate	9 Monate	12 Monate
Primärer Endpunkt						
Funktionale Gesundheit	Körperliche Funktionsfähigkeit (SF-36) [a]	I+K				I+K
Sekundäre Endpunkte						
Funktionale Gesundheit	Körperliche Funktionsfähigkeit (SF-36) [a]		I+K	I+K	I	
Gesundheitsbezogene Lebensqualität	SF-36 (übrige Skalen) [a]	I+K				I+K
Behinderung	Bell [a]	I+K	I	I	I	I+K
Krankheits-charakteristika	CCC [a]	I+K	I	I	I	I+K
Fatigue	Chalder Fatigue Scale [a]	I+K	I	I	I	I+K
Autonome Dysfunktion	COMPASS-31 [a]	I+K	I	I	I	I+K
körperlicher Aktivität	IPAQ [a]	I+K	I	I	I	I+K
Lebensqualität	EQ-5D-5L [a]	I+K	I	I	I	I+K
Psychische Gesundheit	PHQ-9 [a]	I+K				I+K
Schlaf	ESS [a]	I+K				I+K
Objektive Parameter der Körperfunktion	Schritte (Akzelerometrie)	I+K		I+K		I+K
	Handkraft [b]	I+K				I+K
	Schellong-Test [b]	I+K				I+K

Gefördert durch: GBA Innovationsausschuss

Variable	Erhebungs-instrument	Erhebungszeitpunkt				
		Baseline	3 Monate	6 Monate	9 Monate	12 Monate
	Laborwerte [b]	I+K				I+K
Gesundheitsökonomische Kennzahlen (nur von BKK-Versicherungsnehmer:innen)						
	Krankenkassen-kosten [b]					I+K
	Krankentage (nur von BKK) [b]					I+K

I: Interventionsgruppe, K: Kontrollgruppe.
[a] Selbstbericht
[b] Fremdbericht (Erhebende waren nicht verblindet)

2.14 Studienpopulationen

Es wurden die folgenden vier Auswertepopulationen betrachtet:

- Auswertepopulation A: Diese Population umfasste, in Anlehnung an das Intent-to-Treat (ITT) Prinzip, alle Patient:innen, die einem Behandlungsarm zugeordnet wurden, unabhängig von der Adhärenz.
- Auswertepopulation B: Diese Population umfasst alle Patient:innen, die einem Behandlungsarm zugeordnet wurden und diesen auch tatsächlich durchlaufenen haben.
- Auswertepopulation C: Diese Population beinhaltete alle Patient:innen, die die Studie gemäß dem zugewiesenen Protokoll abgeschlossen haben, unter Ausschluss jener, die Lost-to-Follow-up waren, die Studie abgebrochen, den zugewiesenen Behandlungsarm gewechselt haben oder keine Daten für den primären Endpunkt vorlagen.
- Auswertepopulation P: Diese Auswertepopulation entspricht im Wesentlichen Auswertepopulation C; schließt jedoch keine Fälle mit fehlenden Werten im primären Endpunkt aus.

2.15 Statistische Analyse der Hypothesen

2.15.1 Haupthypothese

Für das primäre Zielkriterium wurden die Unterschiede nach 12 Monaten mithilfe von Kovarianzanalyse-Modellen (ANCOVA) ausgewertet. Um potenzielle Störfaktoren (Confounder) zu berücksichtigen, wurden alle Modelle für die Ausgangswerte der jeweiligen Zielkriterien, das Alter, das Geschlecht, den BMI, die Krankheitsdauer zu Baseline, den Schweregrad der Erkrankung (Bell-Skala) und das Bildungsniveau (gemäß der International Standard Classification of Education) adjustiert.

Dem Intention-to-Treat- (ITT) und dem Per-Protocol-Prinzip (PP) folgend wurden verschiedene Analysepopulationen definiert. Die primäre Analyse basierte auf allen Patienten, die die Einschlusskriterien erfüllten und den Gruppen zugewiesen wurden (Population B), was die klinische Wirksamkeit unter realen Bedingungen (*real-world effectiveness*) widerspiegelt.

Gefördert durch: GBA Innovationsausschuss

Zusätzlich wurde eine PP-Analyse durchgeführt, die nur Patienten einschloss, die auswertbare Daten für das primäre Zielkriterium lieferten (Population C).

Als Sensitivitätsanalysen wurde das primäre Zielkriterium anhand von zwei zusätzlichen Ansätzen bewertet: ein strengerer ITT-Ansatz, der alle Patienten wie ursprünglich zugewiesen einschloss, unabhängig vom Beginn des Rehabilitationsprogramms (Population A), und ein weiter gefasster PP-Ansatz, der alle Patienten einschloss, die die Studie abgeschlossen hatten, unabhängig von der Vollständigkeit der Daten zum primären Zielkriterium (Population P). Abbildung 2 stellt diese verschiedenen Populationen dar.

Fehlende Zielkriteriumsdaten im vollständigen Analysesatz (*full analysis set*), welcher als Basis für die ITT-Analyse diente, wurden mittels multipler Imputation durch verkettete Gleichungen (*Multiple Imputation by Chained Equations*) ersetzt. Um Konvergenz zu gewährleisten, wurden nach 50 Iterationen m = 30 vollständige Datensätze generiert. Der Imputationsprozess nutzte das *Predictive Mean Matching*, stratifiziert nach Behandlungsgruppe. Das Imputationsmodell beinhaltete sowohl die ANCOVA-Kovariaten als auch die Ausgangswerte als Prädiktoren. Die Anpassungsgüte (*goodness of fit*) und die Verteilungsstabilität der imputierten Daten wurden mittels Kerndichteschätzungen (*kernel density plots*) bewertet. Wenn für eine Variable mehr als die Hälfte der Datenpunkte fehlte, wurde keine Imputation durchgeführt. Die endgültigen Schätzwerte wurden durch die Zusammenführung (*pooling*) der Ergebnisse aus den 30 imputierten Datensätzen nach den Rubinschen Regeln ermittelt.

2.15.2 Sekundären Hypothesen

Alle sekundären Hypothesen wurden analog zur primären Hypothese und entsprechend der jeweiligen Endpunkte mittels geeigneter statistischer Methoden (ANCOVA, generalisierte lineare Modelle Regressionsmodelle) in Auswertepopulation C analysiert. Sensitivitätsanalysen erfolgten auf den Auswertepopulationen A, B und P. Wie bei der Analyse des primären Endpunkts wurde für die in Abschnitt 2.7 aufgeführten Konfounder adjustiert. Für binäre Endpunkte wurde anstatt des ANCOVA Modells eine logistische Regression verwendet. Die zeitliche Entwicklung der Endpunkte innerhalb der Interventionsgruppe wurde mittels (generalisierter) gemischter linearer Modelle modelliert. Als feste Effekte wurden dabei der Messzeitpunkt (linear) und die in Abschnitt 2.7 aufgeführten Konfounder berücksichtigt. Um für die Messwiederholungen zu kontrollieren, wurden darüber hinaus Individuen spezifische Zufallseffekte („random intercepts“) in das Modell aufgenommen.

Alle sekundären Hypothesen wurden explorativ ausgewertet.

2.16 Darstellung der Ergebnisse

Für alle Endpunkte wurde der Mittelwert (Standardabweichung) zu Baseline und zu Monat 12 sowie die unadjustierte und die um die Kovariate adjustierte Differenz zusammen mit den zugehörigen 95-% Konfidenzintervallen berichtet. Berichtet wurde außerdem der P-Wert des zweiseitigen Tests auf Signifikanz der Mittelwertdifferenz. Zusätzlich wurde das partielle η^2 als Maß der Effektstärke angegeben. Das partielle η^2 quantifiziert den Anteil der Varianz im Endpunkt, der durch den Interventionseffekt erklärt wird, nachdem andere Effekte (Kovariate) berücksichtigt wurden, um die praktische Relevanz des gefundenen Unterschieds zu beurteilen. Für die Entwicklung über die Zeit wurden die adjustierten marginalen Mittelwerte zu jedem Zeitpunkt mit 95-% Konfidenzintervall dargestellt.

Gefördert durch: GBA Innovationsausschuss

Die Werte wurden je nach Größenordnung auf eine oder zwei Nachkommastellen gerundet. P-Werte und η^2 wurden immer auf drei Nachkommastellen gerundet.

2.17 Statistische Analysen der Subgruppen

Für die statistischen Analysen der Subgruppen wurde im ANCOVA-Modell zusätzlich ein Interaktionsterm des Interventionsgruppenindikator mit der Subgruppe aufgenommen, um Hinweise für differentielle Interventionseffekte in den Subgruppen zu analysieren. Die Unterschiede des Interventionseffekt zwischen den Subgruppen wurden zusammen mit dem entsprechenden 95-% KI angegeben.

Alle Subgruppenanalysen sind explorativ und erfolgten nur für den primären Endpunkt in Auswertepopulation C.

2.18 Gesundheitsökonomische Auswertung

Über den Interventionszeitraum wurden die Gesamtkosten der Behandlung über die Krankenkassendaten für die eingeschlossenen BKK-versicherten Patient:innen gesammelt und ausgewertet. Es wurden sämtliche stationäre und ambulante Abrechnungsdaten berücksichtigt, die im Erhebungszeitraum abzüglich der ersten drei Monate anfielen. Dazu gehörten Krankenhausleistungen (Dauer, Diagnosen, Prozeduren, Kosten), Rehabilitationsleistungen (Art der Leistung, Diagnose, Verordnungsdatum, Dauer, Kosten), Leistungen im Sinne der Arzneimittel-, Heilmittel-, Hilfsmittel-, Fahrtkosten-, häusliche Krankenpflege-, Haushaltshilfeverordnungen, aber ebenso vertragsärztliche Leistungen (Datum, Fachgruppe des Arztes oder der Ärztin, Diagnose, Einzelleistungen, Kosten), sowie Leistungen der sozialen Pflegeversicherung gemäß SGB XI. Anschließend wurden die ambulanten und die stationären Gesamtkosten ins Verhältnis zur jeweiligen Beobachtungszeit gesetzt (mittlere Behandlungskosten pro Monat und Patient:in). Die so ermittelten individuellen monatlichen Durchschnittswerte für ambulante und stationäre Gesundheitskosten wurden im nächsten Schritt für jede Patient:in zu einem mittleren Gesamtposten pro Monat aufsummiert. Zusätzlich wurden die Kosten für die neue Versorgungsform ermittelt.

Wie bei den primären und sekundären Endpunkten wurde ein ANCOVA-Modell verwendet, um den Unterschied der Kosten zwischen den Gruppen zu vergleichen. Das Modell wurde dabei für Geschlecht, Alter, BMI, Schweregrad der Erkrankung zu Studienbeginn, Krankheitsdauer zu Studienbeginn, Bildung adjustiert. Weiterhin wurde überprüft, inwiefern sich die Intervention als überlegen in Bezug auf „quality adjusted life years“ (QALYs) erweist. Ein QALY berücksichtigt sowohl Lebenszeit als auch Lebensqualität. QALYs wurden auf Grundlage von Lebensqualitätsfragebögen (EQ-5D-5L) berechnet, die zu Baseline und zu nachfolgenden Zeitpunkten (Monat 3, 6, 9, 12) während der Projektlaufzeit ausgefüllt wurden. Zwischen den Messzeitpunkten wurde eine lineare zeitliche Veränderung der Lebensqualität angenommen. QALYs ergeben sich dann aus der Berechnung der Fläche unter der Kurve. Falls die Intervention zu besseren Lebensqualitätswerten bei gleichzeitig geringeren Kosten führt, gilt sie als kosteneffektiv. Außerdem wurden die „incremental cost-effectiveness ratio“ (ICER) berechnet, indem die mittleren zusätzlichen Kosten für je zusätzlichem QALY geschätzt wurden. Eine Grenze von 50.000 € pro extra QALY wird in Studien oft als Grenzwert für die Kosteneffektivität einer Intervention verwendet. Liegt das ICER unter dieser Grenze, gilt die Intervention als kosteneffektiv.

Gefördert durch: GBA Innovationsausschuss

Zusätzlich werden die Krankentage (nur von BKK) im Projektzeitraum (abzüglich der ersten drei Monate) zwischen den beiden Gruppen verglichen. Dafür wurde ebenfalls ein um oben genannten Konfounder adjustiertes ANCOVA-Modell verwendet.

Für die gesundheitsökonomische Auswertung wurden erste Effekte der Intervention mit Abschluss der Rehabilitation erwartet. Da es sich um kumulierte Größen (Kosten bzw. Krankentage) handelt, wurde der Erhebungszeitraum exklusive der ersten drei Monate gewählt. Um die Vergleichbarkeit zu gewährleisten, wurde in der Kontrollgruppe analog verfahren.

2.19 Prozessevaluation

Zur Evaluation der Reha wurde eine freiwillige Befragung durchgeführt. Der Befragung beinhaltete fünf Blöcke mit folgenden Themen:

- Vor der Rehabilitation.
- Durchführung der Rehabilitation (Pacing und Alltagsgestaltung).
- Informationsvermittlung zu ME/CFS und Therapiestrategien
- Erkenntnisgewinn durch die Rehabilitation.
- Selbstbeurteilung des Rehabilitationserfolges.

Der erste Block stellte Fragen über die Voruntersuchungen, ob sich die Patient:innen beispielsweise ausreichend über das Reha-Konzept informiert fühlten, spezifisch der Nutzen der schlafmedizinischen, der physiotherapeutischen, der sportmedizinischen und der psychosomatischen Untersuchung, sowie der Gefäßdiagnostik. Im Verlauf wurde erkannt, dass Post-Exertional Malaise (PEM) noch besser charakterisiert werden sollte, weshalb der zweite Block noch einen PEM-Teil erweitert wurde. Im zweiten Block wurden Fragen zu Pacing gestellt, ob Pacing als Strategie angewendet wurde, ob ein Grundniveau zur Bewältigung der Reha Programmpunkte gelungen ist, ob es zu PEM gekommen sei, ob die Patient:innen an allen Veranstaltungen teilnehmen konnten, ob es ausreichend Erholungszeit nach der Veranstaltung gab, ob es als Druck empfunden wurde, das Reha-Pensum zu absolvieren und ob die FITbit-Uhr bei den Aktivitätsgrenzen geholfen hat. Im Teilblock Alltagsgestaltung wurde über die Eignung der Unterbringung, der Ernährung gefragt, die Möglichkeit des Kontaktes zu gleich oder ähnlich an ME/CFS-Erkrankten, und ob sich die Patient:innen wahr- und ernstgenommen fühlten. Der in einer zweiten Version ergänzte spezifische Punkt PEM, mit einer Erklärung sowohl zur Abkürzung als auch was PEM beinhaltet, wurde auf Häufigkeit, Schweregrad, Zeitpunkt, Dauer und Symptome erweitert. Der vierte Block erfasste einige Punkte der Reha wie zum Beispiel Informationen über ME/CFS, Informationen und Schulung über Pacing und die Anwendung von Pacing-Strategien, Vermittlung oder Anwendung von Übungen zu Konzentration, Atmungs- und Entspannungstechniken, Behandlungsstrategien für Schmerzen, Schlafstörungen, Kreislaufstörungen, ob eine Psychotherapeutische, sozialmedizinische Beratung. Im vierten Block ging es um Erkenntnisgewinn; zum Beispiel ob sich die Patient:innen über die Erkrankung besser informiert fühlen, ob sie die Erkrankung besser bewältigen können, ob sie Pacing-Strategien und/oder Entspannungstechniken erfolgreich anwenden können etc. Im fünften Block wurden Fragen zur Verbesserung der Gestaltung des Alltags, des Zustands bezogen auf die Erkrankung durch die Rehabilitation, Vereinbarung der Anforderungen von Beruf oder Ausbildung und ob die Patient:innen zu einem späteren Zeitpunkt an einer erneuten stationären Rehabilitation für ME/CFS teilnehmen würden.

Die Auswertung erfolgt rein deskriptiv anhand absoluter und relativer Häufigkeiten.

Gefördert durch: GBA Innovationsausschuss

2.20 Umgang mit fehlenden Werten

Bei fehlenden Werten, bei denen die Annahme, dass sie „missing at random“ (MAR) oder „missing completely at random“ (MCAR) sind, vertretbar war, wurden Werte mittels Predictive Mean Matching (PMM) aus vorhandenen Daten imputiert. Neben PMM wurde für fehlende Werte in Schul- und Ausbildung eine multinomiale Regression verwendet. Die Imputationsmodelle berücksichtigten alle Kovariate aus dem ANCOVA Modell sowie, falls vorhanden, die übrigen Messzeitpunkte des jeweiligen Endpunkts. Die Imputation erfolgte stratifiziert nach Interventions- und Kontrollgruppe. Aufgrund der Datenstruktur war eine sinnvolle Imputation der BKK-Daten nicht möglich.

Die Schätzer wurden anschließend mittels Rubins Regeln (Rubin, 1987) über m = 30 imputierte Datensätze aggregiert, wobei für jede Imputation 50 Iterationen verwendet wurden, um die Wahrscheinlichkeit für Modellkonvergenz zu erhöhen. Die Qualität der imputierten Daten wurde abschließend über Kerndichtschätzer bzw. Dotplots visualisiert und evaluiert.

In allen Auswertepopulationen (A, B, C, P) wurden fehlende Werte imputiert, wobei in Auswertepopulation C ausschließlich die Imputation von Drittvariablen erfolgte; die Anzahl der imputierten Datensätze je Endpunkt und Auswertepopulation ist in Anlage 6 dargestellt.

2.21 Transformationen

Um eine möglichst gute Approximation durch die Normalverteilung zu erreichen, wurden Transformationen (zum Beispiel Logarithmus oder allgemeiner Box-Cox) in Betracht gezogen. Die Notwendigkeit einer Transformation wurde aus Kerndichteschätzern der ANCOVA-Residuen abgeleitet. War die Verteilung der Residuen sehr schief, musste von einer schlechten Approximation durch die Normalverteilung ausgegangen werden, was die Qualität der Schätzer für Konfidenzintervalle und P-Werte unterminiert. Der Grad der Symmetrie der Residualverteilungen wurde mittels Kerndichteschätzer visualisiert und evaluiert.

3 Ergebnisse der Evaluation

3.1 Studienpopulation

3.1.1 CONSORT

Abbildung 22 zeigt den Pfad der Patient:innen durch die Studie.

Bei insgesamt 247 Fällen bestand die Einschlussfähigkeit auf Basis einer bestätigten ME/CFS-Diagnose sowie der Erfüllung aller Einschlusskriterien. Neun Fälle wurden ausgeschlossen, davon sieben aufgrund von Screening-Failures, ein Fall aufgrund eines doppelten Einschlusses und ein Fall aufgrund des Widerrufs der Einwilligungserklärung. Nach der Allokation erfolgte ein Rückzug der Einwilligung mit der Aufforderung der Löschung der Daten. Insgesamt wurden ursprünglich 238 Fälle den beiden Studiengruppen zugeteilt, davon zunächst 137 der Interventionsgruppe und 101 der Kontrollgruppe. Dementsprechend umfasste die Auswertepopulation A insgesamt 237 Patient:innen; davon 100 in der Kontrollgruppe und 137 in der Interventionsgruppe.

Vor Beginn der Intervention mussten 35 Fälle (25,5 %) aus der Interventionsgruppe ausgeschlossen werden. Hauptgründe hierfür waren die Ablehnung der vorgesehenen Intervention (11 Fälle), eine nicht gegebene Zuständigkeit der DRV (12 Fälle) sowie den Verlust der Reha Fähigkeit infolge einer Zustandsverschlechterung (7 Fälle) oder Notwendigkeit einer Gehhilfe (3 Fälle). Bei zwei Fällen blieb der Grund unbekannt. In der Interventionsgruppe

Gefördert durch: GBA Innovationsausschuss

nahmen letztendlich 97 Patient:innen die Reha in Anspruch. Demnach befanden sich in Auswertepopulation B 97 Fälle in der Interventions- und 100 Fälle in der Kontrollgruppe.

Nach der Interventionsphase reduzierten sich die Fallzahlen in der Interventionsgruppe um zwei Patienten (ein Ausschluss wegen Studien-Doppelteilnahme, ein Widerruf der Einwilligung). Die Kontrollgruppe verzeichnete einen Ausfall durch ein „Lost to follow-up“. Die finale Drop-out-Quote war damit ausgehend von Population A in den Armen sehr ungleich verteilt: Während in der Interventionsgruppe 98,0 % der Patient:innen (99/101) die Studie beendeten, lag die Komplettierungsrate in der Kontrollgruppe bei 69,3 % (95/137).

Für die Auswertepopulation C wurde von allen Patient:innen, die die Studie erfolgreich abschlossen noch alle Fälle mit mit fehlendenden Werten im primären Endpunkt ausgeschlossen (jeweils 6 Fälle in Interventions- und Kontrollgruppe). Auswertepopulation C umfasste damit insgesamt 182 Patient:innen; davon 89 in der Interventionsgruppe und 93 in der Kontrollgruppe.

Gefördert durch: GBA Innovationsausschuss

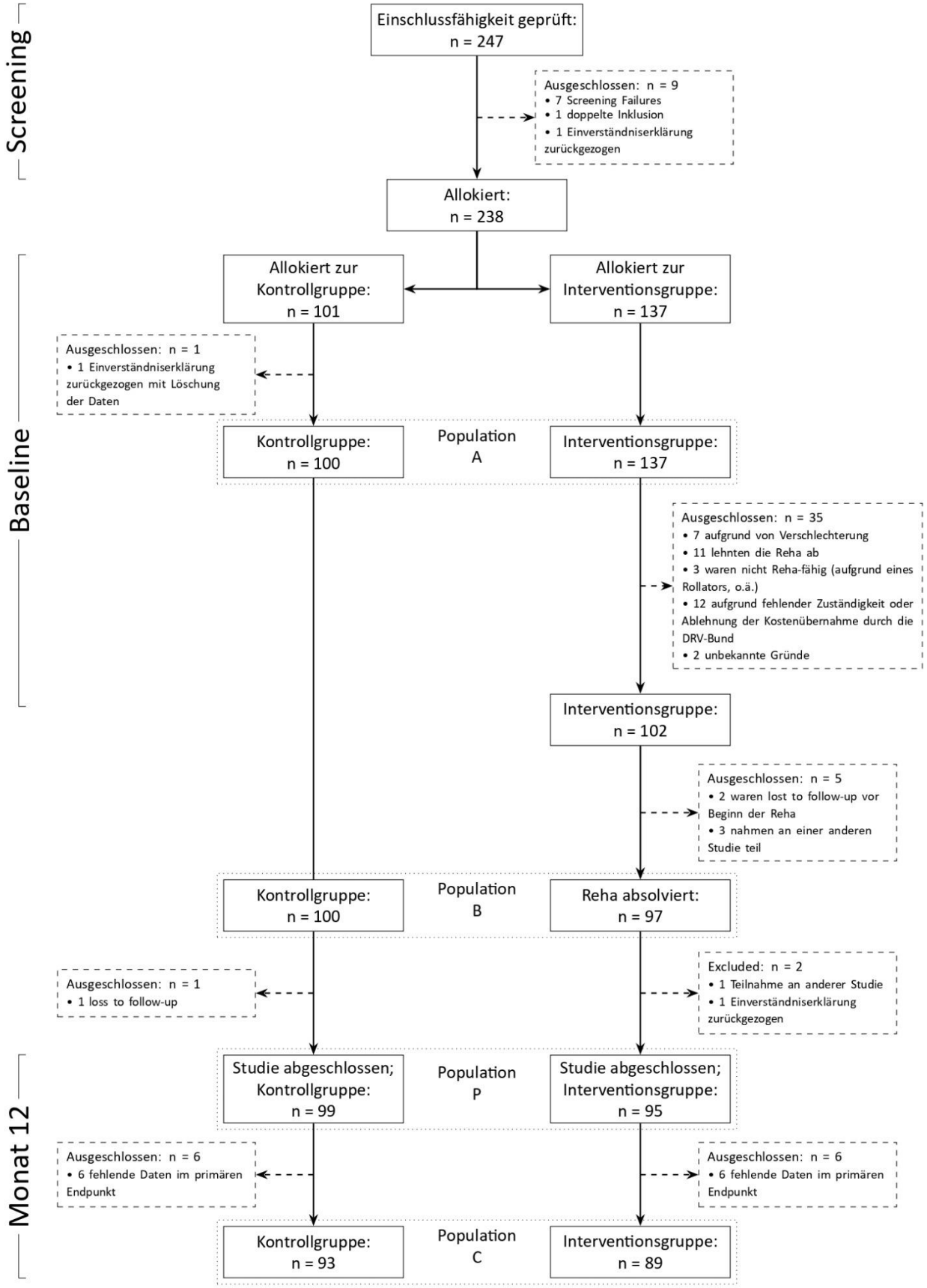


Abbildung 2: CONSORT-Diagramm.

Gefördert durch:
GBA Innovationsausschuss

3.1.2 Baseline-Charakteristika der Patient:innen

Tabelle 2 stellt die Baseline-Charakteristika der Patient:innen in den Auswertepopulationen B und C dar. Die verschiedenen Auswertepopulationen unterschieden in den Baseline-Charakteristika insgesamt nur minimal, weshalb sich im Folgenden aus Auswertepopulation B beschränkt wird.

Beide Gruppen wiesen ein demografisches Profil auf, das typisch für ME/CFS-Patient:innen ist: Das Durchschnittsalter lag bei etwa 42 Jahren, wobei die Interventionsgruppe im Schnitt 5 Jahre älter war als die Kontrollgruppe (Kontrollgruppe: 39,3 Jahre mit Standardabweichung: 11,1 Jahre; Interventionsgruppe: 44,0 Jahre mit Standardabweichung: 9,5 Jahre). Die Verteilung der Geschlechter war mit über 80 % Frauen in beiden Gruppen sehr ähnlich und spiegelt die bekannte höhere Prävalenz weiblicher ME/CFS-Patientinnen wider. Während in der Interventionsgruppe der Frauenanteil überwog (88,7 % vs. 83,0 %), war der Anteil männlicher Teilnehmer in der Kontrollgruppe geringfügig höher (17,2 % vs. 11,3 %). Der BMI war im Mittel in beiden Gruppen nahezu identisch bei 25,1 kg/m^2 (Standardabweichung: etwa 5,1 kg/m^2) und befand sich damit im oberen Normalbereich/knapp im Übergewicht.

Ein Großteil der Teilnehmenden gab an, nie oder lediglich in der Vergangenheit geraucht zu haben; beide Gruppen waren in dieser Hinsicht aber relativ ähnlich. Die Verteilungen der wöchentlichen Konsumfrequenz von Alkohol war in beiden Gruppen ebenfalls relativ ähnlich.

Die Verteilung der selbstberichteten Begleiterkrankungen war zwischen beiden Gruppen weitgehend ausgewogen. Am häufigsten wurden Allergien (56,9 %), pulmonale Erkrankungen (23,7 %) sowie endokrine Störungen (20,8 %) berichtet. Während die Kontrollgruppe einen höheren Anteil an Allergien und psychischen Störungen aufwies, waren in der Interventionsgruppe nephrologische Erkrankungen häufiger vertreten. In allen übrigen Kategorien zeigten sich die Gruppen relativ homogen.

In beiden Gruppen waren die meisten Patient:innen ledig (47,3 %) oder verheiratet (43,6 %). Die Verteilung zwischen den Gruppen war aber relativ ähnlich. Hinsichtlich des Anteils an Kindern zeigten sich die beiden Gruppen nahezu deckungsgleich.

Das durchschnittliche Bildungsniveau (gemessen am ISCED) lag in beiden Gruppen primär im mittleren bis hohen Bereich. Dabei wies die Kontrollgruppe eine stärkere Konzentration im mittleren Bereich auf, während in der Interventionsgruppe der Anteil mit hohem Bildungsabschluss überwog. Die Mehrheit der Patient:innen in beiden Gruppen war in einem Beschäftigungsverhältnis. In der Kontrollgruppe waren 90,8 % der Patient:innen erwerbstätig, in der Interventionsgruppe waren es 92,4 %. Entsprechend gering fiel der Anteil der Nicht-Erwerbstätigen aus (9,2 % in der Kontrollgruppe vs. 7,6 % in der Interventionsgruppe). Die beiden Gruppen unterschieden sich hinsichtlich ihres Erwerbsstatus bei Studieneintritt somit praktisch nicht.

Der durchschnittliche Schweregrad (gemäß der Bell Disability Scale) betrug in beiden Gruppen 36,6 Punkte (Standardabweichung: 7,7 Punkte), was klinisch einer moderat bis schweren Beeinträchtigung entspricht. Die Patient:innen waren im Durchschnitt etwa 1,5 Jahre erkrankt, wobei die Kontrollgruppe mit 21,1 Monaten (Standardabweichung: ca. 14,0 Monate) im Durchschnitt etwas länger erkrankt war als die Interventionsgruppe mit 17,0 Monaten (Standardabweichung: 9,3 Monate). Bei über 90 % der Fälle wurde die Krankheit durch eine Infektion ausgelöst.

Gefördert durch: GBA Innovationsausschuss

Tabelle 2: Baseline-Charakteristika von Patient:innen in Auswertepopulation B und C.

Auswertepopulation B	Gesamt N = 197	Kontrollgruppe N = 100	Interventionsgruppe N = 97	SMD a
Patient:innendaten				
Alter [Jahre], MW (SD)	41,6 (10,5)	39,3 (11,1)	44,0 (9,5)	0,4
Body Mass Index [kg/m2], MW (SD)	25,1 (5,1)	25,0 (5,0)	25,2 (5,3)	0,0
Geschlecht, n (%)				0,1
weiblich	169 (85,8 %)	83 (83,0 %)	86 (88,7 %)	
männlich	28 (14,2 %)	17 (17,0 %)	11 (11,3 %)	
Alkohol und Rauchen				
Derzeitiger Alkoholkonsum, n (%)				0,1
kein derzeitiger Alkoholkonsum	79 (41,6 %)	37 (39,4 %)	42 (43,8 %)	
1x pro Woche	65 (34,2 %)	36 (38,3 %)	29 (30,2 %)	
2x pro Woche	37 (19,5 %)	16 (17,0 %)	21 (21,9 %)	
3x pro Woche	9 (4,7 %)	5 (5,3 %)	4 (4,2 %)	
Keine Angabe	7	6	1	
Rauchen, n (%)				0,1
Nichtraucher	115 (59,6 %)	56 (57,1 %)	59 (62,1 %)	
Ehemalige:r Raucher:in	61 (31,6 %)	33 (33,7 %)	28 (29,5 %)	
Derzeitige:r Raucher:in	17 (8,8 %)	9 (9,2 %)	8 (8,4 %)	
Keine Angabe	4	2	2	
Komorbiditäten (Patient:innenangabe), n (%)				
Allergien	112 (56,9 %)	51 (51,0 %)	61 (62,9 %)	0,2
Kardiovaskuläre Erkrankungen 1	30 (15,2 %)	17 (17,0 %)	13 (13,4 %)	0,1
Gastrointestinale Erkrankungen 2	17 (8,6 %)	8 (8,0 %)	9 (9,3 %)	0,0
Erkrankungen des Urogenitaltraktes 3	14 (7,1 %)	10 (10,0 %)	4 (4,1 %)	0,2
Endokrine Erkrankungen 4	41 (20,8 %)	22 (22,0 %)	19 (19,6 %)	0,1
Muskuloskelettale Erkrankungen 5	11 (5,6 %)	7 (7,0 %)	4 (4,1 %)	0,1
Neurologische Störungen 6	16 (8,1 %)	9 (9,0 %)	7 (7,2 %)	0,1

Gefördert durch: GBA Innovationsausschuss

Auswertepopulation B	Gesamt N = 197	Kontrollgruppe N = 100	Interventionsgruppe N = 97	SMD a
Pulmonale Erkrankungen 7	46 (23,4 %)	23 (23,0 %)	23 (23,7 %)	0,0
Hauterkrankungen 8	11 (5,6 %)	6 (6,0 %)	5 (5,2 %)	0,0
Psychische Störungen 9	20 (10,2 %)	7 (7,0 %)	13 (13,4 %)	0,2
Andere 10	22 (11,2 %)	14 (14,0 %)	8 (8,2 %)	0,2
Familie und Kinder				
Familienstand, n (%)				0,1
Verheiratet	82 (43,6 %)	42 (43,8 %)	40 (43,5 %)	
Getrennt	17 (9,0 %)	6 (6,2 %)	11 (12,0 %)	
Ledig	89 (47,3 %)	48 (50,0 %)	41 (44,6 %)	
Keine Angabe	9	4	5	
Kinder, n (%)				0,0
Ja	59 (31,9 %)	31 (33,0 %)	28 (30,8 %)	
Nein	126 (68,1 %)	63 (67,0 %)	63 (69,2 %)	
Keine Angabe	12	6	6	
Ausbildung und Beschäftigung				
Ausbildung, n (%)				0,1
Kein Abschluss, noch in Ausbildung	12 (6,6 %)	7 (7,4 %)	5 (5,7 %)	
Beruflich-betrieblich	53 (29,1 %)	27 (28,4 %)	26 (29,9 %)	
Beruflich-schulisch, Fachschule	26 (14,3 %)	12 (12,6 %)	14 (16,1 %)	
Fachhochschule, Universität	88 (48,4 %)	48 (50,5 %)	40 (46,0 %)	
Andere	3 (1,6 %)	1 (1,1 %)	2 (2,3 %)	
Keine Angabe	15	5	10	
Internationale Standardklassifikation für das Bildungswesen (ISCED), n (%)				0,2
Niedrig	6 (3,2 %)	5 (5,1 %)	1 (1,1 %)	
Mittel	80 (42,6 %)	38 (38,8 %)	42 (46,7 %)	
Hoch	102 (54,3 %)	55 (56,1 %)	47 (52,2 %)	
Keine Angabe	11	2	9	
Beschäftigungsstatus, n (%)				0,0
Beschäftigt	174 (91,6 %)	89 (90,8 %)	85 (92,4 %)	

Gefördert durch: GBA Innovationsausschuss

Auswertepopulation B	Gesamt N = 197	Kontrollgruppe N = 100	Interventionsgruppe N = 97	SMD a
Nicht beschäftigt	16 (8,4 %)	9 (9,2 %)	7 (7,6 %)	
Keine Angabe	7	2	5	
Krankheitsmerkmale				
Schweregrad [Bell], MW (SD)	36,6 (7,7)	36,5 (7,8)	36,8 (7,6)	0,0
Dauer [Monate], MW (SD)	19,1 (12,1)	21,1 (14,0)	17,0 (9,3)	0,3
Auslöser (Trigger), n (%)				0,1
Covid-19	155 (79,5 %)	77 (78,6 %)	78 (80,4 %)	
Epstein-Barr-Virus	7 (3,6 %)	5 (5,1 %)	2 (2,1 %)	
Andere Atemwegsinfektionen	19 (9,7 %)	9 (9,2 %)	10 (10,3 %)	
Andere Infektion	5 (2,6 %)	3 (3,1 %)	2 (2,1 %)	
Nicht infektiös	9 (4,6 %)	4 (4,1 %)	5 (5,2 %)	
Unbekannt, keine Angabe	2	2	0	

Auswertepopulation C	**Gesamt N = 182**	**Kontrollgruppe N = 93**	**Interventionsgruppe N = 89**	**SMD [a]**
Patient:innendaten				
Alter [Jahre], MW (SD)	41,8 (10,4)	39,3 (10,9)	44,4 (10,9)	0,5
Body Mass Index [kg/m^2], MW (SD)	25,2 (5,2)	25,0 (4,9)	25,4 (5,5)	0,1
Geschlecht, n (%)				0,1
weiblich	158 (86,8 %)	78 (83,9 %)	80 (89,9 %)	
männlich	24 (13,2 %)	15 (16,1 %)	9 (10,1 %)	
Alkohol und Rauchen				
Derzeitiger Alkoholkonsum, n (%)				0,1
kein derzeitiger Alkoholkonsum	71 (40,6 %)	34 (39,1 %)	37 (41,6 %)	
1x Woche	61 (34,9 %)	34 (36,6 %)	27 (30,3 %)	
2x Woche	35 (20,0 %)	15 (16,1 %)	20 (22,5 %)	
3x Woche	8 (4,6 %)	4 (4,6 %)	4 (4,6 %)	
Keine Angabe	7	6	1	
Rauchen, n (%)				0,1
Nichtraucher	106 (59,2 %)	52 (57,1 %)	54 (61,4 %)	

Gefördert durch: GBA Innovationsausschuss

Auswertepopulation C	Gesamt N = 182	Kontrollgruppe N = 93	Interventionsgruppe N = 89	SMD [a]
Ehemalige:r Raucher:in	56 (31,3 %)	30 (33,0 %)	26 (29,5 %)	
Derzeitige:r Raucher:in	17 (9,5 %)	9 (9,9 %)	8 (9,1 %)	
Keine Angabe	3	2	1	
Komorbiditäten (Patient:innenangabe), n (%)				
Allergien	103 (56,6 %)	45 (48,4 %)	58 (65,2 %)	0,3
Kardiovaskuläre Erkrankungen [1]	28 (15,4 %)	16 (17,2 %)	12 (13,5 %)	0,1
Gastrointestinale Erkrankungen [2]	16 (8,8 %)	7 (7,5 %)	9 (10,1 %)	0,1
Erkrankungen des Urogenitaltraktes [3]	13 (7,1 %)	9 (9,7 %)	4 (4,5 %)	0,2
Endokrine Erkrankungen [4]	39 (21,4 %)	22 (23,7 %)	17 (19,1 %)	0,1
Muskuloskelettale Erkrankungen [5]	10 (5,5 %)	6(6,5 %)	4(4,5 %)	0,1
Neurologische Störungen [6]	15 (8,3 %)	9 (9,7 %)	6 (6,8 %)	0,1
Pulmonale Erkrankungen [7]	43 (23,6 %)	21 (22,6 %)	22 (24,7 %)	0,1
Hauterkrankungen [8]	11 (6,1 %)	6 (6,5 %)	5 (5,6 %)	0,0
Psychische Störungen [9]	19 (10,5 %)	7 (7,5 %)	12 (13,6 %)	0,2
Andere [10]	21 (11,5 %)	13 (14,0 %)	8 (9,0 %)	0,2
Familie und Kinder				
Familienstand, n (%)				0,1
Verheiratet [11]	80 (46,0 %)	42 (47,2 %)	38 (44,7 %)	
Getrennt	15 (8,6 %)	5 (5,6 %)	10 (11,8 %)	
Ledig	79 (45,4 %)	42 (47,2 %)	37 (43,5 %)	
Keine Angabe	8	4	4	
Kinder, n (%)				0,0
Nein	114 (66,7 %)	57 (65,5 %)	57 (67,9 %)	
Ja	57 (32,1 %)	30 (34,5 %)	27 (32,1 %)	
Keine Angabe	11	6	5	
Ausbildung und Beschäftigung				
Ausbildung, n (%)				0,1
Kein Abschluss, noch in Ausbildung	12 (7,1 %)	7 (8,0 %)	5 (6,3 %)	
Beruf-betrieblich	50 (29,8 %)	25 (28,4 %)	25 (31,2 %)	

Gefördert durch: GBA Innovationsausschuss

Auswertepopulation C	Gesamt N = 182	Kontrollgruppe N = 93	Interventionsgruppe N = 89	SMD [a]
Beruflich-schulisch, Fachschule	21 (12,5 %)	9 (10,2 %)	12 (15,0 %)	
Fachhochschule, Universität	82 (48,9 %)	46 (52,2 %)	36 (45,0 %)	
Anderer Abschluss	3 (1,8 %)	1 (1,1 %)	2 (2,5 %)	
Keine Angabe	14	5	9	
Internationale Standardklassifikation für das Bildungswesen (ISCED), n (%)				0,1
Niedrig [12]	6 (3,4 %)	5 (5,5 %)	1 (1,2 %)	
Mittel [13]	76 (43,7 %)	35 (38,5 %)	41 (49,4 %)	
Hoch [14]	92 (52,92 %)	51 (56,0 %)	41 (49,4 %)	
Keine Angabe	8	2	6	
Beschäftigungsstatus, n (%)				0,0
Beschäftigt	161 (91,5 %)	82 (90,1 %)	79 (92,9 %)	
Nicht beschäftigt	15 (8,50 %)	9 (9,9 %)	6 (7,1 %)	
Keine Angabe	6	2	4	
Krankheitsmerkmale				
Schweregrad [Bell], MW (SD)	36,6 (7,6)	36,8 (8,0)	36,5 (7,3)	0,0
Dauer [Monate], MW (SD)	19,0 (12,2)	21,1 (14,1)	16,9 (9,3)	0,4
Auslöser (Trigger), n (%)				0,1
Covid-19	141 (77,5 %)	71 (78,0 %)	70 (78,7 %)	
Epstein-Barr-Virus	7 (3,8 %)	5 (5,5 %)	2 (2,2 %)	
Andere Atemweg-sinfektionen [15]	18 (9,9 %)	8 (8,8 %)	10 (11,2 %)	
Andere Infektionen [16]	5 (2,7 %)	3 (3,3 %)	2 (2,2 %)	
Keine Infektion [17]	9 (4,9 %)	4 (4,4 %)	5 (5,6 %)	
Unbekannt, keine Angabe	2	2	0	

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Absolute Häufigkeit.
[a] Standardisierte Mittelwertsdifferenz (berechnet auf den relativen Häufigkeiten für kategoriale Merkmale).
[1] Arterieller Hypertonus, Herzrhythmusstörungen, usw.
[2] Reizdarmsyndrom, Gastritis, usw.
[3] Endometriose, Blasenstörungen, usw.
[4] Schilddrüsenüber- oder -unterfunktion, Diabetes mellitus, usw.
[5] Arthrose, Fibromyalgie Syndrom, usw.
[6] Migräne, Polyneuropathie, usw.
[7] Asthma, chronische Sinusitis, usw.
[8] Neurodermitis, Ekzem, usw.
[9] Depression, Angststörungen, usw.
[10] Tinnitus, Lip-/Lymphödem, u. a.
[11] Inklusive von Ehepartner:in getrennt lebend.
[12] Definiert als: Kein Schulabschluss, noch keinen Schulabschluss, Haupt-, Volks-, oder Realschulabschluss oder Abschluss an einer polytechnischen Oberschule; keine beruflich-betrieblich oder beruflich-schulische Ausbildung.

Gefördert durch: GBA Innovationsausschuss

[13] Definiert als: Beruflich-betriebliche Ausbildung, beruflich-schulische Ausbildung, Fachhochschulreife, Abitur, allgemeine oder fachgebundene Hochschulreife, oder erweiterte Oberschule; keinen Abschluss an einer Fachschule, Fachakademie, Fachhochschule, Ingenieurschule, Universität oder Hochschule.
[14] Definiert als: Abschluss an einer Fachschule, Fachakademie, Fachhochschule, Ingenieurschule, Universität oder Hochschule.
[15] Pneumonie, Bronchitis, Sinusitis, Otitis, Influenza, usw.
[16] Meningitis, Pyelonephritis, Borreliose, Herpes, Zytomegalovirus.
[17] Impfung, körperliche Belastung.

3.1.3 Baseline-Charakteristika der ausgeschlossenen Patient:innen

Ein Vergleich der Baseline-Charakteristika zeigt, dass die aus der Auswertepopulation B und C ausgeschlossenen Patient:innen nur geringfügig von den verbliebenen Patient:innen abwichen. So gaben die ausgeschlossenen Patient:innen beispielsweise einen niedrigeren Alkoholkonsum an. Zudem waren sie seltener verheiratet (dafür häufiger ledig) und wiesen einen geringeren Anteil neurologischer, kardiovaskulärer sowie endokriner Erkrankungen auf. Diese Unterschiede sind insgesamt jedoch als vernachlässigbar einstufbar. Auffällig war dagegen die Geschlechterverteilung. Während in Auswertepopulation C 24 Männer verblieben, wurden 13 ausgeschlossen (ca. 35 %). Bei den Frauen hingegen lag die Ausschlussquote mit 46 Frauen bei etwa 22 %. Aus Auswertepopulation B wurden dagegen 9 von insgesamt 37 Männern (24,3 %) und 35 von insgesamt 204 Frauen (17,2 %) ausgeschlossen.

3.2 Primärer Endpunkt: Körperliche Funktionsfähigkeit (SF-36)

Die Skala der körperlichen Funktionsfähigkeit (SF-36) reicht von 0 bis 100 Punkten, wobei höhere Werte einer besseren körperlichen Funktionsfähigkeit entsprechen. Tabelle 3 zeigt die Ergebnisse der Analyse in Auswertepopulation B und C. Sowohl die Kontrollgruppe (40,4 Punkte) als auch die Interventionsgruppe (ebenfalls 40,4 Punkte) wiesen zu Baseline in Auswertepopulation B identische Mittelwerte auf; die Standardabweichung betrug in beiden Fällen etwa 19 Punkte. Zu Monat 12 hatte die Interventionsgruppe mit einem Mittelwert von 36,0 Punkten (Standardabweichung: 20,6 Punkte) niedrigere (d.h. schlechtere) Werte als die Kontrollgruppe mit 41,8 Punkten (Standardabweichung: 25,9 Punkte). Entsprechend betrug die unadjustierte Differenz −5,8 Punkte (95 %-KI: [-12,6; 1,0]). Nach Adjustierung betrug die Differenz nur noch −3,8 Punkte (95-% KI: [−9,9; 2,2]) und war statistisch nicht signifikant (P-Wert = 0,211). Im Durchschnitt hatten Patient:innen in der Interventionsgruppe nach 12 Monaten also eine um 3,9 Punkte schlechtere körperliche Funktionsfähigkeit als Patient:innen der Kontrollgruppe. Die Effektstärken (partielles η^2) war mit 0,009 aber nur gering. Die Ergebnisse in Auswertepopulation C waren nahezu deckungsgleich.

Gefördert durch: GBA Innovationsausschuss

Tabelle 3: Ergebnisse der Analyse der körperlichen Funktionsfähigkeit (SF-36).

Körperliche Funktionsfähigkeit (SF-36) [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert [a, b]
Auswertepopulation B						
Baseline	99	40,4 (19,2)	94	40,4 (18,5)		
Monat 12	94	41,8 (25,9)	92	36,0 (20,6)	-5,8 [-12,6; 1,0] η^2 = 0,015 P = 0,094	-3,8 [-9,9; 2,2] η^2 = 0,008 P = 0,211
Auswertepopulation C						
Baseline	93	41,6 (18,7)	89	40,0 (18,3)		
Monat 12	93	41,6 (25,9)	89	36,6 (20,3)	-4,8 [-11,6; 2,1] η^2 = 0,010 P = 0,171	-1,9 [-8,2; 4,5] η^2 = 0,002 P = 0,564

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.
[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.
[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

3.2.1 Sensitivitätsanalysen

Um die Robustheit der Ergebnisse zu untersuchen, wurde die Analyse in den Auswertepopulationen A und P sowie mit einer Complete-Case Analyse überprüft. Die Ergebnisse der Analysen sind in Tabelle 4

Tabelle 4 dargestellt. Die Ergebnisse bestätigen das negative Vorzeichen des Punktschätzers aus der primären Analyse und finden teils nominal sogar etwas stärkere Effekte (−4,1 Punkte in A, −3,2 in P und −2,0 Punkte im Rahmen der Complete-Case). Neben dem Vorzeichen bestätigten die Sensitivitätsanalysen auch die erhebliche statistische Unsicherheit der Hauptanalyse: Alle berechneten 95-%-KIs schlossen die Null ein.

Tabelle 4: Ergebnisse der Sensitivitätsanalysen der körperlichen Funktionsfähigkeit (SF-36).

Körperliche Funktionsfähigkeit (SF-36) [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert [a]
Auswertepopulation A [b]						
Baseline	99	40,4 (19,2)	134	38,5 (18,5)		
Monat 12	94	41,8 (25,9)	111	35,2 (22,0)	-6,6 [-13,2; 0,0] η^2 = 0,019 P = 0,050	-4,1 [-9,5; 1,2] η^2 = 0,010 P = 0,128
Auswertepopulation P [b]						
Baseline	98	40,6 (19,2)	92	40,8 (2,1)		

Gefördert durch: GBA Innovationsausschuss

Körperliche Funktionsfähigkeit (SF-36) [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert [a]
Monat 12	94	41,8 (25,9)	92	36,0 (20,6)	-5,8 [-12,6; 1,0] η^2 = 0,015 P = 0,094	-3,2 [-9,3; 2,9] η^2 = 0,006 P = 0,302
Complete-Case						
Baseline	91	41,9 (18,8)	83	40,7 (18,7)		
Monat 12	91	42,1 (26,0)	83	37,4 (20,5)	-4,7 [-11,7; 2,4] η^2 = 0,001 P = 0,192	-2,0 [-8,5; 4,6] η^2 = 0,002 P = 0,552

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.
[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.
[b] Die berichteten Differenzen, partielle η^2 sowie P-Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

3.2.2 Subgruppenanalysen

Um zu untersuchen, ob der Effekt der Intervention möglicherweise zwischen verschiedenen Gruppen von ME/CFS-Patient:innen variiert, wurde eine Subgruppenanalyse in Auswertepopulation B und C durchgeführt. Diese Analyse zielte darauf ab, potenzielle Heterogenität der Behandlungseffekte zu identifizieren. Dabei wurden die folgenden zwei Subgruppenkriterien betrachtet:

- Geschlecht: Vergleich des Interventionseffekts zwischen Männern und Frauen.
- Schwere der Erkrankung: Vergleich des Interventionseffekts zwischen schwerer Erkrankung (Bell = 30) und weniger schweren Erkrankung (Bell ≥ 40).

Die Subgruppenanalysen zeigten je Subgruppe Effekte mit unterschiedlichem Vorzeichen (Tabelle 5). Bei Männern in Auswertepopulation B ergab sich eine positive adjustierte Differenz, was auf günstigere Werte in der Interventionsgruppe hindeutet, während die adjustierte Differenz bei Frauen negativ ausfiel (4,2 Punkte mit 95-%-KI: [−11,7; 20,0] vs. −5,2 Punkte mit 95-%-KI: [−11,7; 1,3]). Hinsichtlich der Erkrankungsschwere zeigte sich bei schwerer erkrankten Patient:innen nur ein minimaler Unterschied zwischen den Gruppen (adjustierte Differenz: -0,5 Punkte mit 95-%-KI: [−9,0; 8,1]). Demgegenüber war bei weniger schwer erkrankten Patient:innen eine negative unadjustierte Differenz von −11,3 Punkten (mit 95-%-KI: [−20,4; −2,3]) zu beobachten, was auf eine Verschlechterung in der Interventionsgruppe hinweist. Nach Adjustierung verringerte sich diese Differenz jedoch deutlich auf −6,9 Punkte (95-%-KI: [−15,3; 1,4]). Insgesamt sind die geschätzten Effekte mit erheblicher statistischer Unsicherheit behaftet; in allen betrachteten Subgruppen schloss das 95-%-KI der adjustierten Differenzen die Null ein. Bezogen auf Auswertepopulation C wichen diese Werte nominal nur leicht ab.

Gefördert durch: GBA Innovationsausschuss

Tabelle 5: Ergebnisse der Subgruppenanalyse nach Geschlecht und Schwere der Erkrankung der körperlichen Funktionsfähigkeit (SF-36).

Körperliche Funktionsfähigkeit (SF-36) [Punkte]		Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
		n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert [a, b]
Auswertepopulation B							
Geschlecht							η^2 = 0,006
Männer	Baseline	17	43,5 (19,5)	11	45,0 (17,5)		
	Monat 12	15	45,3 (27,3)	9	51,1 (24,2)	5,8 [-13,6; 25,1] P = 0,557	4,2 [-11,7; 20,0] P = 0,604
Frauen	Baseline	82	39,7 (19,2)	83	39,8 (18,7)		
	Monat 12	79	41,1 (25,7)	83	34,4 (19,4)	-6,7 [-14,0; 0,5] P = 0,067	-5,2 [-11,7; 1,3] P = 0,115
Schwere der Erkrankung							η^2 = 0,006
Bell = 30	Baseline	51	29,5 (15,5)	44	32,7 (16,7)		
	Monat 12	46	32,5 (24,2)	43	32,2 (19,6)	-0,3 [-9,7; 9,1] P = 0.951	-0,5 [-9,0; 8,1] P = 0,915
Bell ≥ 40	Baseline	48	51,9 (15,9)	50	47,2 (17,6)		
	Monat 12	48	50,7 (24,4)	49	39,4 (21,2)	-11,3 [-20,4; -2,3] P = 0,014	-6,9 [-15,3; 1,4] P = 0,101
Auswertepopulation C							
Geschlecht							η^2 = 0,012
Männer	Baseline	15	46,7 (27,3)	9	42,8 (13,9)		
	Monat 12	15	45,3 (27,3)	9	51,1 (24,2)	5,8 [-25,1; 13,5] P = 0,556	8,2 [-10,0; 26,3] P = 0,376
Frauen	Baseline	78	40,6 (18,7)	80	39,7 (18,7)		
	Monat 12	78	40,9 (25,8)	80	35,2 (19,4)	-5,7 [-12;9, 1,6] P = 0,128	-3,5 [-10,5; 3,5] P = 0,329
Schwere der Erkrankung							η^2 = 0,008
Bell = 30	Baseline	46	30,4 (14,9)	42	31,9 (15,2)		
	Monat 12	46	32,5 (24,2)	42	33,0 (19,2)	0,5 [-9,0; 9,9] P = 0.921	0,4 [-9,0; 9,8] P = 0,932
Bell ≥ 40	Baseline	47	52,6 (15,3)	47	47,2 (17,9)		
	Monat 12	47	50,5 (24,6)	47	40,3 (20,9)	-10,2 [-19,3; -1,1] P = 0,029	-4,1 [-13,2; 5,0] P = 0,372

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.
[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.
[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

3.3 Sekundäre Endpunkte

Alle sekundären Endpunkte wurden in Auswertepopulation B und C evaluiert. Sensitivitätsanalysen auf den Auswertepopulationen A und P sowie einer Complete-Case Analyse sind im Anhang tabellarisch dargestellt.

Gefördert durch: GBA Innovationsausschuss

3.3.1 Bell Disability Scale

Die Bell Disability Scale reicht von 0 bis 100 Punkten, wobei höhere Werte einen geringeren Grad der Behinderung anzeigen. Tabelle 6 zeigt die Ergebnisse der Analyse in Auswertepopulation B und C. Die Interventionsgruppe in Auswertepopulation B wies zu Monat 12 mit einem Mittelwert von 33,0 Punkten niedrigere (d.h. schlechtere) Werte auf als die Kontrollgruppe (37,9 Punkte), was einer Mittelwertdifferenz von −4,9 Punkten (95 %-KI: [−8,7; −1,1]) entsprach. Durch die Adjustierung verringerte sich diese Differenz leicht auf −4,4 Punkte (95 %-KI: [−8,1; -0,6]). Die Effektstärke (partielles η^2) waren mit 0,022 aber nur gering. Die Ergebnisse in der Auswertepopulation C waren nahezu deckungsgleich.

Tabelle 6: Ergebnisse der Analyse der Bell Disability Scale.

Bell Disability Scale [Punkte]	Kontrollgruppe n	Kontrollgruppe MW (SD)	Interventionsgruppe n	Interventionsgruppe MW (SD)	Differenz [95-% Konfidenzintervall] Unadjustiert	Differenz [95-% Konfidenzintervall] Adjustiert [a, b]
Auswertepopulation B						
Baseline	100	36,5 (7,8)	97	36,8 (7,6)		
Monat 12	95	37,9 (16,0)	94	33,0 (9,5)	-4,9 [-8,7; -1,1] η^2 = 0,034 P = 0,011	-4,4 [-8,1; -0,6] η^2 = 0,029 P = 0,022
Auswertepopulation C						
Baseline	93	36,8 (8,0)	89	36,5 (7,2)		
Monat 12	93	38,0 (16,1)	89	33,1 (9,5)	-4,8 [-8,7; -0,9] η^2 = 0,032 P = 0,015	-3,7 [-7,5; 0,1] η^2 = 0,011 P = 0,058

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.
[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.
[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

3.3.2 Chalder Fatigue Scale

Die Chalder Fatigue Scale ist ein zweidimensionales Instrument zur Erfassung von mentaler Fatigue (Wertebereich: 0-12 Punkte) und physischer Fatigue (Wertebereich: 0-21 Punkte), wobei höhere Punktwerte ein höheres Ausmaß an Fatigue indizieren. Die Ergebnisse der Analyse sind in Tabelle 7 dargestellt. Im Folgenden wird sich auf den Gesamtscore (Summe der Subskalen; Wertebereich: 0-33 Punkte) in Auswertepopulation B beschränkt. Zu Monat 12 wiesen beide Gruppen nahezu identische Mittelwerte auf (25,8 bzw. 26,0 Punkte). Die adjustierte und die unadjustierte Differenz von 0,2 Punkten (95 %-KI: [−1,2; 1,6]) bzw. −0,2 Punkten (95 %-KI: [−1,5; 1,1]) sind äußerst gering. Negative Werte der Differenz entsprechen dabei einem geringeren Level an Fatigue.

Gefördert durch: GBA Innovationsausschuss

Tabelle 7: Ergebnisse der Analyse der Chalder Fatigue Scale (mental, körperlich und Gesamtscore).

Chalder Fatigue Scale [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert [a, b]
Auswertepopulation B						
Gesamtscore						
Baseline	95	27,8 (3,5)	93	27,8 (3,0)		
Monat 12	94	25,8 (4,5)	93	26,0 (4,9)	0,2 [-1,2; 1,6] η^2 < 0,001 P = 0,769	-0,2 [-1,5; 1,1] η^2 < 0,001 P = 0,796
Mental						
Baseline	97	9,3 (1,9)	94	9,5 (1,7)		
Monat 12	95	8,9 (2,0)	93	8,9 (2,2)	0,0 [-0,6; 0,6] η^2 < 0,001 P = 0,951	-0,2 [-0,8; 0,4] η^2 = 0,003 P = 0,472
Körperlich						
Baseline	92	18,6 (2,3)	88	18,3 (2,1)		
Monat 12	93	17,0 (3,1)	89	17,0 (3,1)	-0,0 [-0,9; 1,0] η^2 < 0,001 P = 0,923	-0,2 [-1,1; 0,8] η^2 < 0,001 P = 0,751
Auswertepopulation C						
Gesamtscore						
Baseline	89	27,7 (3,5)	86	27,7 (2,9)		
Monat 12	93	25,9 (4,5)	88	25,9 (4,9)	0,0 [-1,4; 1,4] η^2 < 0,001 P = 0,995	-0,3 [-1,7; 1,1] η^2 < 0,001 P = 0,686
Mental						
Baseline	90	9,2 (1,9)	87	9,5 (1,7)		
Monat 12	93	8,9 (2,0)	88	8,9 (2,1)	-0,0 [-0,6; 0,6] η^2 < 0,001 P = 0,924	-0,2 [-0,8; 0,4] η^2 = 0,003 P = 0,481
Körperlich						
Baseline	92	18,6 (2,3)	88	18,3 (2,1)		
Monat 12	93	17,0 (3,1)	89	17,0 (3,1)	-0,0 [-0,9; 1,0] η^2 < 0,001 P = 0,923	-0,2 [-1,1; 0,8] η^2 < 0,001 P = 0,751

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.
[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.
[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

3.3.3 CCC-Kriterien

Die Analyse der Symptomschwere (nach CCC) zeigte kaum relevante Unterschiede in Auswertepopulation B (Tabelle 8). Die Wahrscheinlichkeit für das Vorliegen von Muskelschmerz war in der Interventionsgruppe erhöht (unadjustiertes Odds Ratio: 4,84 mit 95-% KI: [1,15; 15,10]; adjustiertes Odds Ratio: 3,76 mit 95-% KI: [1,01; 14,05]). Auffällig weit fallen die 95-%-KIs aus, was teilweise auf das starke Ungleichgewicht zwischen den Klassen zurückzuführen ist (über 80 % positiv). Bei Kopfschmerzen waren die Klassen deutlich besser balanciert (50 bis 60 % positiv) und die Weite der 95-% KIs geringer. Aber auch die Odds Ratios sind näher an eins (unadjustiertes Odds Ratio: 0,97 mit 95-% KI: [0,54; 1,74]; adjustiertes Odds

Gefördert durch: GBA Innovationsausschuss

Ratio: 1,13 mit 95-% KI: [0,55; 2,31]). Die Ergebnisse in Auswertepopulation C waren im Wesentlichen deckungsgleich.

Tabelle 8: *Ergebnisse der Analyse der Symptome Muskelschmerz, Kopfschmerzen, nach den Kanadischen Konsensus Kriterien*.

CCC-Kriterien (Auswahl)	Kontrollgruppe		Interventionsgruppe		Odds Ratio [95-% Konfidenzintervall]	
	n	Positiv (%)	n	Positiv (%)	Unadjustiert	Adjustiert [a, b]
Auswertepopulation B						
Kopfschmerzen						
Baseline	100	65 (65,0)	96	59 (61,5)		
Monat 12	96	59 (61,5)	94	57 (60,6)	0,97 [0,54; 1,74] P = 0,908	1,13 [0,55; 2,31] P = 0,735
Muskelschmerz						
Baseline	100	86 (86,0)	96	87 (90,6)		
Monat 12	96	79 (82,3)	94	90 (95,7)	4,84 [1,15; 15,10] P = 0,007	3,76 [1,01; 14,05] P = 0,049
Auswertepopulation C						
Kopfschmerzen						
Baseline	93	61 (65,6)	89	54 (60,7)		
Monat 12	93	57 (61,3)	89	52 (58,4)	0,89 [0,49; 1,61] P = 0,694	1,08 [0,52; 2,25] P = 0,839
Muskelschmerz						
Baseline	93	80 (86,0)	89	82 (92,1)		
Monat 12	93	78 (83,9)	89	85 (95,5)	4,09 [1,29; 12,94] P = 0,017	2,65 [0,70; 10,00] P = 0,148

n: Fallzahl.
[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.
[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

3.3.4 COMPASS-31 Score

Die COMPASS-31 Score besteht aus insgesamt sechs Subskalen (Blase, Gastrointestinaltrakt, Orthostase, Pupillomotorik, Sekretomotorik und Vasomotorik) mit unterschiedlichen Wertebereichen, wobei höhere Werte einer höheren Belastung durch autonome Symptome anzeigen. Die Ergebnisse der Analyse sind in Tabelle 9 dargestellt. Die folgende Interpretation der Ergebnisse beschränkt sich auf den Gesamtscore (Wertebereicht: 0-100 Punkte) in Auswertepopulation B. Zu Monat 12 wiesen beide Gruppen im Durchschnitt fast identische Scores auf (Kontrollgruppe 39,4 Punkte vs. Interventionsgruppe 39,7 Punkte), was einer unadjustierten Mittelwertdifferenz von 0,4 Punkten (mit 95-% KI: [−4,0; 4,7]) entspricht. Die adjustierte Differenz war nominal etwas größer und betrug −1,4 Punkte (95 %-KI: [−5,1; 2,4]). Negative Werte der Differenz deuten dabei auf eine geringere Gesamtbelastung durch autonome Symptome in der Interventionsgruppe im Vergleich zur Kontrollgruppe hin. Die Differenzen sind jedoch äußerst gering.

Gefördert durch: GBA Innovationsausschuss

Tabelle 9: Ergebnisse der Analyse des Composite Autonomic Symptom Score 31 (Blase, Gastrointestinaltrakt, Orthostase, Pupillomotorik, Sekretomotorik und Vasomotorik und Gesamtscore).

COMPASS-31 [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	N	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert [a, b]
Auswertepopulation B						
Gesamtscore						
Baseline	93	40,3 (14,0)	83	38,6 (15,2)		
Monat 12	87	39,4 (14,4)	87	39,7 (14,8)	0,4 [-4,0; 4,7] η^2 < 0,001 P = 0,870	-1,4 [-5,1; 2,4] η^2 = 0,003 P = 0,470
Blase						
Baseline	99	1,2 (1,7)	95	1,2 (1,5)		
Monat 12	95	1,1 (1,4)	94	1,2 (1,4)	0,1 [-0,3; 0,5] η^2 < 0,001 P = 0,729	-0,1 [-0,4; 0,2] η^2 = 0,002 P = 0,537
Gastrointestinaltrakt						
Baseline	98	7,7 (3,8)	95	7,6 (3,7)		
Monat 12	92	8,5 (3,9)	90	8,1 (4,2)	-0,4 [-1,6; 0,8] η^2 = 0,002 P = 0,542	-0,1 [-1,1; 0,8] η^2 < 0,001 P = 0,759
Orthostase						
Baseline	99	22,5 (9,9)	95	21,7 (11,4)		
Monat 12	93	21,0 (9,7)	94	20,3 (10,8)	-0,7 [-3,7; 2,2] η^2 = 0,001 P = 0,625	-1,0 [-3,8; 1,8] η^2 = 0,003 P = 0,479
Pupillomotorik						
Baseline	98	2,4 (1,2)	90	2,4 (1,1)		
Monat 12	93	2,7 (1,1)	94	2,7 (1,0)	0,0 [-0,3; 0,3] η^2 < 0,001 P = 0,913	-0,1 [-0,3; 0,2] η^2 = 0,003 P = 0,485
Sekretomotorik						
Baseline	97	5,2 (3,7)	92	5,3 (3,6)		
Monat 12	93	5,4 (3,5)	93	5,5 (3,8)	0,2 [-0,9; 1,2] η^2 < 0,001 P = 0,765	-0,2 [-1,1; 0,7] η^2 = 0,001 P = 0,608
Vasomotorik						
Baseline	100	1,0 (1,4)	94	0,9 (1,4)		
Monat 12	94	1,2 (1,5)	91	1,4 (1,6)	0,2 [-0,3; 0,7] η^2 = 0,004 P = 0,372	0,2 [-0,3; 0,6] η^2 = 0,003 P = 0,470
Auswertepopulation C						
Gesamtscore						
Baseline	86	39,3 (13,5)	77	38,1 (15,2)		

Gefördert durch: GBA Innovationsausschuss

COMPASS-31 [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	N	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert [a, b]
Monat 12	95	39,7 (13,9)	82	38,9 (14,5)	-0,7 [-5,1; 3,6] η^2 < 0,001 P = 0,737	-0,9 [-5,0; 3,2] η^2 = 0,001 P = 0,656
Blase						
Baseline	92	1,0 (1,3)	88	1,2 (1,5)		
Monat 12	93	1,1 (1,4)	89	1,2 (1,4)	0,1 [-0,4; 0,5] η^2 < 0,001 P = 0,809	-0,2 [-0,5; 0,2] η^2 = 0,005 P = 0,348
Gastrointestinaltrakt						
Baseline	91	7,5 (3,8)	88	7,5 (3,7)		
Monat 12	90	8,5 (3,8)	85	7,9 (4,2)	-0,6 [-1,9; 0,6] η^2 = 0,006 P = 0,293	-0,4 [-1,4; 0,7] η^2 = 0,003 P = 0,485
Orthostase						
Baseline	92	22,0 (9,9)	89	21,4 (11,6)		
Monat 12	91	21,1 (9,5)	89	19,8 (10,7)	-1,4 [-4,3; 1,6] η^2 = 0,005 P = 0,365	-2,0 [-4,9; 1,0] η^2 = 0,010 P = 0,187
Pupillomotorik						
Baseline	91	2,3 (1,2)	83	2,4 (1,2)		
Monat 12	91	2,7 (1,1)	89	2,7 (1,0)	-0,0 [-0,4; 0,3] η^2 < 0,001 P = 0,802	-0,2 [-0,4; 0,1] η^2 = 0,009 P = 0,233
Sekretomotorik						
Baseline	90	5,1 (3,7)	86	5,4 (3,6)		
Monat 12	91	5,5 (3,5)	88	5,5 (3,8)	0,1 [-1,0; 1,1] η^2 < 0,001 P = 0,906	-0,1 [-1,1; 0,8] η^2 < 0,001 P = 0,793
Vasomotorik						
Baseline	93	0,9 (1,4)	88	0,9 (1,4)		
Monat 12	92	1,2 (1,5)	86	1,4 (1,6)	0,2 [-0,3; 0,7] η^2 = 0,007 P = 0,355	0,2 [-0,3; 0,6] η^2 = 0,003 P = 0,517

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.
[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.
[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

3.3.5 SF-36 weitere Skalen

Tabelle 10 stellt die Ergebnisse der Analyse des Gesamt SF-36 für die Auswertepopulationen B und C dar. Die einzelnen SF-36 Skalen reichen von 0 bis 100 Punkten, wobei höhere Werte eine bessere gesundheitsbezogene Lebensqualität im jeweiligen Bereich anzeigen. In den Bereichen körperliche Schmerzen (adjustierte Differenz: –3,1 Punkte mit 95-% KI: [–9,3; 3,0]),

Gefördert durch: GBA Innovationsausschuss

allgemeine Gesundheitswahrnehmung (adjustierte Differenz: -1,4 Punkte mit 95-% KI: [-5,1; 2,3]), körperliche Rollenfunktion (adjustierte Differenz: −4,4 Punkte mit 95-% KI: [−9,1; 0,2]) und Vitalität (adjustierte Differenz: −1,6 Punkte mit 95-% KI: [−6,4; 3,2]) waren die beobachteten Gruppenunterschiede in Auswertepopulation B negativ (d.h. nachteilig für die Interventionsgruppe im Vergleich zur Kontrollgruppe) aber nominal eher klein. Bei der Änderung des Gesundheitszustands lag die adjustierte Differenz bei -9,1 Punkten (95-% KI: [-17,3; -0,9]) und bei der sozialen Funktion bei -6,9 Punkten (95-% KI: [-13,7; -0,1]). Lediglich die emotionale Rollenfunktion und das emotionale Wohlbefinden zeigten mit einer positiven adjustierten Differenz von 12,1 bzw. 2,8 Punkten (95-% KI: [1,6; 22,7] bzw. [-1,5; 7,0]) einen gewissen Vorteil für die Interventionsgruppe. Die Effektstärken (partielles η^2) waren allesamt aber lediglich gering. Die entsprechenden Ergebnisse in der Auswertepopulation waren bis auf kleine Abweichungen deckungsgleich.

Tabelle 10: Ergebnisse der Analyse des Short Form-36 (körperliche Schmerzen, emotionales Wohlbefinden, allgemeine Gesundheitswahrnehmung, Änderung des Gesundheitszustandes, emotionale Rollenfunktion, körperliche Rollenfunktion, soziale Funktion, Vitalität).

SF-36 [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert [a, b]
Auswertepopulation B						
Körperliche Schmerzen						
Baseline	98	30,7 (20,6)	94	33,2 (23,6)		
Monat 12	93	38,1 (26,9)	94	33,3 (19,6)	-4,8 [-11,6; 2,0] η^2 = 0,010 P = 0,167	-3,1 [-9,3; 3,0] η^2 = 0,005 P = 0,319
Emotionales Wohlbefinden						
Baseline	96	59,2 (18,3)	96	58,2 (17,8)		
Monat 12	93	63,1 (19,1)	92	63,6 (18,3)	0,4 [-5,0; 5,9] η^2 < 0,001 P = 0,877	2,8 [-1,5; 7,0] η^2 = 0,009 P = 0,200
Allgemeine Gesundheitswahrnehmung						
Baseline	98	27,2 (13,0)	95	27,4 (13,3)		
Monat 12	93	28,0 (14,1)	93	26,2 (11,7)	-1,8 [-5,5; 2,0] η^2 = 0,005 P = 0,351	-1,4 [-5,1; 2,3] η^2 = 0,003 P = 0,458
Änderung des Gesundheitszustands						
Baseline	100	25,5 (29,9)	96	20,6 (28,3)		
Monat 12	93	53,2 (28,7)	93	42,5 (25,8)	-10,7 [-18,6; -2,8] η^2 = 0,038 P = 0,008	-9,1 [-17,3; -0,9] η^2 = 0,026 P = 0,031
Emotionale Rollenfunktion						
Baseline	99	65,6 (43,6)	95	62,5 (43,8)		
Monat 12	92	69,2 (43,7)	94	75,2 (40,6)	6,0 [-6,2; 18,2] η^2 = 0,005 P = 0,335	12,1 [1,6; 22,7] η^2 = 0,028 P = 0,024
Körperliche Rollenfunktion						
Baseline	99	1,5 (7,8)	96	1,3 (6,7)		
Monat 12	93	6,8 (20,2)	93	2,1 (7,9)	-4,6 [-9,0; -0,1] η^2 = 0,022 P = 0,044	-4,4 [-9,1; 0,2] η^2 = 0,020 P = 0,059

Gefördert durch: GBA Innovationsausschuss

SF-36 [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert [a, b]
Soziale Funktion						
Baseline	97	25,5 (19,8)	96	23,6 (18,7)		
Monat 12	89	32,6 (26,8)	93	22,6 (21,7)	-10,0 [-17,1; -2,9] η^2 = 0,041 P = 0,006	-6,9 [-13,7; -0,1] η^2 = 0,022 P = 0,046
Vitalität						
Baseline	98	16,4 (13,7)	96	16,0 (12,8)		
Monat 12	94	21,4 (18,4)	91	18,3 (16,5)	-3,1 [-8,2; 2,0] η^2 = 0,008 P = 0,231	-1,6 [-6,4; 3,2] η^2 = 0,002 P = 0,515
Auswertepopulation C						
Körperliche Schmerzen						
Baseline	91	31,3 (20,9)	87	32,6 (21,9)		
Monat 12	92	37,8 (26,9)	89	33,6 (19,7)	-4,2 [-11,1; 2,7] η^2 = 0,008 P = 0,233	-2,2 [-8,7; 4,3] η^2 = 0,003 P = 0,502
Emotionales Wohlbefinden						
Baseline	89	60,9 (17,0)	89	57,6 (18,1)		
Monat 12	92	63,2 (19,2)	88	64,1 (17,9)	0,9 [-4,6; 6,3] η^2 < 0,001 P = 0,753	3,0 [-1,1; 7,2] η^2 = 0,012 P = 0,152
Allgemeine Gesundheitswahrnehmung						
Baseline	92	27,5 (12,8)	88	27,2 (13,4)		
Monat 12	93	28,0 (14,1)	88	26,2 (11,3)	-1,8 [-5,5; 2,0] η^2 = 0,005 P = 0,356	-1,5 [-5,3; 2,4] η^2 = 0,003 P = 0,451
Änderung des Gesundheitszustands						
Baseline	93	26,6 (30,6)	89	19,1 (28,0)		
Monat 12	93	53,0 (28,8)	89	43,3 (25,8)	-9,7 [-17,7; 1,7] η^2 = 0,031 P = 0,018	-7,3 [-15,6; 1,0] η^2 = 0,017 P = 0,086
Emotionale Rollenfunktion						
Baseline	92	68,8 (42,2)	88	61,1 (44,1)		
Monat 12	91	68,9 (43,8)	89	76,4 (40,0)	7,5 [-4,8; 19,9] η^2 = 0,008 P = 0,230	14,6 [3,8; 25,4] η^2 = 0,040 P = 0,009
Körperliche Rollenfunktion						
Baseline	92	1,4 (7,7)	89	1,4 (6,9)		
Monat 12	92	6,8 (20,3)	88	2,3 (8,2)	-4,2 [-9,1; 0,1] η^2 = 0,021 P = 0,054	-4,3 [-9,2; 0,6] η^2 = 0,018 P = 0,082
Soziale Funktion						
Baseline	90	26,5 (19,7)	89	23,6 (18,7)		
Monat 12	89	32,6 (26,8)	88	21,9 (20,6)	-10,7 [-17,8; -3,6] η^2 = 0,007 P = 0,003	-7,9 [-15,1; -0,8] η^2 = 0,029

Gefördert durch: GBA Innovationsausschuss

SF-36 [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert [a, b]
						P = 0,029
Vitalität						
Baseline	91	16,4 (13,4)	89	15,7 (12,8)		
Monat 12	93	21,5 (18,5)	87	18,7 (16,6)	-2,8 [-8,0; 2,4] η^2 = 0,006 P = 0,292	-1,1 [-6,2; 4,0] η^2 = 0,001 P = 0,667

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.
[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.
[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

3.3.6 Handgriffkraft

Tabelle 11 zeigt die Ergebnisse der Handgriffkraft-Messungen. Die Analyse umfasst die mittlere Kraft (Fmean), die Maximalkraft (Fmax), das Ermüdungsverhältnis (Fatigue Ratio = Fmax / Fmean) über zwei Zeitpunkte (Sessions) sowie das Erholungsverhältnis (Recovery Ratio = Fmean2 / Fmean1). Höhere Werte bei Fmax und Fmean entsprechend einer höheren Handgriffkraft. Bei den Kraftmessungen schnitt die Interventionsgruppe schlechter ab, was sich in einem negativen Unterschied widerspiegelt. Der Unterschied war jedoch nur gering; Fmean und Fmax der ersten Session in Auswertepopulation B zeigten beispielsweise beide eine adjustierte mittlere Differenz von lediglich −1,6 kg bzw. −2,5 kg (mit 95-% KI: [−3,1; −0,2] bzw. [−4,6; −0,5]). Ermüdungsverhältnis und Erholungsverhältnis unterschieden sich zwischen den Gruppen leicht. Die berechneten Effektstärken (partielles η^2) waren jedoch nur gering. Die Unterschiede in den Handgriffkraft-Messungen im Vergleich zu Auswertepopulation C sind im Wesentlichen auf die Imputation zurückzuführen.

Tabelle 11: Ergebnisse der Analyse der Handgriffkraft Messungen (Fmean, Fmax, Fatigue Ratio, Recovery Ratio).

Handgriff-kraft	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert [a, b]
Auswertepopulation B						
Fmean (Session 1) [kg]						
Baseline	100	16,7 (9,1)	97	16,6 (7,7)		
Monat 12	89	16,1 (8,2)	89	13,6 (7,9)	-2,5 [-4,9; -0,1] η^2 = 0,024 P = 0,039	-1,6 [-3,1; -0,2] η^2 = 0,025 P = 0,029
Fmean (Session 2) [kg]						
Baseline	99	15,1 (9,2)	97	15,1 (7,3)		
Monat 12	89	14,6 (8,2)	88	12,2 (8,0)	-2,4 [-4,8; 0,0] η^2 = 0,021 P = 0,052	-1,6 [-3,0; -0,1] η^2 = 0,025 P = 0,031
Fmax (Session 1) [kg]						
Baseline	100	20,7 (9,6)	97	20,4 (8,9)		

Gefördert durch: GBA Innovationsausschuss

Handgriff-kraft	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert [a, b]
Monat 12	89	20,1 (9,7)	89	16,7 (9,5)	-3,4 [-6,2; -0,6] η^2 = 0,031 P = 0,018	-2,5 [-4,6; -0,5] η^2 = 0,033 P = 0,014
Fmax (Session 2) [kg]						
Baseline	99	18,4 (10,2)	97	18,5 (8,2)		
Monat 12	89	17,8 (9,3)	88	15,1 (9,8)	-2,6 [-5,5; 0,2] η^2 = 0,019 P = 0,067	-1,5 [-3,4; 0,3] η^2 = 0,014 P = 0,106
Fatigue Ratio (Session 1) [c]						
Baseline	100	1,32 (0,28)	97	1,27 (0,19)		
Monat 12	89	1,30 (0,28)	89	1,28 (0,27)	-1,8 % [-6,7; 3,4] η^2 = 0,003 P = 0,483	-1,7 % [-7,1; 3,4] η^2 = 0,002 P = 0,531
Fatigue Ratio (Session 2) [c]						
Baseline	100	1,35 (0,48)	97	1,27 (0,20)		
Monat 12	85	1,27 (0,20)	88	1,33 (0,52)	-2,4 % [-3,0; 8,1] η^2 = 0,004 P = 0,387	-2,0 % [-4,0; 8,4] η^2 = 0,002 P = 0,517
Recovery Ratio						
Baseline	99	0,86 (0,21)	97	0,91 (0,17)		
Monat 12	89	0,88 (0,18)	88	0,87 (0,20)	-0,02 [-0,07; 0,04] η^2 = 0,002 P = 0,575	-0,01 [-0,07; 0,05] η^2 = 0,001 P = 0,676
Auswertepopulation C						
Fmean (Session 1) [kg]						
Baseline	93	17,0 (9,2)	89	15,9 (7,7)		
Monat 12	85	16,3 (8,3)	83	13,5 (7,7)	-2,6 [-5,0; -0,1] η^2 = 0,025 P = 0,042	-1,2 [-2,8; 0,4] η^2 = 0,013 P = 0,143
Fmean (Session 2) [kg]						
Baseline	92	15,4 (9,3)	89	14,9 (7,3)		
Monat 12	85	14,8 (8,2)	83	12,2 (7,8)	-2,6 [-5,0; -0,1] η^2 = 0,025 P = 0,041	-1,3 [-2,9; 0,3] η^2 = 0,015 P = 0,119
Fmax (Session 1) [kg]						
Baseline	93	20,9 (9,7)	89	20,2 (9,0)		
Monat 12	85	20,3 (9,7)	83	16,8 (9,4)	-3,4 [-6,4; 0,5] η^2 = 0,032 P = 0,021	-1,9 [-3,9; 0,1] η^2 = 0,021 P = 0,065
Fmax (Session 2) [kg]						
Baseline	92	18,9 (10,4)	89	18,3 (8,2)		
Monat 12	85	17,9 (9,3)	83	15,2 (9,6)	-2,7 [-5,6; 0,2] η^2 = 0,021 P = 0,063	-1,5 [-3,4; 0,5] η^2 = 0,014 P = 0,133

Gefördert durch: GBA Innovationsausschuss

Handgriff-kraft	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert [a, b]
Fatigue Ratio (Session 1) [c]						
Baseline	93	1,31 (0,28)	83	1,28 (0,19)		
Monat 12	85	1,30 (0,28)	89	1,27 (0,26)	-2,1 % [-7,1; 3,1] η^2 = 0,004 P = 0,417	-2,6 % [-7,4; 2,5] η^2 = 0,007 P = 0,307
Fatigue Ratio (Session 2) [c]						
Baseline	92	1,34 (0,47)	89	1,28 (0,21)		
Monat 12	85	1,26 (0,19)	83	1,27 (0,49)	2,0 % [-3,1; 7,4] η^2 = 0,003 P = 0,447	-0,0 % [-5,1; 5,4] η^2 < 0,001 P = 1,000
Recovery Ratio						
Baseline	92	0,96 (0,21)	89	0,91 (0,18)		
Monat 12	85	0,89 (0,18)	83	0,87 (0,18)	-0,02 [-0,07; 0,04] η^2 = 0,002 P = 0,536	-0,01 [-0,07; 0,04] η^2 = 0,002 P = 0,618

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.
[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.
[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.
[c] Die Differenzen stellen prozentuale Veränderungen dar; abgeleitet aus einem Modell mit logarithmisch transformierter abhängiger Variable.

3.3.7 Schellong-Test (Anlehn-Stehtest)

Ein positiver Schellong-Test war definiert als Abfall des systolischen Blutdrucks um mindestens 20 mmHg oder des diastolischen Blutdrucks um mindestens 10 mmHg während der Stehphase oder ein Anstieg der Herzfrequenz um mindestens 30 Schlägen pro Minute oder eine absolute Herzfrequenz von mindestens 120 Schlägen pro Minute im Stehen. Die Kontrollgruppe wies in allen vier Messzeitpunkten (nach 0, 2, 5 und 10 Minuten Stehen) bereits zu Baseline einen höheren Anteil positiver Testergebnisse auf als die Interventionsgruppe. Beispielsweise betrug der Anteil der positiv getesteten Patient:innen nach 5 Minuten Stehen in Auswertepopulation B zur Baseline 54,6 % in der Kontrollgruppe, verglichen mit 31,9 % in der Interventionsgruppe. Dieses Ungleichgewicht blieb auch nach 12 Monaten bestehen. Entsprechend lagen die Odds Ratios < 1. Lediglich das adjustierte Odds Ratio für das Vorliegen eines positiven Schellong-Tests nach 0 Minuten Stehen in Auswertepopulation C war leicht größer als eins (adjustiertes Odds Ratio: 1,00 mit 95-% KI: [0,51; 2,33]). Diese Ergebnisse sind jedoch mit erheblicher statistischer Unsicherheit behaftet; die 95-% KIs der Odds Ratios reichen von etwa 0,3 bis knapp 2,4.

Gefördert durch: GBA Innovationsausschuss

Tabelle 12: Ergebnisse der Analyse der Schellong-Tests (nach 0, 2, 5 und 10 Minuten Stehen).

Schellong-Test	Kontrollgruppe		Interventionsgruppe		Odds Ratio [95-% Konfidenzintervall]	
	n	Positiv (%)	n	Positiv (%)	Unadjustiert	Adjustiert [a, b]
Auswertepopulation B						
Nach 0 Minuten Stehen						
Baseline	99	51 (51,5)	96	27 (28,1)		
Monat 12	87	33 (37,9)	88	25 (29,6)	0,69 [0,36; 1,29] P = 0,243	1,00 [0,49; 2,25] P = 0,679
Nach 2 Minuten Stehen						
Baseline	98	47 (48,0)	95	25 (26,3)		
Monat 12	85	41 (48,2)	88	30 (34,1)	0,56 [0,30; 1,03] P = 0,061	0,84 [0,40; 1,75] P = 0,700
Nach 5 Minuten Stehen						
Baseline	97	53 (54,6)	94	30 (31,9)		
Monat 12	82	47 (57,3)	81	27 (33,3)	0,40 [0,21; 0,73] P = 0,003	0,69 [0,35; 1,36] P = 0,279
Nach 10 Minuten Stehen						
Baseline	95	56 (59,0)	92	33 (35,9)		
Monat 12	80	42 (52,5)	76	28 (36,8)	0,53 [0,28; 1,01] P = 0,052	0,89 [0,43; 1,86] P = 0,755
Auswertepopulation C						
Nach 0 Minuten Stehen						
Baseline	92	50 (54,4)	88	26 (29,6)		
Monat 12	85	31 (36,5)	83	25 (30,1)	0,75 [0,39; 1,44] P = 0,385	1,09 [0,51; 2,33] P = 0,815
Nach 2 Minuten Stehen						
Baseline	91	43 (47,3)	87	23 (26,4)		
Monat 12	83	39 (46,7)	83	29 (34,9)	0,61 [0,32; 1,14] P = 0,117	0,89 [0,42; 1,89] P = 0,757
Nach 5 Minuten Stehen						
Baseline	90	51 (56,7)	86	28 (32,6)		
Monat 12	80	45 (56,3)	76	27 (35,5)	0,43 [0,22; 0,82] P = 0,011	0,62 [0,30; 1,30] P = 0,204
Nach 10 Minuten Stehen						
Baseline	88	55 (62,5)	85	31 (36,5)		
Monat 12	78	40 (51,3)	71	26 (36,6)	0,55 [0,28; 1,06] P = 0,075	0,93 [0,42; 2,06] P = 0,854

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.
[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.
[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Gefördert durch: GBA Innovationsausschuss

3.3.8 Durchschnittliche Schrittzahl pro Tag

Die objektive körperliche Aktivität, gemessen anhand der durchschnittlichen Schritte pro Tag mit einem Akzelerometer, ist in Tabelle 13 dargestellt. Da fast 90 % der Daten für Monat 6 aufgrund technischer und logistischer Probleme fehlten, wurde auf eine longitudinale Auswertung verzichtet. Stattdessen wurde im Rahmen der Auswertung lediglich der Vergleich zu Monat 12 durchgeführt. Für die Analyse wurden die mittels des Troiano-Algorithmus (Troiano, et al., 2008) ermittelten Werte herangezogen, da diese im Vergleich zu alternativen Berechnungen die plausibelsten Schätzungen der Schrittzahl lieferten. Zu Baseline zeigten beide Gruppen in Auswertepopulation B eine sehr ähnliche mittlere Schrittzahl: Die Kontrollgruppe lief im Durchschnitt etwa 5.050 Schritte pro Tag (Standardabweichung: knapp 2.650 Schritte pro Tag) und die Interventionsgruppe bei etwa 4.680 Schritten (Standardabweichung: etwa 2.550 Schritte pro Tag). Die hohen Standardabweichungen deuten auf eine sehr große Heterogenität der körperlichen Aktivität innerhalb der Gruppen hin oder können Ausdruck möglicher Messungenauigkeiten sein (z.B. Auslösen der Schrittzählung bei Erschütterungen beim Autofahren o.ä.). Nach 12 Monaten hatte die Interventionsgruppe im Vergleich zur Kontrollgruppe in Auswertepopulation C keinen relevanten Vorteil erzielt. Die adjustierte Differenz der mittleren Schrittzahl pro Tag betrug lediglich 149 Schritte pro Tag (mit 95-% KI: [−648; 946]) zugunsten der Interventionsgruppe. Diese Ergebnisse müssen jedoch äußerst vorsichtig interpretiert werden, da viele Messungen (zu Monat 12) fehlten und es zu Ungenauigkeiten der Messungen durch die erwähnten Störfaktoren kommen konnte. So sind die Unterschiede in den Ergebnissen zu Auswertepopulation B sind im Wesentliche auf die hohe Streuung zurückzuführen.

Tabelle 13: Ergebnisse der Analyse der durchschnittlichen Schrittzahl pro Tag (mittels Troiano-Algorithmus ermittelt).

Schrittzahl [Schritte/Tag] nach Troiano	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert [a, b]
Auswertepopulation B						
Baseline	88	5.053 (2.648)	91	4.679 (2.560)		
Monat 6	11	4.409 (2.425)	14	3.492 (2.353)		
Monat 12	62	4.304 (2.685)	50	4.655 (2.876)	352 [-693; 1.396] η^2 = 0,004 P = 0,506	-459 [-3.186; 2268] η^2 = 0,005 P = 0,735
Auswertepopulation C						
Baseline	84	5.113 (2.685)	84	4.643 (2.463)		
Monat 6	11	4.409 (2.425)	14	3.492 (2.353)		
Monat 12	58	4.421 (2.697)	47	4.558 (2.880)	137 [-944; 1.220] η^2 < 0,001 P = 0,802	149 [-648; 946] η^2 = 0,002 P = 0,711

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.
[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.
[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Gefördert durch: GBA Innovationsausschuss

3.3.9 IPAQ

Die mittels IPAQ erhobene subjektive körperliche Aktivität wurden in MET-Minuten gemessen. MET-Minuten sind das Produkt aus dem Metabolischen Äquivalent (MET) einer Aktivität und deren Dauer. Wegen einer missverständlichen Implementierung des IPAQ-Fragebogens konnte Dauer der täglichen Aktivitätslevel nicht wie üblich bei 240 Minuten begrenzt werden. Da 1 MET dem Ruheumsatz entspricht, reflektiert eine höhere Anzahl an MET-Minuten eine höhere körperliche Belastung. Die Ergebnisse der Analyse sind in Tabelle 14 für die Auswertepopulationen B und C dargestellt. Auffallend sind die hohen Standardabweichungen und die Tatsache, dass die Anzahl fehlender Werte mit steigendem Aktivitätsniveau abnimmt. Aufgrund eines hohen Anteils an fehlenden Werten für Gehen/leichte Aktivität sind für dieses Aktivitätslevel nur die deskriptiven Statistiken dargestellt. Nach 12 Monaten lag die adjustierte Differenz der anstrengenden körperlichen Aktivität (prozentuale Änderung der MET-Minuten/Woche) zwischen Interventions- und Kontrollgruppe in Auswertepopulation B bei −12,9 % (95-%-KI: [−37,8; 39,5]). Ein negatives Vorzeichen steht hierbei für ein geringeres Aktivitätsniveau in der Interventionsgruppe. Demgegenüber lag die adjustierte Differenz der moderaten körperlichen Aktivität bei 45,0 % (95-% KI: [−11,8; 138,3]) und war damit leicht positiv. Die Unterschiede im Vergleich zu Auswertepoluation C ließen sich auf fehlende Werte bzw. die Imputation zurückführen. Aufgrund sehr vieler fehlender Werte bei den selbstberichteten Aktivitätsangaben und der missverständlichen Implementierung des Fragebogens sind die Ergebnisse des IPAQ jedoch insgesamt mit äußerster Vorsicht zu interpretieren.

Tabelle 14: Ergebnisse der Analyse des Physical Activity Questionnaire (Gehen/leichte körperliche Aktivität, moderate körperliche Aktivität, anstrengende körperliche Aktivität).

IPAQ [MET-Minuten/ Woche]	Kontrollgruppe n	Kontrollgruppe MW (SD) Median (IQR)	Interventionsgruppe n	Interventionsgruppe MW (SD) Median (IQR)	Differenz [95-% Konfidenzintervall] Unadjustiert [c]	Differenz [95-% Konfidenzintervall] Adjustiert [a, b, c]
Auswertepopulation B						
Gehen/leichte körperliche Aktivität [d]						
Baseline	11	865 (1.599) 297 (173 - 594)	17	562 (624) 248 (198 - 594)		
Monat 12	17	670 (955) 396 (198 - 594)	13	680 (1.011) 396 (264 - 693)		
Moderate körperliche Aktivität						
Baseline	55	1.355 (2.639) 480 (190 - 960)	52	1.265 (1.172) 880 (360 - 1.680)		
Monat 12	54	969 (1.478) 480 (240 - 960)	36	1.186 (1.075) 750 (450 - 1.680)	61,8 % [-0,8; 164] η^2 = 0,042 P = 0,054	45,0 % [-11,8; 138,3] η^2 = 0,022 P = 0,140
Anstrengende körperliche Aktivität						
Baseline	83	3.166 (4.097) 1.680 (680 - 3.684)	80	2.611 (2.947) 1.680 (812 - 3.360)		
Monat 12	77	2.747 (3.640) 1.200 (480 - 3.360)	73	2.545 (3.309) 1.200 (760 - 3360)	2,2 % [-33,2; 56,4] η^2 < 0,001 P = 0,918	-12,9 % [-37,8; 39,5] η^2 = 0,004 P = 0,629

Gefördert durch: GBA Innovationsausschuss

IPAQ [MET-Minuten/ Woche]	Kontrollgruppe n	Kontrollgruppe MW (SD) Median (IQR)	Interventionsgruppe n	Interventionsgruppe MW (SD) Median (IQR)	Differenz [95-% Konfidenzintervall] Unadjustiert [c]	Differenz [95-% Konfidenzintervall] Adjustiert [a, b, c]
Auswertepopulation C						
Gehen/leichte körperliche Aktivität [d]						
Baseline	10	943 (1.664) 347 (198 - 693)	15	598 (657) 248 (182 - 693)		
Monat 12	16	709 (973) 396 (272 - 594)	12	721 (1.045) 396 (289 - 693)		
Moderate körperliche Aktivität						
Baseline	53	1.403 (2.677) 480 (240 - 960)	48	1.230 (1.188) 760 (344 – 1.680)		
Monat 12	52	993 (1.500) 480 (240 - 960)	35	1.145 (1.061) 720 (420 - 1.600)	49,7 % [-8,2; 144] η^2 = 0,031 P = 0,104	1,5 % [-44,7; 86,3] η^2 < 0,001 P = 0,961
Anstrengende körperliche Aktivität						
Baseline	77	3.287 (4.172) 1.680 (720 - 3.760)	75	2350 (624) 1680 (784 - 3.360)		
Monat 12	75	2.761 (3.679) 1.200 (480 – 3.360)	71	1.302 (624) 1.200 (764 - 3360)	13,0 % [-23,3; 66,6] η^2 = 0,003 P = 0,533	-6,9 % [-37,8; 39,5] η^2 < 0,001 P = 0,729

MW: Mittelwert (arithmetisch), SD: Standardabweichung, IQR: Interquartilsspannweite, n: Fallzahl.
[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.
[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.
[c] Die Differenzen stellen prozentuale Veränderungen dar; abgeleitet aus einem Modell mit logarithmisch transformierter abhängiger Variable.
[d] Aufgrund des hohen Anteils an fehlenden Werten sind nur die deskriptiven Statistiken dargestellt.

3.3.10 Blutparameter

Die Blutparameterwerte sind in Tabelle 15 für die Auswertepopulationen B und C dargestellt und waren bei den meisten Patienten normwertig, Interleukin-8 nach Erylyse war bei einem relevanten Anteil erhöht. Über den Beobachtungszeitraum zeigten sich nahezu konstante Albumin-, anorganische Phosphat- und Calciumwerte und in diesen Parametern zeigten sich keine nennenswerten Differenzen zwischen den Gruppen. Bemerkenswert ist die geringe Streuung der Calciumwerte; diese ist ursächlich für die engen 95-%-KIs der Differenzen sowie für das relativ hohe partielle η^2. Die Differenzen aller Parameter zu Monat 12 zeigten mit Ausnahme von Albumin in Auswertepopulation B höhere Werte in der Interventionsgruppe an. Die größten prozentualen Differenzen in Auswertepopulation B zeigten sich bei NT-pro-BNP gefolgt von Interleukin-8 nach Erylyse und Creatinkinase. Die adjustierten Differenzen betrugen 10,4 %, 7,4 % und 6,8 %, wobei die zugehörigen 95-%-Konfidenzintervalle ([−19,8; 51,9], [−18,4; 41,4] bzw. [−7,7; 23,5]) allesamt relativ breit waren. Unterschiede zu Auswertepopulation C ließen sich auf fehlende Werte bzw. die Imputation zurückführen.

Gefördert durch: GBA Innovationsausschuss

Tabelle 15: Ergebnisse der Blutparameter Albumin, anorganisches Phosphat, Calcium, Creatinkinase, Ferritin, Interleukin-8 nach Erylyse, Laktatdehydrogenase und NT-pro-BNP.

Blut-parameter	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD) Median (IQR)	n	MW (SD) Median (IQR)	Unadjustiert	Adjustiert [a, b]
Auswertepopulation B						
Albumin [g/l]						Referenz: 35,0-52,0 g/l
Baseline	60	45,5 (2,8) 45,0 (44,0 – 47,1)	51	45,1 (2,5) 44,9 (43,6 - 47,1)		
Monat 12	67	45,2 (2,5) 45,0 (43,7 - 46,5)	61	44,9 (2,6) 44,8 (43,1 - 46,1)	-0,3 [-1,2; 0,6] η^2 = 0,004 P = 0,493	-0,0 [-0,8; 0,7] η^2 = 0,002 P = 0,959
Anorganisches Phosphat [mmol/l]						Referenz: 0,87-1,45 mmol/l
Baseline	60	0,90 (0,16) 0,88 (0,78 - 1,01)	51	0,94 (0,18) 0,94 (0,81 - 1,07)		
Monat 12	67	0,91 (0,18) 0,87 (0,78 - 0,99)	61	0,96 (0,18) 0,95 (0,84 - 1,08)	0,05 [-0,01; 0,12] η^2 = 0,022 P = 0,094	0,04 [-0,02; 0,10] η^2 = 0,016 P = 0,222
Calcium [mmol/l]						Referenz: 2,15-2,50 mmol/l
Baseline	58	2,35 (0,09) 2,35 (2,31 - 2,39)	47	2,35 (0,09) 2,36 (2,29 - 2,42)		
Monat 12	66	2,33 (0,09) 2,33 (2,28 - 2,38)	58	2,36 (0,09) 2,36 (2,30 - 2,42)	0,03 [-0,00; 0,06] η^2 = 0,026 P = 0,069	0,03 [0,00; 0,06] η^2 = 0,038 P = 0,026
Creatinkinase [U/l] [c]						Referenz: < 167 U/l bei Frauen, < 190 U/l bei Männern
Baseline	58	81,3 (48,6) 72,0 (53,7 - 92,2)	51	73,9 (27,5) 70,0 (54,0 - 89,0)		
Monat 12	66	75,7 (34,9) 69,0 (57,0 - 81,5)	61	81,0 (50,3) 73,0 (57,0 - 89,0)	3,2 % [-10,2; 18,8] η^2 = 0,002 P = 0,652	6,8 % [-7,7; 23,5] η^2 = 0,012 P = 0,371
Ferritin [µg/l] [c]						Referenz: 13,0-150,0 µg/l bei Frauen, 30-400 µg/l bei Männern
Baseline	60	87,6 (80,1) 58,6 (37,8 - 112,4)	51	79,1 (56,6) 57,0 (40,7 - 117,0)		
Monat 12	67	80,4 (67,2) 59,4 (34,2 - 114,1)	61	73,0 (54,2) 55,0 (34,4 - 98,0)	3,3 % [-26,2; 26,7] η^2 < 0,001 P = 0,807	0,0 % [-15,3; 18,2] η^2 = 0,002 P = 0,996
Interleukin-8 nach Erylyse [pg/ml] [c]						Referenz: 135-250 U/l
Baseline	58	203 (313) 140 (111 - 190)	49	179 (113) 149 (120 - 191)		

Gefördert durch: GBA Innovationsausschuss

Blut-parameter	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD) Median (IQR)	n	MW (SD) Median (IQR)	Unadjustiert	Adjustiert [a, b]
Monat 12	68	167 (64) 156 (125 - 198)	60	215 (285) 170 (135 - 203)	11,9 % [-4,3; 30,9] η^2 = 0,016 P = 0,157	7,4 % [-18,4; 41,4] η^2 = 0,016 P = 0,593
Laktatdehydrogenase [U/l] [c]						Referenz: 135-250 U/l
Baseline	60	204 (38) 199 (183 - 230)	47	196 (31) 194 (173 - 221)		
Monat 12	67	215 (37) 217 (188 - 233)	58	230 (86) 222 (194 - 249)	3,8 % [-3,8; 12,0] η^2 = 0,007 P = 0,334	5,7 % [-2,0; 14,1] η^2 = 0,019 P = 0,150
NT-pro-BNP [ng/l] [c]					Referenz: nach Alter und Geschlecht allenfalls ≤ 386 ng/l	
Baseline	58	56,2 (49,6) 42,5 (19,0 - 79,5)	47	72,1 (63,7) 63,0 (27,5 - 88,0)		
Monat 12	67	57,4 (55,4) 50,0 (22,0 - 67,5)	58	82,4 (77,8) 59,0 (28,0 - 114,0)	41,6 % [1,4; 97,7] η^2 = 0,033 P = 0,041	10,4 % [-19,8; 51,9] η^2 = 0,007 P = 0,537
Auswertepopulation C						
Albumin [g/l]						Referenz: 35,0-52,0 g/l
Baseline	58	45,4 (2,8) 44,9 (43,9 - 47,1)	47	44,9 (2,5) 44,6 (43,1 - 46,3)		
Monat 12	66	45,2 (2,5) 45,1 (43,7 - 46,5)	58	44,9 (2,7) 44,9 (43,1 - 46,2)	-0,3 [-1,2; 0,6] η^2 = 0,003 P = 0,524	0,1 [-0,8; 0,9] η^2 < 0,001 P = 0,917
Anorganisches Phosphat [mmol/l]						Referenz: 0,87-1,45 mmol/l
Baseline	58	0,90 (0,16) 0,88 (0,78 - 1,01)	47	0,93 (0,18) 0,94 (0,81 - 1,06)		
Monat 12	66	0,91 (0,18) 0,87 (0,78 - 1,00)	58	0,96 (0,17) 0,95 (0,84 - 1,08)	0,05 [-0,01; 0,11] η^2 = 0,020 P = 0,117	0,02 [-0,04; 0,09] η^2 = 0,005 P = 0,486
Calcium [mmol/l]						Referenz: 2,15-2,50 mmol/l
Baseline	58	2,35 (0,09) 2,35 (2,30 - 2,39)	47	2,35 (0,09) 2,36 (2,29 - 2,42)		
Monat 12	66	2,33 (0,09) 2,33 (2,28 - 2,38)	58	2,36 (0,09) 2,36 (2,30 - 2,42)	0,03 [-0,00; 0,06] η^2 = 0,029 P = 0,060	0,04 [0,01; 0,08] η^2 = 0,059 P = 0,020
Creatinkinase [U/l] [c]						Referenz: < 167 U/l bei Frauen, < 190 U/l bei Männern
Baseline	58	79,8 (48,4) 71,0 (53,3 - 84,8)	47	75,0 (27,8) 75,0 (54,0 - 93,0)		
Monat 12	66	75,7 (35,2) 68,5 (57,0 - 82,8)	58	82,1 (51,1) 73,0 (59,0 - 92,0)	4,8 % [-9,1; 20,9] η^2 = 0,003 P = 0,515	11,0 % [-2,0; 25,6] η^2 = 0,030 P = 0,099
Ferritin [µg/l] [c]						Referenz: 13,0-150,0 µg/l bei Frauen, 30-400 µg/l bei Männern
Baseline	58	85,6 (80,2)	47	79,2 (58,5)		

Gefördert durch: GBA Innovationsausschuss

Blut-parameter	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD) Median (IQR)	n	MW (SD) Median (IQR)	Unadjustiert	Adjustiert [a, b]
		57,8 (37,3 - 103,9)		57,0 (42,3 - 117,1)		
Monat 12	66	79,4 (67,2) 59,4 (34,2 - 108,6)	58	71,3 (51,0) 54,3 (34,5 - 97,8)	-3,2 % [-26,4; 27,2] η^2 < 0,001 P = 0,812	6,5 % [-11,6; 28,3] η^2 = 0,005 P = 0,503
Interleukin-8 nach Erylyse [pg/ml] [c]						Referenz: < 150,0 pg/ml
Baseline	56	200 (318) 140 (110 - 187)	45	160 (68) 149 (120 - 189)		
Monat 12	67	167 (64) 154 (125 - 195)	57	209 (289) 164 (134 - 199)	9,2 % [-6,6; 27,8] η^2 = 0,010 P = 0,268	10,0 % [-7,5; 30,7] η^2 = 0,014 P = 0,279
Laktatdehydrogenase [U/l] [c]						Referenz: 135-250 U/l
Baseline	58	204 (39) 196 (180 - 234)	47	198 (31) 195 (179 - 224)		
Monat 12	66	215 (37) 216 (188 - 233)	58	229 (86) 223 (194 - 249)	3,9 % [-3,9; 12,1] η^2 = 0,008 P = 0,319	6,5 % [-1,9; 15,6] η^2 = 0,025 P = 0,131
NT-pro-BNP [ng/l] [c]						Referenz: nach Alter und Geschlecht allenfalls ≤ 386 ng/l
Baseline	58	56,6 (50,0) 42,5 (19,3 - 80,3)	47	69,3 (61,5) 63,0 (27,5 - 87,5)		
Monat 12	66	56,8 (55,6) 49,5 (21,0 - 66,3)	58	79,6 (77,1) 54,5 (28,0 - 112,8)	38,0 % [-1,8; 93,8] η^2 = 0,028 P = 0,063	22,9 % [-5,7; 60,0] η^2 = 0,026 P = 0,125

MW: Mittelwert (arithmetisch), SD: Standardabweichung, IQR: Interquartilsspannweite, n: Fallzahl.
[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.
[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.
[c] Die Differenzen stellen prozentuale Veränderungen dar; abgeleitet aus einem Modell mit logarithmisch transformierter abhängiger Variable.

3.3.11 PHQ-9

Die PHQ-9-Skala umfasst Werte von 0 bis 27 Punkten, wobei höhere Werte einen stärkeren Schweregrad depressiver Symptome anzeigen. Beim PHQ-9 muss berücksichtigt werden, dass drei Items (Fatigue, Schlafstörungen und Konzentrationsprobleme) mit den Kernsymptomen von ME/CFS überlappen. Der PHQ-9 ist daher zur Beurteilung der Schwere depressiver Symptome in dieser Population nur eingeschränkt bzw. nicht valide interpretierbar.

Tabelle 16 zeigt die Ergebnisse der Analyse für die Auswertepopulationen B und C. Die Mittelwerte in beiden Auswertepopulation unterschieden sich in der Interventions- und Kontrollgruppe zu beiden Zeitpunkten kaum voneinander und lagen zwischen 11,4 und 12,9 Punkten (mit einer Standardabweichung von etwa 4 bis 5 Punkten). Nach Adjustierung betrug die mittlere Differenz zu Monat 12 −0,1 Punkte (95 %-KI: [−1,3; 1,0]) in Auswertepopulation C und −0,0 Punkte (95 %-KI: [−1,1; 1,1]) in Auswertepopulation B, wobei ein negativer Wert der Differenz einer Besserung entspricht.

Gefördert durch: GBA Innovationsausschuss

Tabelle 16: Ergebnisse der Analyse des Patient Health Questionnaires.

PHQ-9 [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert [a, b]
Auswertepopulation B						
Baseline	98	13,2 (4,7)	92	12,9 (4,1)		
Monat 12	94	11,4 (4,9)	94	11,9 (4,3)	0,5 [-0,9; 1,8] η^2 = 0,003 P = 0,491	-0,0 [-1,1; 1,1] η^2 < 0,001 P = 0,987
Auswertepopulation C						
Baseline	91	12,8 (4,3)	86	12,9 (4,2)		
Monat 12	92	11,4 (4,9)	89	11,9 (4,3)	0,5 [-0,9; 1,8] η^2 = 0,002 P = 0,513	-0,1 [-1,3; 1,0] η^2 < 0,001 P = 0,806

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.
[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.
[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

3.3.12 ESS

Die Ergebnisse der Analyse der ESS, deren Werte von 0 bis 24 Punkten reichen und höhere Werte einen ausgeprägteren Schweregrad der Schläfrigkeit anzeigen, sind in Tabelle 17 dargestellt. Mit einem Mittelwert von 9,6 Punkten in der Kontrollgruppe und 9,7 und 9,5 Punkten in der Interventionsgruppe in den Auswertepopulationen B bzw. C waren die Scores zur Tagesschläfrigkeit in beiden Gruppen zu Monat 12 nahezu identisch. Die mittlere Differenz betrug nach Adjustierung −0,5 Punkte (95 %-KI: [−1,6; 0,6]) bzw. −0,7 Punkte (95 %-KI: [−1,9; 0,5]), wobei negative Werte des Punktschätzers auf bessere Werte der Interventionsgruppe im Vergleich zur Kontrollgruppe hindeuten.

Tabelle 17: Ergebnisse der Analyse der Epworth Sleepiness Scale.

ESS [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert [a, b]
Auswertepopulation B						
Baseline	92	10,2 (5,1)	92	10,2 (5,3)		
Monat 12	93	9,6 (5,5)	94	9,7 (5,7)	-0,1 [-1,5; 1,7] η^2 < 0,001 P = 0,882	-0,5 [-1,5; 0,6] η^2 = 0,004 P = 0,383
Auswertepopulation C						
Baseline	85	9,9 (5,0)	86	10,2 (5,4)		
Monat 12	91	9,6 (5,4)	89	9,5 (5,5)	-0,1 [-1,8; 1,5] η^2 < 0,001 P = 0,861	-0,7 [-1,9; 0,5] η^2 = 0,009 P = 0,221

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.
[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.
[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Gefördert durch: GBA Innovationsausschuss

3.3.13 EQ-5D-5L Indexwert

Der EQ-5D-5L-Index zur Beschreibung des Gesundheitszustandes weist einen Wertebereich von negativen Werten bis zum Maximalwert 1,0 auf. Negative Indexwerte repräsentieren einen Gesundheitszustand, der subjektiv als schlechter als der Tod eingestuft wird. Der Wert 1,0 entspricht dem besten möglichen Gesundheitszustand. Die Ergebnisse der Analyse sind in Tabelle 18 dargestellt. Zu Monat 12 lagen die Mittelwerte beider Gruppen nahe beieinander (0,51 Punkte in der Kontrollgruppe vs. 0,49 Punkte in der Interventionsgruppe). Die adjustierten und unadjustieren Differenzen waren nahezu identisch; die adjustierte Differenz betrug im Mittel 0,02 Punkte (95 %-KI: [−0,10; 0,10]) in Auswertepopulation C und −0,02 Punkte (95 %-KI: [−0,09; 0,07]) in Auswertepopulation B, wobei negative Differenzen für eine schlechtere gesundheitsbezogene Lebensqualität in der Interventionsgruppe stehen. Die Effektstärken (partielles η^2) waren mit ungefähr 0,001 aber nur gering.

Tabelle 18: Ergebnisse der Analyse des European Quality of Life 5 Dimensions 5 Level Version Indexwerts.

EQ-5D-5L Indexwert [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert [a, b]
Auswertepopulation B						
Baseline	99	0,46 (0,25)	89	0,48 (0,25)		
Monat 12	95	0,51 (0,29)	94	0,49 (0,26)	-0,02 [-0,10; 0,05] η^2 = 0,002 P = 0,550	-0,01 [-0,09; 0,07] η^2 < 0,001 P = 0,758
Auswertepopulation C						
Baseline	92	0,47 (0,25)	82	0,47 (0,24)		
Monat 12	93	0,51 (0,29)	89	0,49 (0,26)	-0,01 [-0,09; 0,07] η^2 < 0,001 P = 0,742	0,02 [-0,07; 0,10] η^2 = 0,001 P = 0,684

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.
[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.
[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

3.4 Analyse der Krankenkassendaten

3.4.1 EQ-5D-5L Indexwerte und Qualität-adjustierte Lebensjahre (QALY)

Im Rahmen der gesundheitsökonomischen Evaluation wurden die EQ-5D-5L Indexwerte sowie die „quality adjusted life years“ (QALYs) in den Auswertepopulationen B und C zwischen Interventions- und Kontrollgruppe verglichen (Tabelle 19).

Zu Studienbeginn wiesen die BKK-Versicherungsnehmer:innen in Auswertepopulation B in der Interventionsgruppe und in der Kontrollgruppe einen mittleren EQ-5D-5L Indexwert von 0,56 bzw. 0,43 Punkten auf, wobei die Standardabweichung bei 0,21 bzw. 0,26 lag. Nach 12 Monaten zeigte sich ein Anstieg in der Kontrollgruppe und ein Abfall in der Interventionsgruppe. Der Wert stieg auf 0,53 Punkte (Standardabweichung: 0,31 Punkte) in

Gefördert durch: GBA Innovationsausschuss

der Kontrollgruppe und sank auf 0,48 Punkte (Standardabweichung: 0,25 Punkte) in der Interventionsgruppe. Die unadjustierte mittlere Differenz zwischen Interventions- und Kontrollgruppe betrug entsprechend –0,05 Punkte (95 %-KI: [–0,23; 0,12]) und die adjustierte Differenz –0,03 Punkte (95 %-KI: [–0,24; 0,18]). In Auswertepopulation C zeigte sich dagegen ein Anstieg in beiden Gruppen von 0,48 Punkten zu Baseline auf 0,52 Punkte (Standardabweichung: 0,31 Punkte) in der Kontrollgruppe und 0,55 Punkte (Standardabweichung: 0,25 Punkte) in der Interventionsgruppe. Die adjustierte und unadjustierte Differenz zu Monat 12 waren jedoch nahezu identisch zu Auswertepopulation B.

Die über 12 Monate berechneten QALYs lagen (unabhängig von der verwendeten Berechnungsmethode oder Auswertepopulation) im Mittel bei etwa 0,50 Punkten mit einer Standardabweichung von etwa 0,2 Punkten in der Kontrollgruppe und 0,16-0,17 in der Interventionsgruppe. Entsprechend gering fielen die sowohl die unadjustierten als auch die adjustierten Differenzen aus.

Tabelle 19: Ergebnisse der Analyse der EQ-5D-5L-Indexwerte und quality-adjusted life-years (QALYs).

EQ-5D-5L Indexwert und QALYs [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert [a, b]
BKK-Versicherungsnehmer:innen in Auswertepopulation B						
EQ-5D-5L Indexwerte						
Baseline	25	0,43 (0,26)	20	0,56 (0,21)		
Monat 12	22	0,53 (0,31)	20	0,48 (0,25)	-0,05 [-0,23; 0,12] η^2 = 0,010 P = 0,538	-0,03 [-0,24; 0,18] η^2 = 0,004 P = 0,747
QALYs						
Variante 1 [c]	22	0,51 (0,20)	19	0,52 (0,16)	0,00 [-0,11; 0,12] η^2 < 0,001 P = 0,934	0,04 [-0,10; 0,17] η^2 = 0,008 P = 0,592
Variante 2 [d]	25	0,45 (0,25)	21	0,49 (0,17)	0,02 [-0,11; 0,16] η^2 = 0,003 P = 0,703	-0,01 [-0,14; 0,12] η^2 = 0,002 P = 0,888
BKK-Versicherungsnehmer:innen in Auswertepopulation C						
EQ-5D-5L Indexwerte						
Baseline	21	0,48 (0,22)	19	0,48 (0,20)		
Monat 12	21	0,52 (0,31)	20	0,55 (0,25)	-0,04 [-0,22; 0,14] η^2 = 0,005 P = 0,661	-0,03 [-0,25; 0,20] η^2 = 0,002 P = 0,810
QALYs						

Gefördert durch: GBA Innovationsausschuss

EQ-5D-5L Indexwert und QALYs [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert [a, b]
Variante 1 [c]	21	0,50 (0,19)	19	0,52 (0,16)	0,02 [-0,10; 0,13] η^2 = 0,002 P = 0,762	-0,01 [-0,13; 0,12] η^2 < 0,001 P = 0,860
Variante 2 [d]	21	0,50 (0,19)	20	0,50 (0,17)	-0,02 [-0,13; 0,10] η^2 = 0,002 P = 0,773	-0,05 [-0,17; 0,08] η^2 = 0,018 P = 0,442

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.
[a] Adjustiert für Baseline (nur bei EQ-5D-5L Indexwert), Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.
[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.
[c] QALYs wurden über Baseline und Monat 12 für beide Gruppen geschätzt.
[d] QALYs wurden über Baseline und Monat 12 für die Kontrollgruppe und über Baseline, Monat 3, 6, 9 und 12 für die Interventionsgruppe geschätzt.

3.4.2 Gesundheitskosten im Zeitraum 3 Monate bis 12 Monate

Tabelle 20 zeigt die Gegenüberstellung der Gesundheitskosten im Zeitraum 3 Monate bis 12 Monate, die sich aus sämtlichen stationäre und ambulante Abrechnungsdaten, die im Erhebungszeitraum anfielen, zusammensetzen. Dazu gehören insbesondere Krankenhausleistungen (Dauer, Diagnosen, Prozeduren, Kosten), Rehabilitationsleistungen (Art der Leistung, Diagnose, Verordnungsdatum, Dauer, Kosten), Leistungen im Sinne der Arzneimittel-, Heilmittel-, Hilfsmittel-, Fahrtkosten-, Häusliche Krankenpflege-, Haushaltshilfeverordnungen, aber ebenso vertragsärztliche Leistungen (Datum, Fachgruppe des Arztes oder der Ärztin, Diagnose, Einzelleistungen, Kosten), sowie Leistungen der sozialen Pflegeversicherung gemäß SGB XI. Der Median der monatlichen Kosten lag in Auswertepopulation B in der Interventionsgruppe bei 400 € (Interquartilsspannweite: 226 – 649 €) und in der Kontrollgruppe bei 219 € (Interquartilsspannweite: 123 – 361 €). Die unadjustierte Differenz zwischen Interventions- und Kontrollgruppe betrug 50,6 % (95 %-KI: [–16,9; 173,0]); d.h. die Interventionsgruppe tendierte zu höheren Kosten. Nach Adjustierung verringerte sich diese Differenz auf 35,1 % (95 %-KI: [–33,3; 173,8]). Der große Unterschied zu Auswertepopulation C (adjustierte Differenz von 3,5 % mit 95 %-KI: [–52,3; 124,4]) ist auf einzelne einflussreiche Beobachtungen zurückzuführen.

Gefördert durch: GBA Innovationsausschuss

Tabelle 20: Ergebnisse der Analyse der monatlichen Gesundheitskosten.

Kosten [€/Monat]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD) Median (IQR)	n	MW (SD) Median (IQR)	Unadjustiert [c]	Adjustiert [a, b, c]
BKK-Versicherungsnehmer:innen in Auswertepopulation B						
Monat 3 bis Monat 12	25	431 (728) 219 (123 - 361)	21	664 (922) 400 (226 - 649)	50,6 % [-16,9; 173,0] η^2 = 0,042 P = 0,172	35,1 % [-33,3; 173,8] η^2 = 0,019 P = 0,393
BKK-Versicherungsnehmer:innen in Auswertepopulation C						
Monat 3 bis Monat 12	20	471 (789) 219 (123 - 465)	21	662 (946) 378 (222 - 594)	39,5 % [-27,4; 168,0] η^2 = 0,027 P = 0,309	3,5 % [-52,3; 124,4] η^2 < 0,001 P = 0,928

MW: Mittelwert (arithmetisch), SD: Standardabweichung, IQR: Interquartilsspannweite, n: Fallzahl.
[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.
[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.
[c] Die Differenzen stellen prozentuale Veränderungen dar; abgeleitet aus einem Modell mit logarithmisch transformierter abhängiger Variable.

3.4.3 Kosteneffizienzanalyse

Tabelle 21 zeigt die Ergebnisse der Kosteneffizienzanalyse in den Auswertepopulationen B und C. Die unadjustierten durchschnittlichen Gesamtkosten pro Patient:in und Monat betrugen in Auswertepopulation B 659 € (95-% KI: [380; 1.059]) in der Interventionsgruppe und 438 € (95-% KI: [258; 691]) in der Kontrollgruppe. Die Interventionsgruppe erzielte im Durchschnitt einen minimal höheren QALY-Wert (0,52 vs. 0,51). Daraus ergab sich eine inkrementelle Kostendifferenz von 222 € (95 %-KI: [–77; 594]) und eine inkrementelle QALY-Differenz von < 0,01 (95 %-KI: [–0,10; 0,11]). Der daraus berechnete ICER lag bei 46.821 € / QALY. Nach Adjustierung für Baseline und Kovariate verringerten sich die durchschnittlichen Gesamtkosten auf 587 € (95-% KI: [459; 609]) in der Interventionsgruppe und blieben mit 434 € (95-% KI: [342; 462]) in der Kontrollgruppe nahezu unverändert. Die adjustierten QALY-Werte waren mit 0,47 vs. 0,52 etwas geringer in der Kontrollgruppe. Dies entspricht einer inkrementellen Kostendifferenz von 153 € (95 %-KI: [73; 190]) und einer inkrementellen QALY-Differenz von 0,04 (95-% KI: [0,01; 0,05]). Alle berichteten Konfidenzintervalle im Rahmen der Kosteneffizienzanalyse basieren auf einem nicht-parametrischen Bootstrap mit 2.000 Replikationen. Abbildung 3 zeigt die Kosten-Effektivitäts-Ebene der mittels Bootstrap ermittelten Streupunkte (blaue Punkte). Jeder Punkt stellt eine simulierte Schätzung der inkrementellen Kosten (y-Achse) und inkrementellen QALYs (x-Achse) der Interventionsgruppe im Vergleich zur Kontrollgruppe dar. Die gelbe Ellipse markiert den 95-% Konfidenzellipse der simulierten Inkrementalschätzungen. Das bedeutet, dass 95 % der simulierten Streupunkte innerhalb dieser Ellipse liegen und somit die Unsicherheit der Kosteneffektivität widerspiegeln. Die rote Linie zeigt die Kosten-Effektivitäts-Schwelle von 50.000 € pro QALY an. Punkte unterhalb dieser Linie gelten als kosteneffektiv, da sie eine geringere Kosten-pro-QALY-Rate aufweisen. Etwa die Hälfte der Bootstrap-Schätzungen sowie der Punktschätzer (gelbe Raute) liegen im linken oberhalb der Kosten-Effektivitäts-Schwelle. Folglich verursacht die Intervention höhere inkrementelle Kosten und führt aber gleichzeitig zu einem (leichten) Gewinn an QALYs und ist als nicht kosteneffizient zu bewerten. In

Gefördert durch: GBA Innovationsausschuss

Auswertepopulation C führt die Intervention dagegen bei etwas höheren Kosten zu einem (leichten) Verlust an QALYs. Diese Unterschiede zwischen den Auswertepopulationen sind auf einzelne einflussreiche Beobachtungen zurückzuführen.

Tabelle 21: Ergebnisse der Kosteneffizienzanalyse.

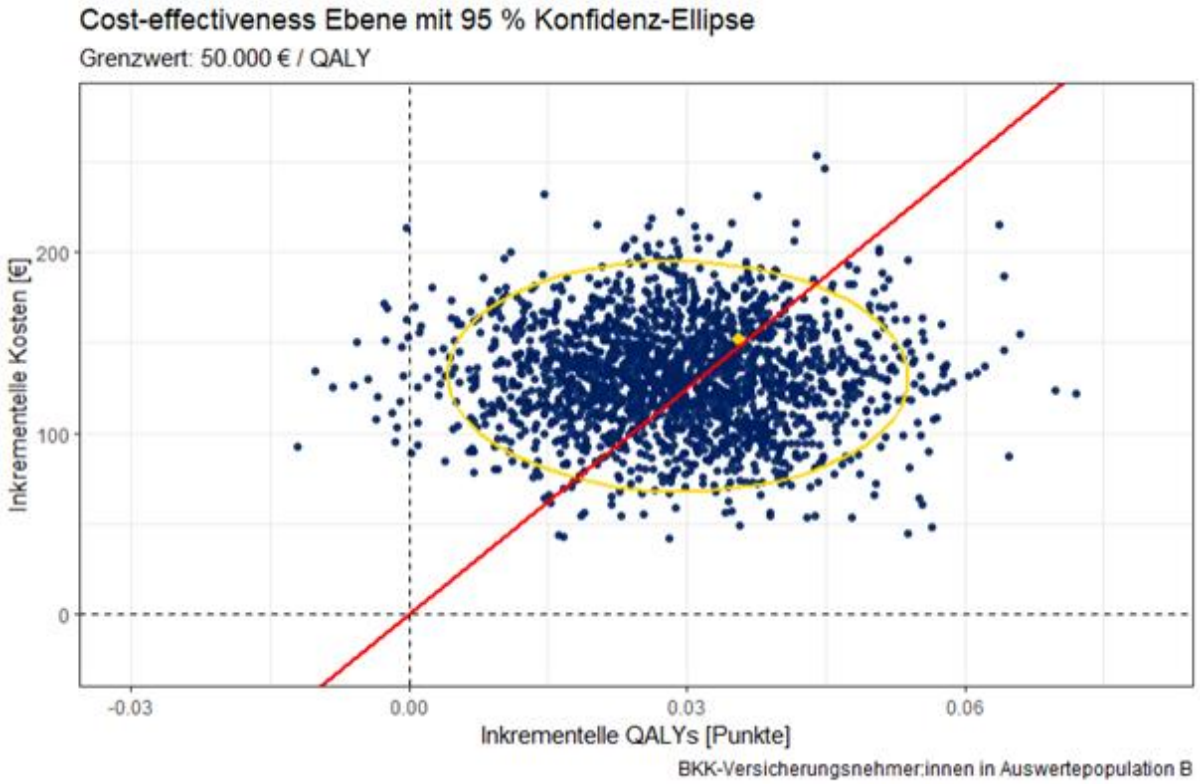
BKK-Versicherungsnehmer:innen in Auswertepopulation B	Mittelwert [95-% Konfidenzintervall] Kontrollgruppe N = 25	Interventionsgruppe N = 21
Unadjustierter Durchschnitt [95-% Konfidenzintervall]		
Gesamtkosten [€] [a, b]	438 [258; 691]	659 [380; 1.059]
QALYs [Punkte] [b]	0,51 [0,43; 0,59]	0,52 [0,45; 0,59]
Inkrementelle Kosten [€] [b]	222 [-77; 594]	
Inkrementelle QALYs [Punkte] [b]	0,00 [-0,10; 0,11)	
ICER [€ / QALY] [b]	46.821 [-70.619; 47.311])	
Adjustierter Durchschnitt [95-% Konfidenzintervall] [c, d]		
Gesamtkosten [€] [a]	434 [342; 462]	587 [459; 609]
QALYs [Punkte] [a]	0,47 [0,46; 0,50]	0,51 [0,49; 0,52]
Inkrementelle Kosten [€] [b]	153 [73; 190]	
Inkrementelle QALYs [Punkte] [b]	0,04 [0,01; 0,05)	
ICER [€ / QALY] [b]	4.275 [1.800; 17.746])	

BKK-Versicherungsnehmer:innen in Auswertepopulation C	Mittelwert [95-% Konfidenzintervall] Kontrollgruppe N = 20	Interventionsgruppe N = 21
Unadjustierter Durchschnitt [95-% Konfidenzintervall]		
Gesamtkosten [€] [a, b]	474 [258; 800]	661 [354; 1.078]
QALYs [Punkte] [b]	0,50 [0,42; 0,58]	0,52 [0,45; 0,59]
Inkrementelle Kosten [€] [b]	188 [-181; 588]	
Inkrementelle QALYs [Punkte] [b]	0,02 [-0,09; 0,12)	
ICER [€ / QALY] [b]	10.928 [-70.848; 66.665])	
Adjustierter Durchschnitt [95-% Konfidenzintervall] [c, d]		
Gesamtkosten [€] [a]	518 [394; 554]	536 [412; 585]
QALYs [Punkte] [a]	0,52 [0,50; 0,54]	0,51 [0,48; 0,53]
Inkrementelle Kosten [€] [b]	19 [-41; 92]	
Inkrementelle QALYs [Punkte] [b]	-0,01 [-0,04; 0,01)	
ICER [€ / QALY] [b]	1.636 [-18.977; 14.314])	

[a] Kosten (€) ist ein geometrisches Mittel, das der rücktransformierte Mittelwert aus transformierten Daten ist (log-Kosten). Der geschätzte Mittelwert der log-Kosten wird auf eine exponentielle Skala rücktransformiert, bereinigt um Duans Verschmierungsfaktor (Duan, 1983).

Gefördert durch: GBA Innovationsausschuss

[b] Die 95-% Konfidenzintervalle mittels nichtparametrischem Bootstrapping geschätzt, um 2.000 Schätzungen der inkrementellen Kosten und Effekte zu erhalten.
[c] Adjustiert für Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.
[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.



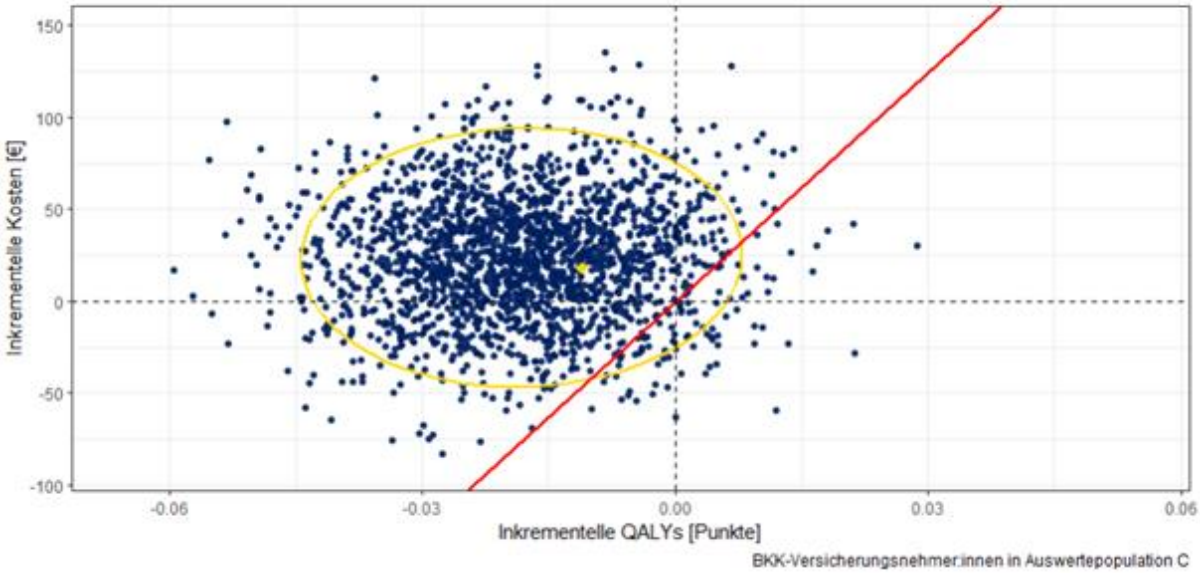


Abbildung 3: Kosten-Effektivitäts-Ebene mit 95-% Konfidenzellipse.

Gefördert durch: GBA Innovationsausschuss

3.4.4 Krankentage im Zeitraum 3 Monate bis 12 Monate

Die Ergebnisse der Krankentage im Zeitraum von 3 bis 12 Monaten nach Studienbeginn sind in Tabelle 22 dargestellt. Die Interventionsgruppe in Auswertepopulation B hatte im Durchschnitt mehr Krankheitstage (175 Tage) als die Kontrollgruppe (132 Tage), was etwa 65 bzw. fast 50 % des Beobachtungszeitraums entspricht. Darüber hinaus ist die Streuung in beiden Gruppen mit Standardabweichungen von deutlich mehr als 100 Tagen auch sehr hoch. Dies deutet auf große individuelle Unterschiede hin. Die Mittelwertdifferenz verringerte sich durch Adjustierung von 44 Tagen (95-% KI: [–26; 112]) auf 11 Tage (95-% KI: [–64; 86]). Die Breite der 95-% KIs legen jedoch nahe, dass es keinen relevanten Unterschied in den Krankentagen in diesem Zeitraum zwischen den beiden Gruppen gibt. Die Unterschiede zu Auswertepopulation C lassen sich durch die hohe Streuung erklären.

Tabelle 22: Krankentage im Zeitraum 3 Monate bis 12 Monate nach Studienbeginn.

Krankentage [Tage]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert [a, b]
BKK-Versicherungsnehmer:innen in Auswertepopulation B						
Monat 3 bis Monat 12	25	132 (124)	21	175 (103)	44 [-26; 112] η^2 = 0,036 P = 0,209	11 [-64; 86] η^2 = 0,002 P = 0,771
BKK-Versicherungsnehmer:innen in Auswertepopulation C						
Monat 3 bis Monat 12	21	125 (125)	20	184 (98)	60 [-12; 130] η^2 = 0,068 P = 0,100	42 [-40; 123] η^2 = 0,032 P = 0,301

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

3.5 Ergebnisse der Prozessevaluation

Das fünfwöchige Rehabilitationsprogramm wurde von allen 97 Patient:innen vollständig und ohne vorzeitige Abbrüche absolviert. Die Mehrheit profitierte subjektiv von der Maßnahme, insbesondere im Hinblick auf das Erlernen von Krankheitsbewältigungsstrategien sowie eine verbesserte Alltagsbewältigung (

Abbildung 4: Selbsteinschätzung des Rehabilitationserfolgs der Patient:innen.). Pacing, sowie Behandlungsstrategien für Schmerzen, Atmungs- und Entspannungstechniken wurden gut vermittelt. Auch die Kontaktmöglichkeiten zu anderen ME/CFS Erkrankten während der Rehabilitation hat den Patient:innen geholfen. Die Patient:innen fühlten sich am Ende der Reha besser informiert über die Erkrankung, fühlten sich auch besser in der Lage ihre Erkrankung bewältigen zu können und Behandlungskonzepte umzusetzen und damit ihren Alltag besser gestalten. Die meisten würden zu einem späteren Zeitpunkt an einer erneuten Rehabilitation für ME/CFS teilnehmen. Auch die Untersuchungen und Beratungen zu Studienbeginn an der Charité wurden als hilfreich empfunden.

Gefördert durch: GBA Innovationsausschuss

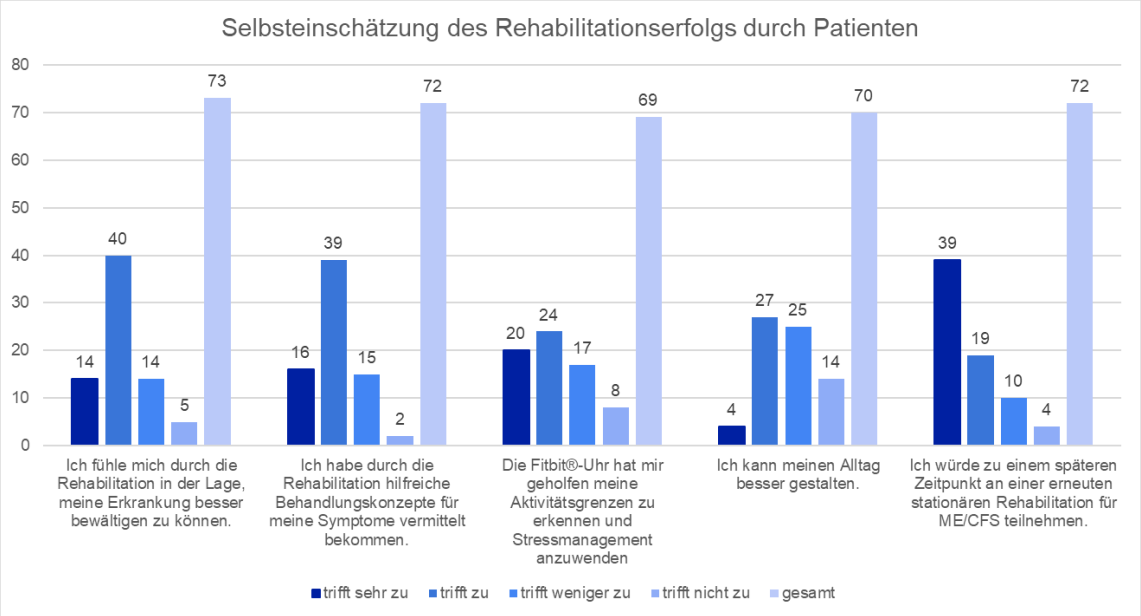


Abbildung 4: Selbsteinschätzung des Rehabilitationserfolgs der Patient:innen.

Demgegenüber stand jedoch bei vielen Teilnehmenden eine temporäre Verschlechterung des Zustands, bedingt durch die Anreise oder im weiteren Verlauf der Rehabilitation. Entsprechend zeigte sich beim Bell-Score nach der Rehabilitation (Monat 6 beziehungsweise 9) ein negatives Bild: Bei 44,7 % der Patient:innen war der Score niedriger als zuvor, während lediglich 13,8 % eine Verbesserung aufwiesen (Differenz *zwischen dem Bell-Score vor und nach der Rehabilitation.*).

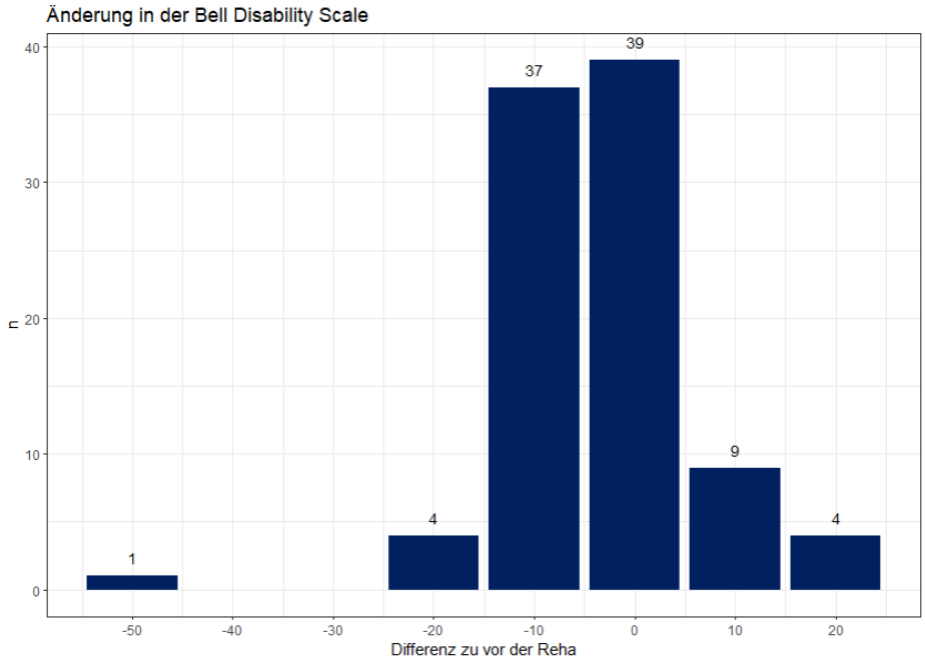


Abbildung 5: Differenz zwischen dem Bell-Score vor und nach der Rehabilitation.

Gefördert durch: GBA Innovationsausschuss

4 Diskussion der Projektergebnisse

4.1 Effektevaluation

4.1.1 Primärer Endpunkt: Körperliche Funktionsfähigkeit (SF-36)

Die körperliche Funktionsfähigkeit (SF-36) nach 12 Monaten wurde in Auswertepopulation B (in Anlehnung an das ITT-Prinzip). Die Interventionsgruppe hatte nach 12 Monaten mit durchschnittlich 36,0 Punkten schlechtere Werte als die Kontrollgruppe mit durchschnittlich 41,8 Punkten, was einer unadjustierten Differenz von −5,8 Punkten (95-% KI: [−12,6; 1,0]) entsprach. Das negative Vorzeichen deutet auf einen potentiell negativen Effekt (bezogen auf die körperliche Funktionsfähigkeit) der Intervention hin. Nach Adjustierung verringerte sich die Differenz auf −3,8 Punkte (95-% KI: [−9,9; 2,2]). Die Reduktion legt nahe, dass der ursprüngliche Unterschied hauptsächlich auf Baseline-Unterschiede oder Konfounder zurückzuführen war, die durch Adjustierung kontrolliert und eliminiert wurden. Der Anteil der Varianz der körperlichen Funktionsfähigkeit, der spezifisch durch die Intervention erklärt werden konnte, nachdem der Effekt aller übrigen Kovariate herausgerechnet wurde, war sehr gering (partielles η^2 = 0,008). Dies legt nahe, dass die Intervention keine klinisch relevante negative Auswirkung auf die körperliche Funktionsfähigkeit der Patient:innen hatte. In der Auswertepopulation C, die den „best case“ abzuschätzen sollte, waren die Ergebnisse nahezu identisch. Die Interventionsgruppe hatte nach 12 Monaten mit durchschnittlich 41,6 Punkten schlechtere Werte als die Kontrollgruppe mit durchschnittlich 36,8. Die adjustierte Differenz war mit −1,9 Punkten (95-% KI: [−8,2 4,5]) aber nominal etwas geringer.

Da das Ziel der Intervention eine Verbesserung der körperlichen Funktionsfähigkeit war, ist das negative Vorzeichen der Differenzen kontraintuitiv. Daher wurde im Rahmen von Sensitivitätsanalysen auf anderen Auswertepopulationen die Stabilität dieser Befunde untersucht. Folgende Sensitivitätsanalysen wurden durchgeführt:

- In Auswertepopulation A: Diese Auswertepopulation schloss alle Patient:innen ein, die zu einem Behandlungsarm allokiert wurden, unabhängig von der tatsächlichen Adhärenz. Fehlende Werte wurden imputiert.
- Auswertepopulation P: Diese Auswertepopulation schloss alle Patient:innen aus, die die Studie nicht wie geplant abgeschlossen hatten. Fehlende Werte wurde imputiert.
- Complete-Case Analyse.

Die Ergebnisse dieser Analysen zeigten, dass die ursprünglich festgestellte Tendenz zu schlechteren Werten in der Interventionsgruppe in allen alternativen Analysen bestehen bleibt. Die Differenzen waren nominal jedoch teils größer. In Auswertepopulation A betrug die adjustierte Differenz −4,1 Punkte (95-% KI: [−9,4; 1,3]). In Auswertepopulation P, die im Wesentlichen einer Variante von Auswertepopulation C mit Imputation entspricht, wurde eine adjustierte Differenz von −3,2 (95-% KI: [−9,3; 2,8]) ermittelt. Die Complete-Case Analyse lieferte mit einer adjustierten Differenz von −2,0 Punkten (95-% KI: [−8,5; 4,6]) ebenfalls ein negatives Vorzeichen. Der Unterschied zur den Auswertepopulationen B und C ergibt sich im Wesentlichen aus der Imputation und insbesondere durch Patient:innen mit fehlendem ISCED. Die Sensitivitätsanalysen bestätigten insgesamt die Robustheit des negativen Vorzeichens. Unabhängig von der Population oder der statistischen Adjustierung konnte kein positiver Effekt der Intervention auf die körperliche Funktionsfähigkeit nachgewiesen werden. Die durchgängig negative Tendenz und geringe Effektstärke (partielles η^2) war in allen Sensitivitätsanalysen gegeben.

Ein möglicher Grund für das negative Vorzeichen des Interventionseffekt könnte die Reha selbst sein. Die Reha wurde in dieser Studie zwar an die Bedürfnisse von ME/CFS-Patient:innen

Gefördert durch: GBA Innovationsausschuss

angepasst, was in Standard- Reha Programmen üblicherweise nicht der Fall ist, jedoch könnte auch diese Anpassung negative Folge einer Überlastung möglicherweise nicht vollständig ausschließen. Die Hauptgefahr liegt im Risiko der PEM einer Zustandsverschlechterung nach bereits geringer körperlicher oder kognitiver Belastung. Schon die Anreise zur Rehaklinik, sowie die Teilnahme am Programm selbst, können eine PEM ausgelöst haben. Selbst eine an die besondere Situation der ME/CFS Erkrankten angepasste Reha, könnte folglich bei einem Teil der Patienten eine akute oder kumulierte Belastung darstellen, die die körperliche Funktionsfähigkeit verschlechtert, was sich dann in einem negativen Vorzeichen widerspiegelt. Alle 97 Patienten konnten das fünfwöchige Rehabilitationsprogramm jedoch vollständig absolvieren; niemand musste die Maßnahme vorzeitig abbrechen. Die Mehrheit der Patienten gab an, dass die Rehabilitation hilfreich war, um Krankheitsbewältigungsstrategien zu erlernen und den Alltag danach besser zu bewältigen. Gleichzeitig berichteten viele Patienten über eine vorübergehende Verschlechterung, entweder durch die Anreise zur Rehabilitationsklinik oder im Verlauf der Rehabilitation selbst.

Eine weitere mögliche Ursache für das schlechtere Abschneiden der Interventionsgruppe könnte die Wartezeit für den Platz in der Rehaklinik sein. Während ursprünglich geplant war, dass die Patient:innen in der Interventionsgruppe nach drei Monaten das Reha-Programm antreten würden oder angetreten haben, kam es aufgrund von Kapazitätsproblemen wegen der höheren Rekrutierung in der 2. Hälfte der Studie in vielen Fällen erst deutlich später dazu. Im Durchschnitt warteten die Patient:innen etwa 158 Tage (mit Standardabweichung: 64 Tage) – also fast doppelt so lange wie ursprünglich geplant – auf den Platz in der Rehaklinik. Die mediane Wartedauer lag bei 154 Tagen (mit Interquartilsspannweite: 112−189 Tage).

Auch die wiederholten Visiten und Untersuchungen in der Interventionsgruppe (Eingangsuntersuchung, manchmal mehrere Besuchstage, Monat 6, Monat 9 und Monat 12) könnte ein Einfluss gehabt haben.

Ungeachtet der Richtung darf das negative Vorzeichen des Interventionseffekts aber auch nicht überbewertet werden. Die nominale Differenz von gerade 3,8 Punkten ist auf der 100-Punkte-Skala des SF-36 sehr klein und entspricht einer Änderung von etwa 4 % im Verhältnis zur Spannweite der Skala. Diese Änderung unterschreitet zudem die minimal clinically important difference (MCID), die in ähnlichen Kontexten (z. B. pädiatrisches ME/CFS), auf etwa 10 Punkte geschätzt wird (Brigden, Parslow, & Gaunt, 2018; Gaunt, Brigden, Metcalfe, Loades, & Crawley, 2023). Eine Änderung von 10 Punkten entspricht auch in etwa der MCID, wenn sie durch die eine halbe Standardabweichung der Baselinewerte (Geoffrey, Sloan, & Wyrwich, 2003) geschätzt wird (18,5 × 0,5 = 9,3 Punkte).

Im Rahmen von zwei Subgruppenanalysen wurde außerdem untersucht, ob der Interventionseffekt bei bestimmten Gruppen von Patient:innen unterschiedlich ausfällt (d.h., ob eine Interaktion zwischen der Gruppenzugehörigkeit und dem Subgruppenmerkmal besteht). Konkret wurden die folgenden Subgruppen untersucht:

- Geschlecht: Männer und Frauen.
- Schwere der Erkrankung: Schwer erkrankte (Bell-Score zu Baseline ≤ 30) und moderat erkrankte (Bell-Score zu Baseline ≥ 40) Patient:innen.

Die Subgruppenanalysen zeigten in der Mehrheit der untersuchten Subgruppen keinen relevanten Unterschied zwischen der Interventions- und der Kontrollgruppe. Bei Männern und schwerer Erkrankten (Bell-Score ≤ 30) war die Differenz zwischen den Gruppen vernachlässigbar. Hingegen zeigten Frauen und weniger schwer erkrankte Patienten (Bell-Score ≥ 40) in den Analysen schlechtere Werte in der Interventionsgruppe (Differenzen von −6,9 bzw. −5,2 Punkte in Auswertepopulation B). Die Punktschätzer sind aber mit einem gewissen Grad an statistischer Unsicherheit behaftet (95-% KI: [−15,3; 1,4] bzw. [−11,7; 1,3]). Im Großen und Ganzen lässt sich festhalten, dass die Intervention auch in keiner der

Gefördert durch: GBA Innovationsausschuss

untersuchten Subgruppen zu einer relevanten Verbesserung der körperlichen Funktionsfähigkeit im Vergleich zur Kontrollgruppe nach 12 Monaten führte und der Interventionseffekt sich nicht wesentlich zwischen den Subgruppen unterschied. Bei der Subgruppenanalyse nach Geschlecht muss jedoch auch beachtet werden, dass insgesamt nur relativ wenig Männer an der Studie teilnahmen; je nach Auswertepopulation betrug der Männeranteil etwa 13 bis 14 %. Diese Gruppengröße ist möglicherweise zu klein, um hier überhaupt Effekte (sofern vorhanden) zu identifizieren.

4.1.2 Sekundäre Endpunkte

Die Analyse der sekundären Endpunkte zeigte ebenfalls keine relevanten Unterschiede zwischen Interventions- und Kontrollgruppe. Die Punktschätzer der (adjustierten) Differenzen wiesen dabei aber keine konsistente Richtung (positiv oder negativ) auf.

Im Detail zeigte die Interventionsgruppe im Bell-Score einen minimal höheren Grad der Behinderung im Vergleich zur Kontrollgruppe (adjustierte Differenz in Auswertepopulation B: −4,4 Punkte mit 95 %-KI: [−8,1; −0,6]; adjustierte Differenz in Auswertepopulation C: −3,7 Punkte mit 95 %-KI: [7,5; 0,1]). Demgegenüber wiesen die Chalder Fatigue Scale und der COMPASS-31-Gesamtscore zwar negative Vorzeichen auf, die tendenziell eine geringere Belastung in der Interventionsgruppe indizierten. Diese Effekte waren jedoch extrem schwach, sodass keine klinisch relevante Veränderung identifiziert werden konnte. Unterstrichen wurde dies durch kleine Effektstärken (partielles η^2).

Die Analyse des SF-36 ergab, dass die Interventionsgruppe in den meisten Bereichen schlechter abschnitt, in der Regel jedoch nur mit vernachlässigbaren Unterschieden. Lediglich die emotionale Rollenfunktion zeigte einen scheinbaren Vorteil mit einer Differenz von 12,1 Punkten (95-% KI: [1,6; 22,7]). Wird wieder die halbe Baseline-Standardabweichung zur Schätzung der MCID herangezogen (Geoffrey, Sloan, & Wyrwich, 2003), ergibt sich jedoch ein Wert von 43,2 × 0,5 = 21,6 als Schätzer für die MCID. Folglich muss auch diese Differenz als klinisch nicht relevant bewertet werden. Zusammenfassend lässt sich sagen, dass die Intervention nicht zu einer überzeugenden Verbesserung der gesundheitsbezogenen Lebensqualität im Rahmen des SF-36 geführt.

Die Differenzen zwischen Interventions- und Kontrollgruppe waren bei den Handkraftmessungen (Fmean und Fmax) negativ, was auf eine schlechtere Leistung in der Interventionsgruppe hindeutet. Die gefundenen Änderungen von unter 2,0 kg sind, gemessen an der halben Baseline-Standardabweichung (von ca. 4,2 bzw. 4,7 kg), jedoch nicht klinisch relevant. Auch bei der orthostatischen Kreislaufregulation (Schellong-Test) wurden keine nachweisbaren Verbesserungen für die Interventionsgruppe gefunden.

Bezüglich der Akzelerometrie-Daten wurde in Auswertepopulation B ein Unterschied von durchschnittlich 459 Schritten pro Tag (95-% KI: [−648; 946]) weniger in der Interventionsgruppe festgestellt. Die Validität dieser Ergebnisse muss jedoch äußerst kritisch beurteilt werden: Die Verteilung der Messwerte wies unplausible Extremwerte auf und die Rate der fehlenden Werte war erheblich (fast 50 % in der Interventionsgruppe). Zusätzliche Korrelationsanalysen zeigten zudem keinerlei Assoziation zwischen den objektiven Messungen und den selbstberichteten Werten des IPAQ-Fragebogens, was auf eine mangelnde Validität der Erfassungsmethoden hindeuten könnte. Die hohe Varianz den

Gefördert durch: GBA Innovationsausschuss

individuellen Tragezeiten des Schrittzählers stellte zudem eine weitere Einschränkung für die Vergleichbarkeit der Messwerte dar.

Die Ergebnisse des IPAQs müssen aufgrund mehrerer methodischen und datenbezogenen Mängel äußerst vorsichtig interpretiert werden. Einerseits war der Datensatz durch ein hohes Maß an fehlenden Werten gekennzeichnet (insbesondere bei den niedrigen Aktivitätslevels; für die Kategorien Gehen/leichte Aktivität lagen insgesamt lediglich 12 valide Werte vor). Mit steigendem Aktivitätslevel nahm der Anteil an fehlenden Werten ab. Diese Beobachtung lässt sich möglicherweise dadurch erklären, dass bereits einfache Tätigkeiten von Patient:innen mit ME/CFS als erheblich anstrengend empfunden werden, was dazu geführt haben könnte, dass sie diese in den Fragebögen nicht als leichte oder moderate Aktivität angaben, da fast alle Aktivitäten für diese Patient:innen anstrengend und möglicher Weise nicht für mindestens 10 Minuten ausführbar waren. Bei der Erklärung des IPAQ-Fragebogens werden „anstrengende Aktivitäten“ als folgendes erklärt: „Aktivitäten, die starke körperliche Anstrengungen erfordern und bei denen Sie (Patient:in) deutlich stärker atmen als normal. Bitte denken Sie nur an diese körperlichen Aktivitäten, die Sie mindestens für 10 Minuten ohne Unterbrechung verrichtet haben.“ Zudem unterstreichen die hohen Standardabweichungen den hohen Grad an individueller Variation innerhalb der erfassten Aktivitätslevels. Darüber hinaus gab es Probleme mit der der Erfassung des IPAQ, die mögliche Messfehler in den Fragebogen begünstigte. Der im eCRF implementierte IPAQ erfragte die Tätigkeit zuerst nach der Häufigkeit (Tage pro Woche) und unmittelbar danach nach der Gesamtzeit (Minuten pro Woche) für die jeweilige Aktivität. Ein Blick auf die Verteilung der Rohdaten legt eine heterogene Interpretation dieser Formulierung nahe, da unplausibel hohe Werte auf die Angabe der Gesamtminuten pro Woche hindeuten, während andere vermutlich eher die Minuten pro Tag angaben. Dies könnte Verzerrungen zur Folge haben. Abgesehen davon war es dadurch nicht möglich die Dauer wie empfohlen auf 240 Minuten je Aktivitätslevel zu begrenzen. Zusätzliche Korrelationsanalysen zeigten außerdem keinerlei Assoziation zwischen den objektiven Messungen des Akzelerometers und den Angaben im IPAQ-Fragebogen. Angesichts dieser kumulierten Mängel in der Datenqualität und der Validität sollten die Ergebnisse des IPAQ daher also äußerst kritisch gesehen und wenn, dann nur mit erheblichen Einschränkungen interpretiert werden. Dies spricht dafür, dass dieser Fragebogen für diese Patient:innenkohorte nicht geeignet ist.

Die Blutparameter Albumin, anorganisches Phosphat sowie Calcium blieben über den Studienzeitraum unverändert. Bei den übrigen Parametern (Creatinkinase, Ferritin, Interleukin-8 nach Erylyse, Laktatdehydrogenase und NT-pro-BNP) wurde dagegen eine Veränderung beobachtet und die Differenzen zu Monat 12 zwischen Interventions- und Kontrollgruppe waren durchweg positiv (d.h. die Interventionsgruppe wies in den einzelnen Parametern immer größere Werte als die Kontrollgruppe auf). Allerdings sind die Punktschätzer aufgrund der breiten 95 %-KIs, wie exemplarisch bei Interleukin-8 in Auswertepopulation C ([−7,5; 30,7]), mit hoher statistischer Unsicherheit behaftet.

Die Ergebnisse des ESS und EQ-5D-5L zeigten, dass viele Patient:innen unter Schlafstörungen und einer schlechten Lebensqualität leiden. Die adjustierten Differenzen zwischen Interventions- und Kontrollgruppe für PHQ-9, ESS und EQ-5D-5L zu Monat 12 von höchstens −0,1, −0,7 bzw. 0,02 Punkten fielen äußerst gering aus. Werden für eine Bewertung der klinischen Relevanz die MCIDs wieder über die halbe Standardabweichung zu Baseline geschätzt (2,1 Punkte für den PHQ-9, 2,6 Punkte für die ESS und rund 0,1 Punkte für den EQ-5D-5L Indexwert), zeigt sich, dass die beobachteten Differenzen vermutlich keine für

Gefördert durch: GBA Innovationsausschuss

Patient:innen wahrnehmbare Veränderung widerspiegeln. Die Intervention ist demnach kein relevanter Nutzen hinsichtlich depressiver Symptomatik, Tagesschläfrigkeit oder gesundheitsbezogener Lebensqualität zuschreiben.

4.2 Gesundheitsökonomische Evaluation

Im Rahmen der gesundheitsökonomischen Evaluation wurden Unterschiede in Lebensqualität, QALYs, Gesundheitskosten oder Krankentagen zwischen Interventions- und Kontrollgruppe untersucht.

Beide Gruppen starteten mit ähnlichen EQ-5D-5L Indexwerten und stiegen über den Verlauf der Studie leicht an. Die resultierenden Gruppenunterschiede (sowohl unadjustiert als auch adjustiert) waren aber klein mit relativ breiten Konfidenzbändern. Auch bei den über 12 Monate berechneten QALYs ergaben sich keine bedeutsamen Differenzen. Die mittleren Werte lagen sehr nah beieinander und weder unadjustierte noch adjustierte Modelle zeigten statistisch oder klinisch relevante Abweichungen.

Die Gesundheitskosten zwischen Monat 3 und 12 waren in der Interventionsgruppe zwar etwas höher, doch die großen Unsicherheiten und die breite Spanne der Konfidenzintervalle lassen keine belastbaren Unterschiede erkennen. Nach Adjustierung ergab sich praktisch kein Kostenunterschied.

Die Intervention erwies sich in der Kosteneffektivitätsanalyse als ökonomisch nicht vorteilhaft. Ein leichter QALY-Gewinn bei höheren Kosten in Auswertepopulation B ließ sich in Population C nicht bestätigen. Bei einem Schwellenwert von 50.000 €/QALY ist die Maßnahme somit nicht als kosteneffektiv einzustufen.

Auch bei den Krankheitstagen (3 bis 12 Monate) ergab sich kein messbarer Unterschied zwischen den Gruppen. Trotz etwas höherer Durchschnittswerte in der Interventionsgruppe waren Streuung und Konfidenzintervalle so groß, dass kein belastbarer Gruppenunterschied erkennbar ist.

Insgesamt sprechen alle Analysen dafür, dass die Intervention weder hinsichtlich Lebensqualität, Kosten oder Krankheitstagen einen Vorteil gegenüber der Kontrollgruppe bietet.

4.3 Limitationen der Studie

Die Einordnung der vorliegenden Ergebnisse muss jedoch unter dem Gesichtspunkt mehrerer methodischer Einschränkungen und Limitationen dieser Studie erfolgen.

Erstens zeigte sich eine Problematik im Bereich der Datenqualität. Die Datensätze enthielten eine Reihe fehlender, widersprüchlicher oder offensichtlich unplausibler Angaben. Dafür waren verschiedene Faktoren verantwortlich – unter anderem die Einschränkungen durch ME/CFS selbst, die Patient:innen oftmals daran hinderten, Fragebögen vollständig oder konsistent auszufüllen. Aber auch organisatorische Probleme zu Beginn der Studie führten zu fehlenden Werten. Trotz intensiver Bereinigungs- und Validierungsversuche konnten letztendlich nicht alle Unstimmigkeiten aufgelöst werden. Zudem erwies sich die Qualität Daten aus dem IPAQ sowie den Akzelerometern als fragwürdig, sodass deren Aussagekraft eingeschränkt bleibt.

Zweitens war die Adhärenz zur zugewiesenen Studienbedingung beeinträchtigt. Eine beträchtliche Zahl von Teilnehmenden (35 von 137) musste anfangs aus der Interventionsgruppe wieder ausgeschlossen werden. Darüber hinaus mussten deutlich mehr Fälle als ursprünglich geplant ausgeschlossen werden, sodass letztendlich nur 93

Gefördert durch: GBA Innovationsausschuss

Patient:innen in der Kontroll- und 89 Patient:innen in der Interventionsgruppe die Studie wie geplant abschlossen. Ursprünglich geplant waren 120 Patient:innen je Gruppe.

Drittens betrug die durchschnittliche Wartezeit der Patient:innen auf einen Reha-Platz nahezu das Doppelte der ursprünglichen Planung. Diese Verzögerung könnte die Effektivität der Intervention reduziert haben. Durch diese Verzögerung wurde insbesondere die Definition der gesundheitsökonomischen Endpunkte (z. B. der Zeitraum von 3 bis 12 Monaten), die explizit darauf abzielte, den Effekt nach Abschluss der stationären Rehabilitation zu erfassen, unterminiert.

Viertens lagen für mehrere der verwendeten Instrumente keine krankheitsspezifisch validierten MCID-Werte vor. Dadurch war die klinische Interpretation beobachteter Differenzen erheblich erschwert. Mangels externer Referenzwerte mussten die MCIDs daher über die Variabilität der eigenen Baseline-Daten geschätzt werden, was die Vergleichbarkeit einschränkt und Unsicherheit erhöht.

Fünftens war aufgrund des Studiendesigns weder eine Randomisierung noch eine Verblindung möglich. Dadurch besteht das Risiko für Verzerrung – insbesondere bei Endpunkten, die auf subjektiven Einschätzungen beruhen, wie der primäre Endpunkt oder viele der sekundären Endpunkte.

Limitationen der gesundheitsökonomischen Analyse: Ein einschränkender Faktor der gesundheitsökonomischen Auswertung liegt darin, dass keine Routinedaten (wie Behandlungskosten oder Arbeitsunfähigkeitszeiten) für den Zeitraum *vor* Studieneintritt (Baseline) erhoben wurden. Ein retrospektiver Vorher-Nachher-Vergleich der Kostenentwicklung war somit nicht möglich. Der Erhebungszeitraum wurde bewusst ab Monat 3 gewählt, um die kumulierten Effekte und Krankentage exklusive der direkten Interventionsphase zu erfassen und post-interventionelle Effekte abzubilden.

5 Schlussfolgerungen und Empfehlungen des Evaluators

Diese Studie evaluierte ein interdisziplinäres Versorgungsmodell für Patient:innen mit ME/CFS, das auf umfassender Diagnostik sowie individuell abgestimmten Therapieempfehlungen basiert. Bestandteile des Programms waren unter anderem Aktivitätsmanagement (Pacing), Entspannungsverfahren, symptomorientierte Maßnahmen wie Schlaf- und Schmerzmanagement, Ernährungsberatung, psychosoziale Unterstützung und Hilfe bei der beruflichen Reintegration.

Im primären Endpunkt, der körperlichen Funktionsfähigkeit (SF-36), zeigte sich nach 12 Monaten jedoch kein statistisch signifikanter Unterschied zwischen Interventions- und Kontrollgruppe. Auch die sekundären Endpunkte wiesen keine klinisch relevanten Vorteile zugunsten der Intervention auf. Die gesundheitsökonomische Analyse ergab ebenfalls keinen Nutzen. Zwar berichteten in der Befragung im Anschluss der Reha viele Patient:innen, dass sie durch das Programm Pacing-Strategien besser verstanden und im Alltag anwenden konnten, doch führte dies scheinbar nicht zu einer Verbesserung der Funktionalität oder der Symptomatik im Verlauf. Zudem deutet die Studie darauf hin, dass die Belastungen durch Anreise und Teilnahme an Rehabilitationsmaßnahmen möglicherweise den potenziellen Nutzen der Intervention infrage stellen. Selbst eine Subgruppenanalyse nach Schwere der Erkrankung (Bell ≤ 30 vs. Bell ≥ 40) zeigte keine Unterschiede im primären Endpunkt. Besonders für schwerer betroffene Patient:innen erscheint daher die derzeitige Praxis, Rehabilitation als Voraussetzung für eine Erwerbsminderungsrente zu verlangen, als nicht angemessen.

Gefördert durch: GBA Innovationsausschuss

Spezialisierte Inhalte wie Pacing-Schulung sind zwar für Betroffene hilfreich, reichen jedoch nicht aus, um die Krankheitsentwicklung zu beeinflussen. Die Ergebnisse legen nahe, dass bestehende Rehabilitationskonzepte für ME/CFS kritisch überprüft werden sollten.

Die negativen Ergebnisse dieser Studie unterstreichen den dringenden Bedarf an der Erforschung kausaler Therapien, die auf den inzwischen gut belegten immunologischen Fehlregulationen, sowie der autonomen, endothelialen und mitochondrialen Dysfunktion bei ME/CFS basieren. Parallel dazu müssen ambulante Versorgungsstrukturen verbessert werden, um eine flächendeckende Diagnostik, sowie eine symptombezogene und adaptive Unterstützung (unter anderem Pacing) sicherstellen. Zukünftige Forschung sollte sich daher auf Mechanismus-basierte Interventionen und Versorgungsmodelle konzentrieren, die die Belastungsintoleranz der Patient:innen konsequent berücksichtigen.

IV Literaturverzeichnis

EuroQol Group. (3 1990). EuroQol-a new facility for the measurement of health-related quality of life. *Health Policy*.

Bell, D. S. (1995). *The Doctor's Guide to Chronic Fatigue Syndrome: Understanding, Treating and Living with CFIDS.* Da Capo Lifelong Books.

Borm, G. F., Fransen, J., & Lemmens, W. A. (12 2007). A simple sample size formula for analysis of covariance in randomized clinical trials. *Journal of Clinical Epidemiology*, S. 1234–1238.

Brigden, A., Parslow, R. M., & Gaunt, D. (12 2018). Defining the minimally clinically important difference of the SF-36 physical function subscale for paediatric CFS/ME: Triangulation using three different methods. *Health and Quality of Life Outcomes*.

Brown, M., Khorana, N., & Jason, L. A. (3 2011). The role of changes in activity as a function of perceived available and expended energy in nonpharmacological treatment outcomes for ME/CFS. *Journal of Clinical Psychology*, S. 253–260.

Carruthers, B. M., Jain, A. K., De Meirleir, K. L., Peterson, D. L., Klimas, N. G., Lerner, A. M., . . . van de Sande, M. I. (2 2003). Myalgic Encephalomyelitis/Chronic Fatigue Syndrome: Clinical Working Case Definition, Diagnostic and Treatment Protocols. *Journal of Chronic Fatigue Syndrome*, S. 7–115.

Chalder, T., Berelowitz, G., Pawlikowska, T., Watts, L., Wessely, S., Wright, D., & Wallace, E. P. (2 1993). Development of a fatigue scale. *Journal of Psychosomatic Research*, S. 147–153.

Craig, C. L. (August 2003). International physical activity questionnaire: 12-country reliability and validity. *Medicine and science in sports and exercise*, S. 1381–1395.

Craig, C. L., Marshall, A. L., Sjöström, M., Bauman, A. E., Booth, M. L., Ainsworth, B. E., . . . Oja, P. (8 2003). International physical activity questionnaire: 12-country reliability and validity. *Medicine & Science in Sports & Exercise*, S. 1381–1395.

Duan, N. (September 1983). Smearing Estimate: A Nonparametric Retransformation Method. *Journal of the American Statistical Association*, S. 605-610.

Gaunt, D., Brigden, A., Metcalfe, C., Loades, M., & Crawley, E. (2 2023). Investigating the factors associated with meaningful improvement on the SF-36-PFS and exploring the appropriateness of this measure for young people with ME/CFS accessing an NHS specialist service: a prospective cohort study. *BMJ Open*, S. e069110.

Gefördert durch: GBA Innovationsausschuss

Geoffrey, N. R., Sloan, J. A., & Wyrwich, K. W. (5 2003). Interpretation of changes in health-related quality of life: The remarkable universality of half a standard deviation. *Medical Care*, S. 458-467.

Jäkel, B. K. (19. April 2021). Hand grip strength and fatigability: correlation with clinical parameters and diagnostic suitability in ME/CFS. *Journal of translational medicine*, S. 159.

Jason, L., Benton, M., Torres-Harding, S., & Muldowney, K. (2 2009). The impact of energy modulation on physical functioning and fatigue severity among patients with ME/CFS. *Patient Education and Counseling*, S. 237–241.

Johns, M. W. (11 1991). A new method for measuring daytime sleepiness: the Epworth sleepiness scale. *Sleep*, S. 540–545.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (9 2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine*, S. 606–613.

Nacul, L. A.-L.-M. (19. May 2021). European Network on Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (EUROMENE): Expert Consensus on the Diagnosis, Service Provision, and Care of People with ME/CFS in Europe. *Medicina (Kaunas)*, S. 510.

Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys.* John Wiley and Sons.

Sletten, D. M., Suarez, G. A., Low, P. A., Mandrekar, J., & Singer, W. (12 2012). COMPASS 31: a refined and abbreviated Composite Autonomic Symptom Score. *Mayo Clinic Proceedings*, S. 1196–1201.

Troiano, R. P., Berrigan, D., Dodd, K. W., Mâsse, L. C., Tilert, T., & McDowell, M. (1 2008). Physical activity in the united states measured by accelerometer. *edicine & Science in Sports & Exercise*, S. 181–188.

UNESCO. (2011). *International standard classification of education: ISCED.*

Ware Jr, J. E., & Sherbourne, C. D. (6 1992). The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care*, S. 473–483.

V Anlagen

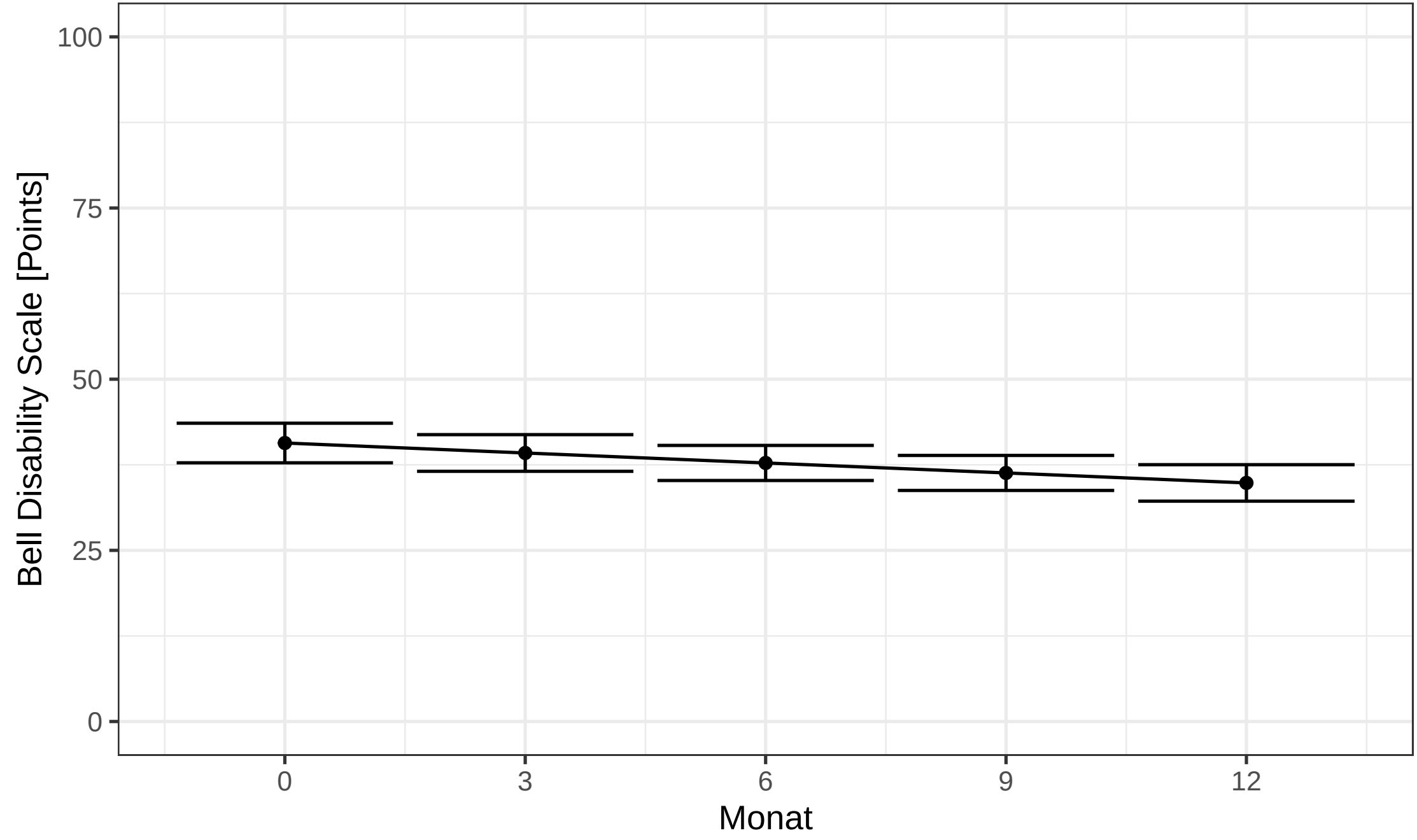
Anlage 1:	Graphische Darstellung der marginalen Mittelwerte (mit 95-%-Konfidenzintervall) der ausgewählten Endpunkte der Interventionsgruppe im zeitlichen Verlauf.
Anlage 2:	Ergebnisse aller Auswertepopulationen
Anlage 3:	Statistischer Analyseplan - Original
Anlage 4:	Statistischer Analyseplan – Änderungen
Anlage 5:	Baseline-Charakteristika der übrigen Auswertepopulationen
Anlage 6:	Anzahl der imputierten Werte je Outcome und Auswertepopulation

Gefördert durch: GBA Innovationsausschuss

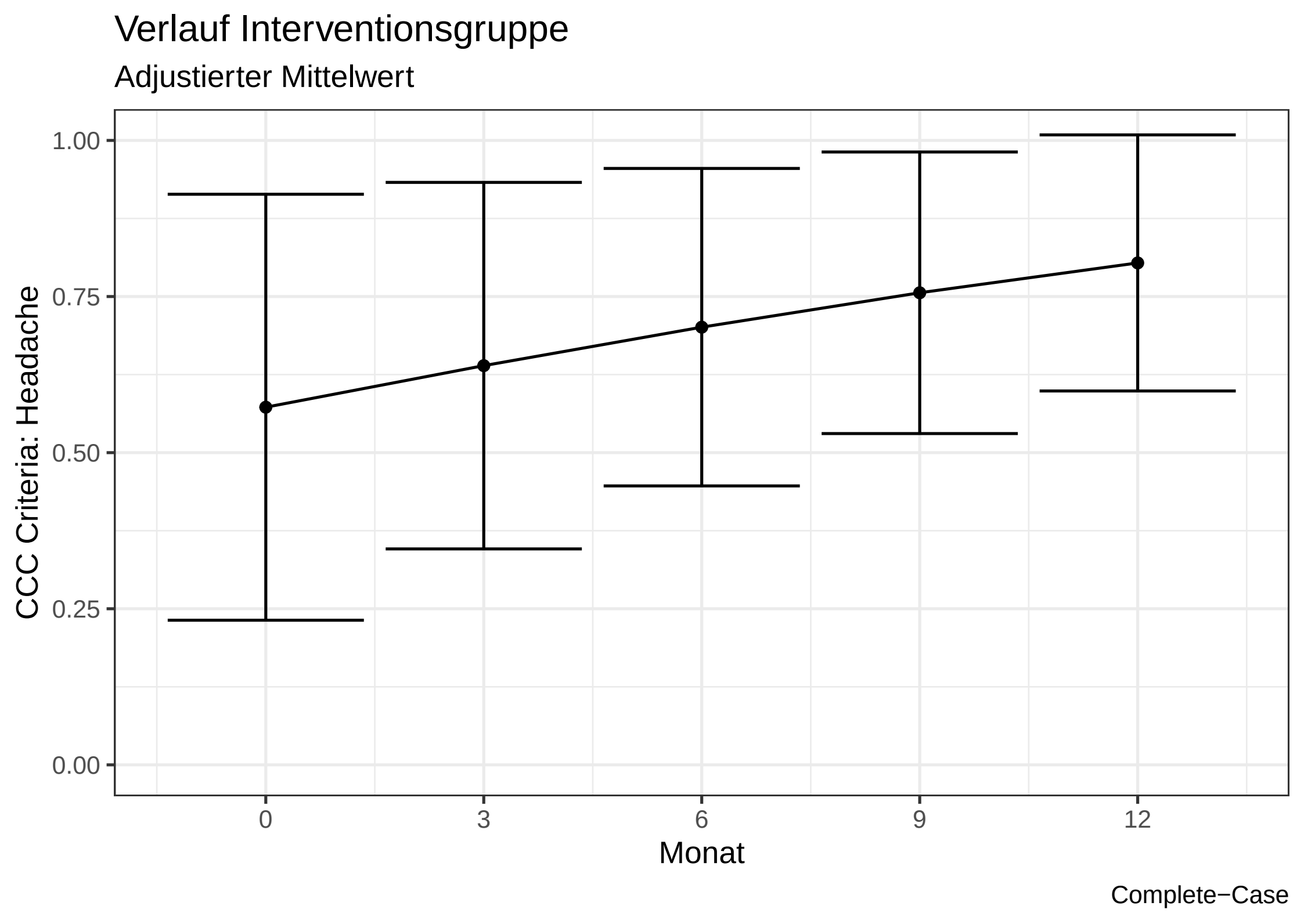
Verlauf Interventionsgruppe

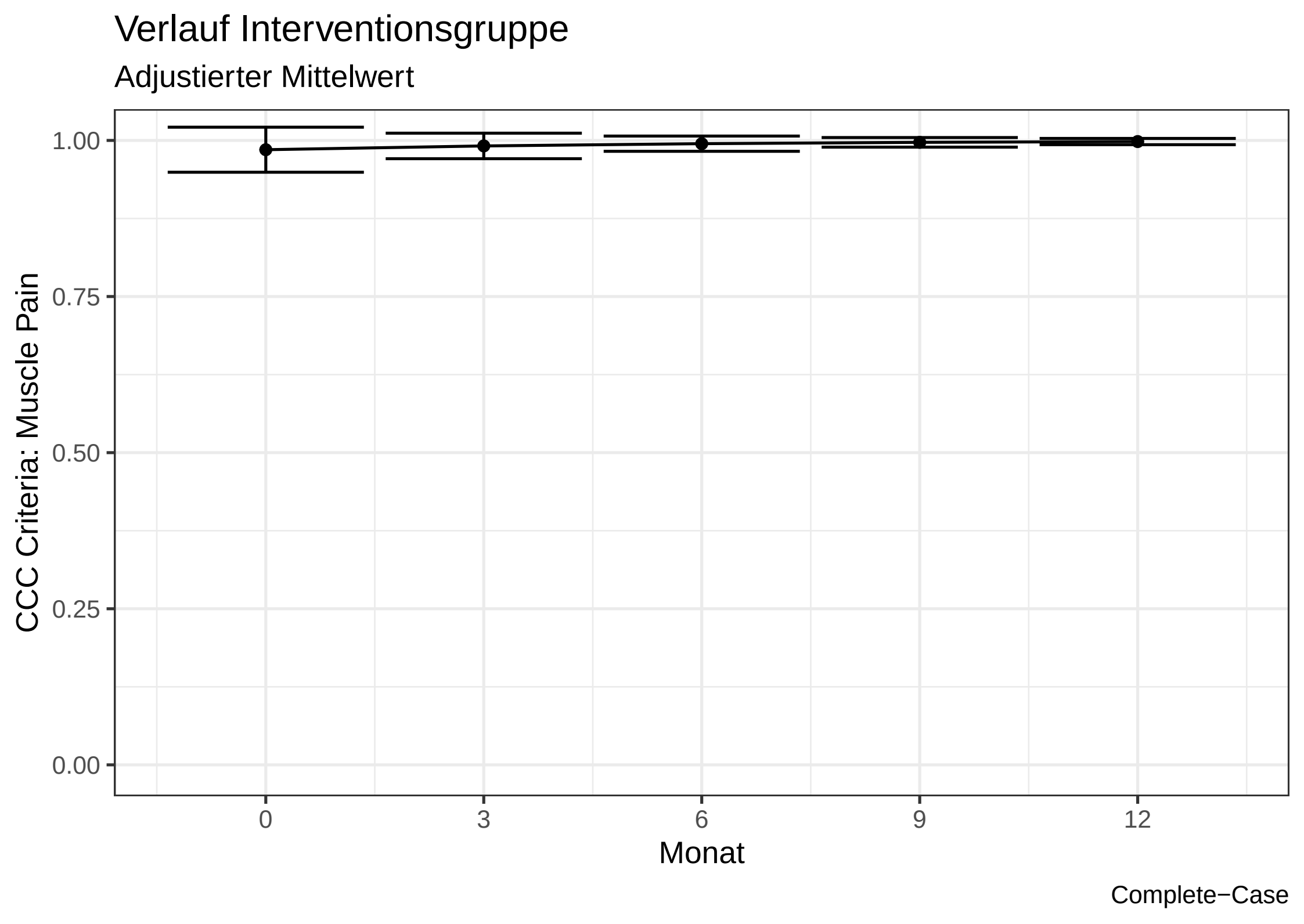
Adjustierter Mittelwert

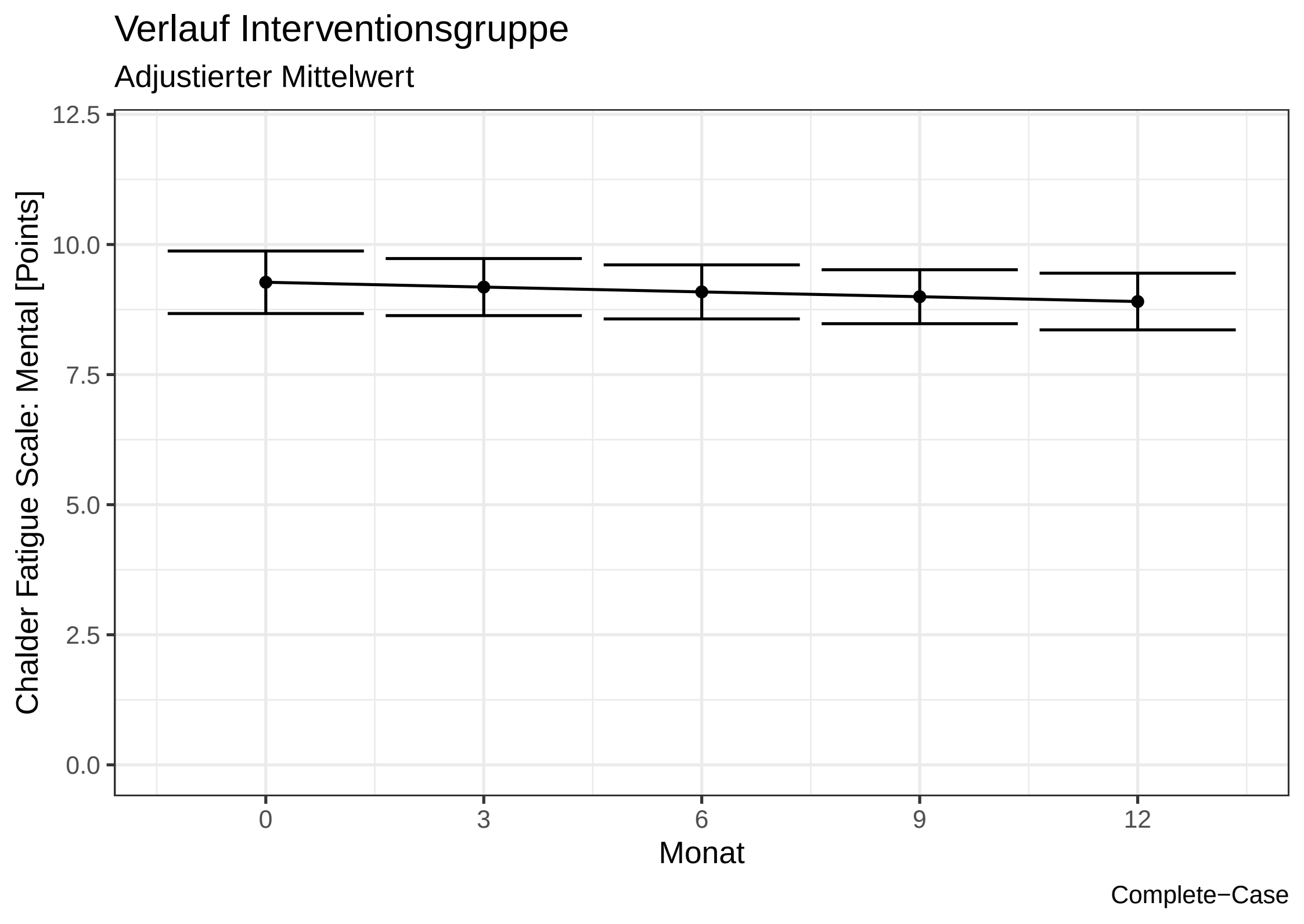
Anlage 1: Graphische Darstellung der marginalen Mittelwerte (mit 95-%-Konfidenzintervall) der ausgewählten Endpunkte der Interventionsgruppe im zeitlichen Verlauf.

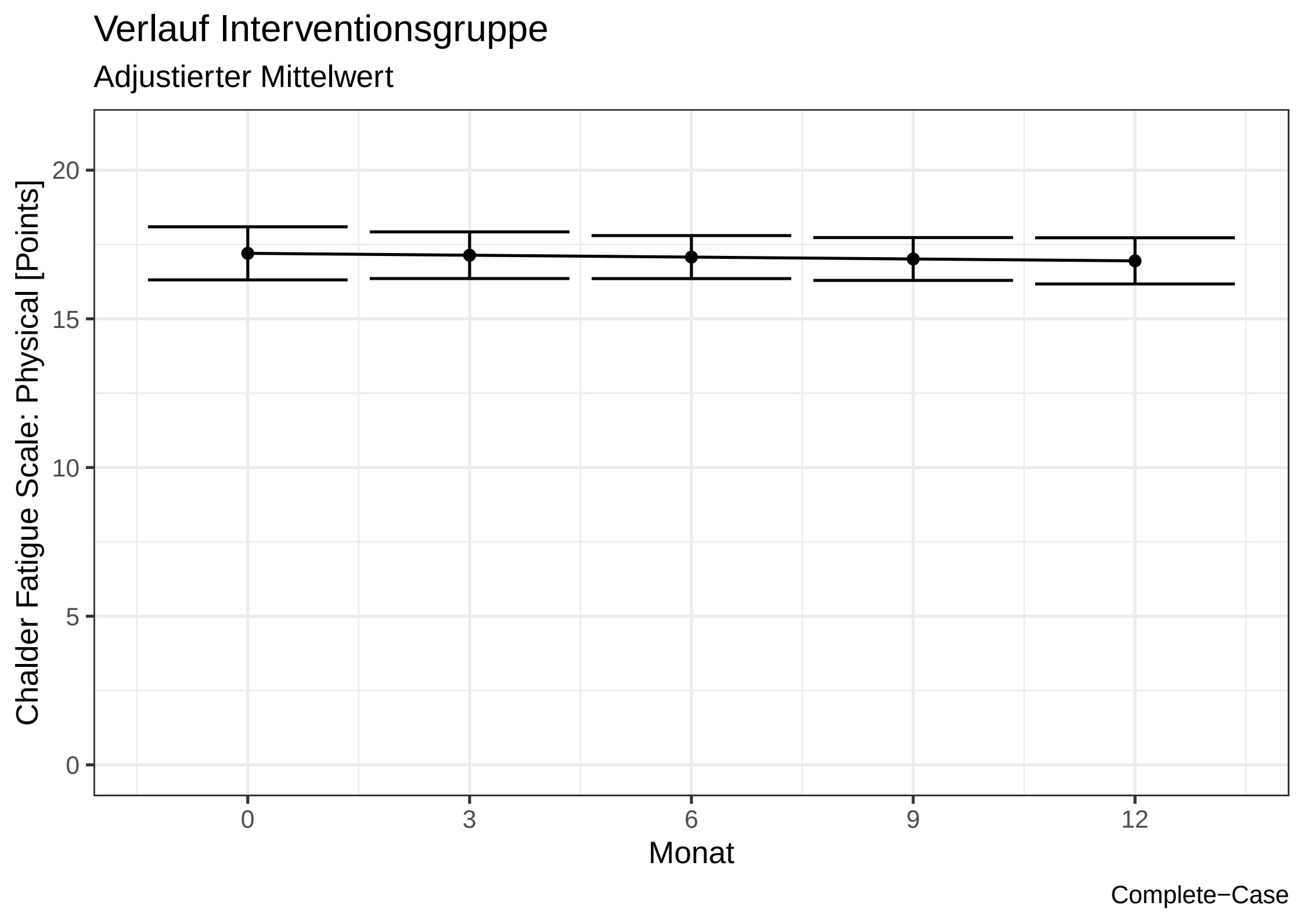


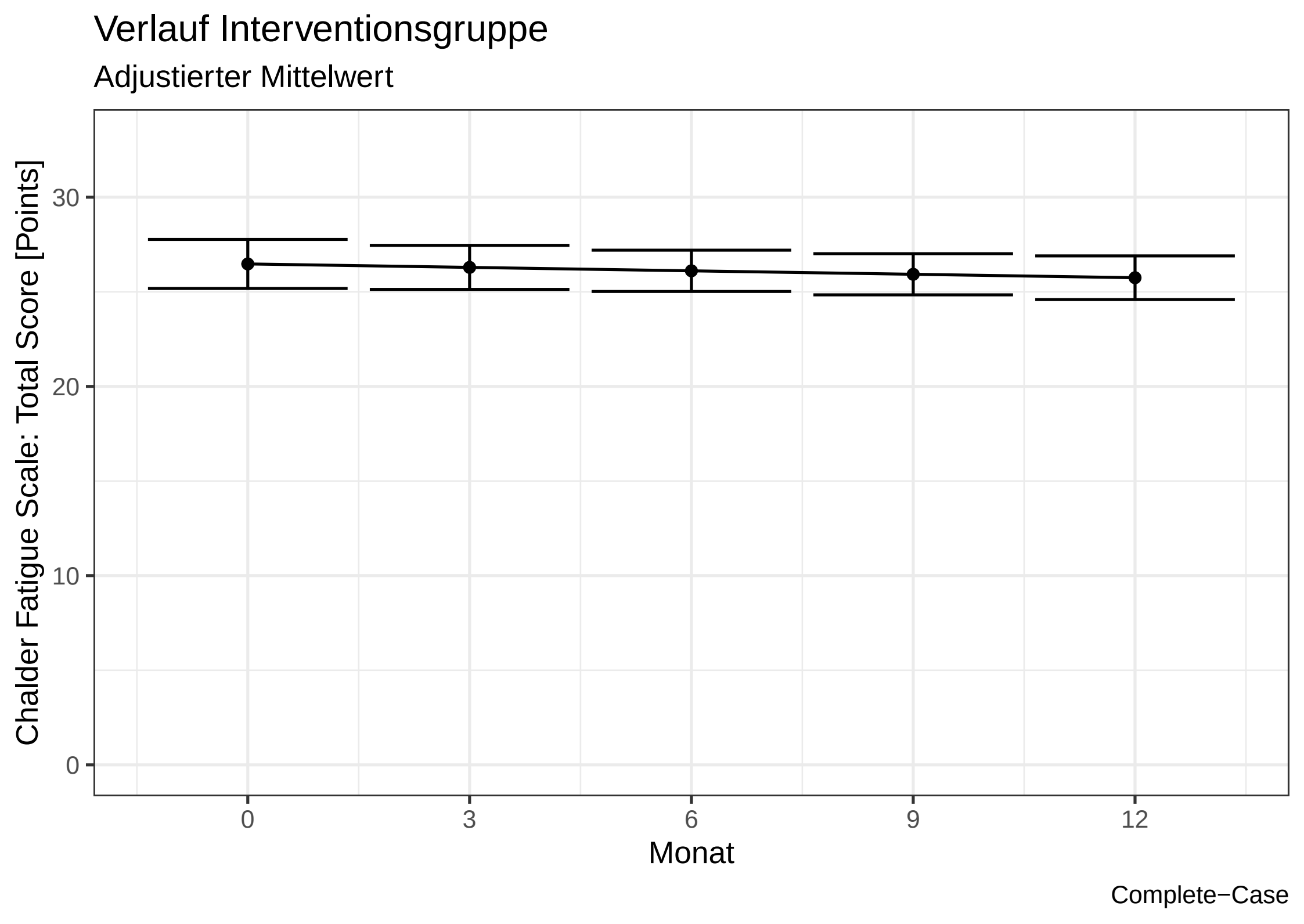
Complete−Case

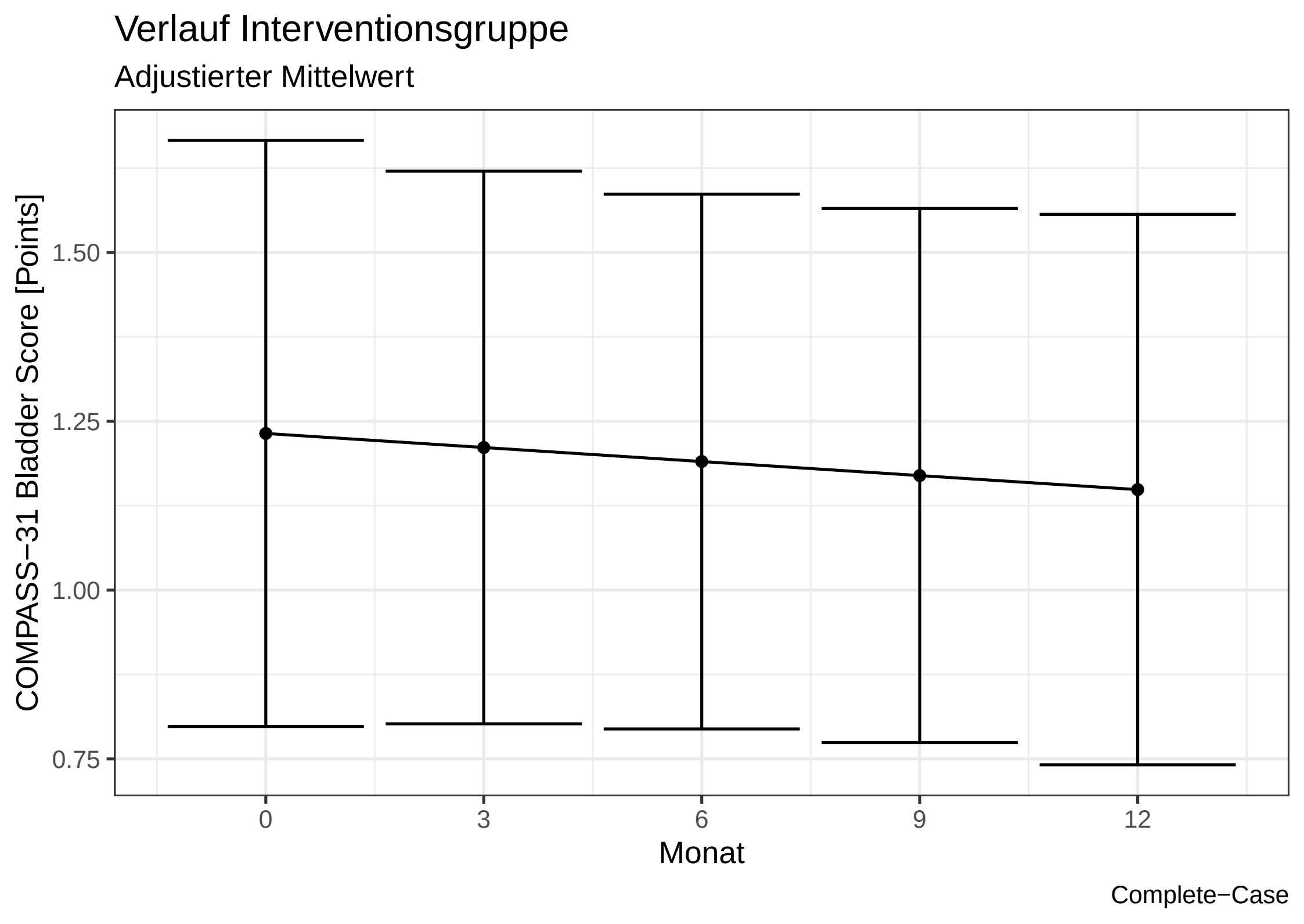

Verlauf Interventionsgruppe
Adjustierter Mittelwert
CCC Criteria: Headache
1.00
0.75
0.50
0.25
0.00
0
3
6
9
12
Monat
Complete−Case

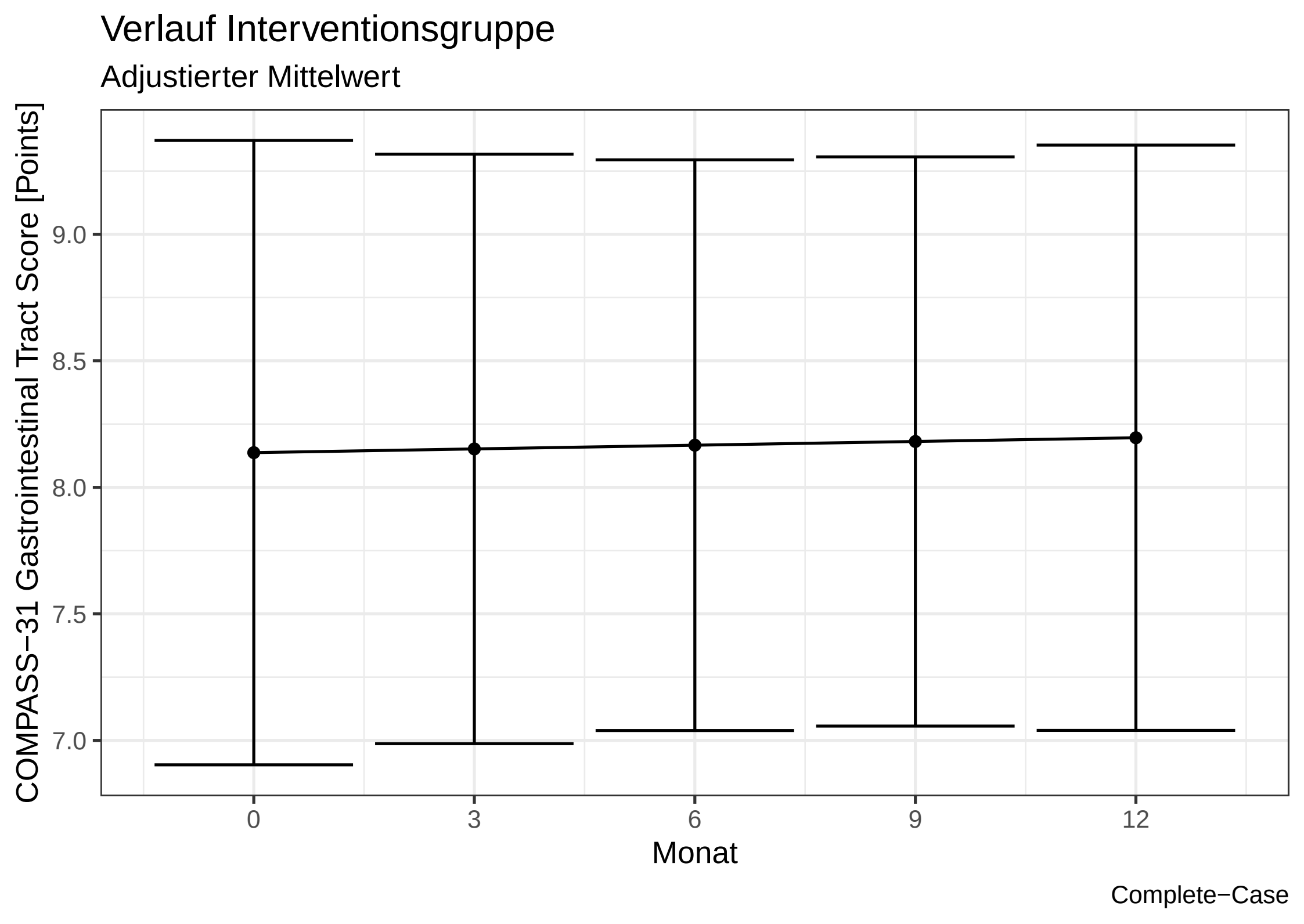

Verlauf Interventionsgruppe
Adjustierter Mittelwert
CCC Criteria: Muscle Pain
1.00
0.75
0.50
0.25
0.00
0
3
6
9
12
Monat
Complete−Case

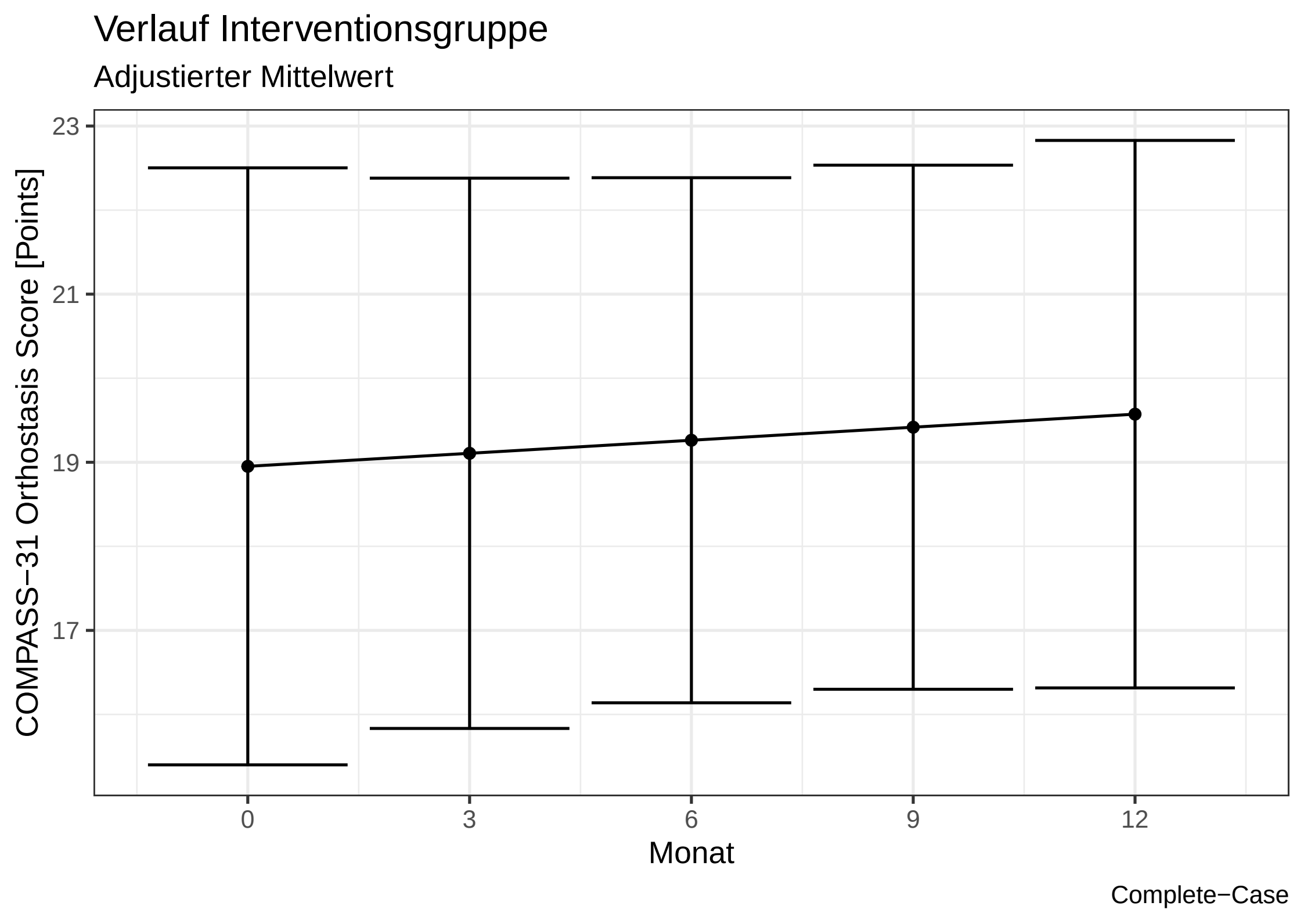

Verlauf Interventionsgruppe
Adjustierter Mittelwert
Chalder Fatigue Scale: Mental [Points]
12.5
10.0
7.5
5.0
2.5
0.0
0
3
6
9
12
Monat
Complete−Case

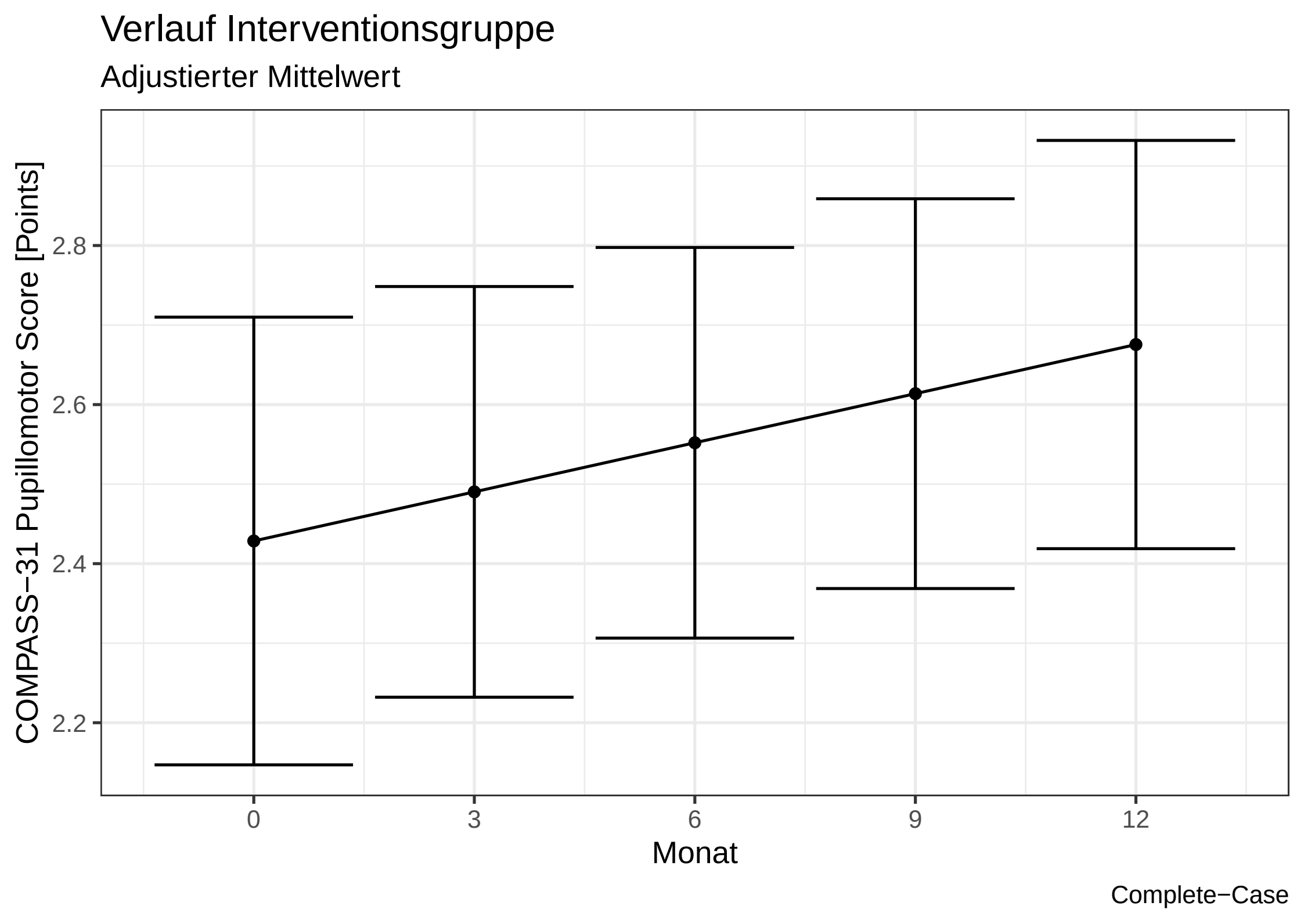

Verlauf Interventionsgruppe
Adjustierter Mittelwert
Chalder Fatigue Scale: Physical [Points]
20
15
10
5
0
0
3
6
9
12
Monat
Complete−Case

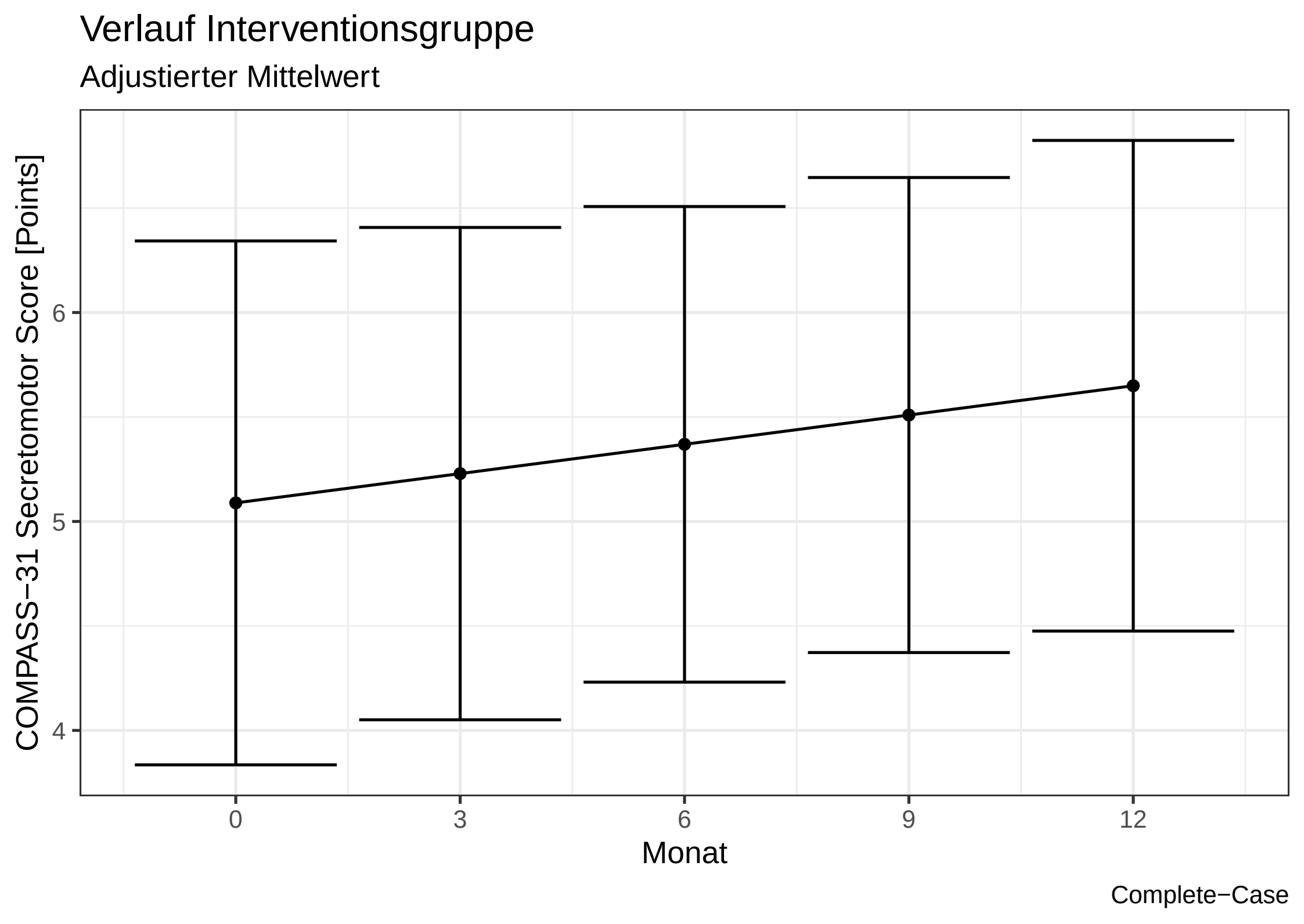

Verlauf Interventionsgruppe
Adjustierter Mittelwert
Chalder Fatigue Scale: Total Score [Points]
30
20
10
0
0
3
6
9
12
Monat
Complete−Case

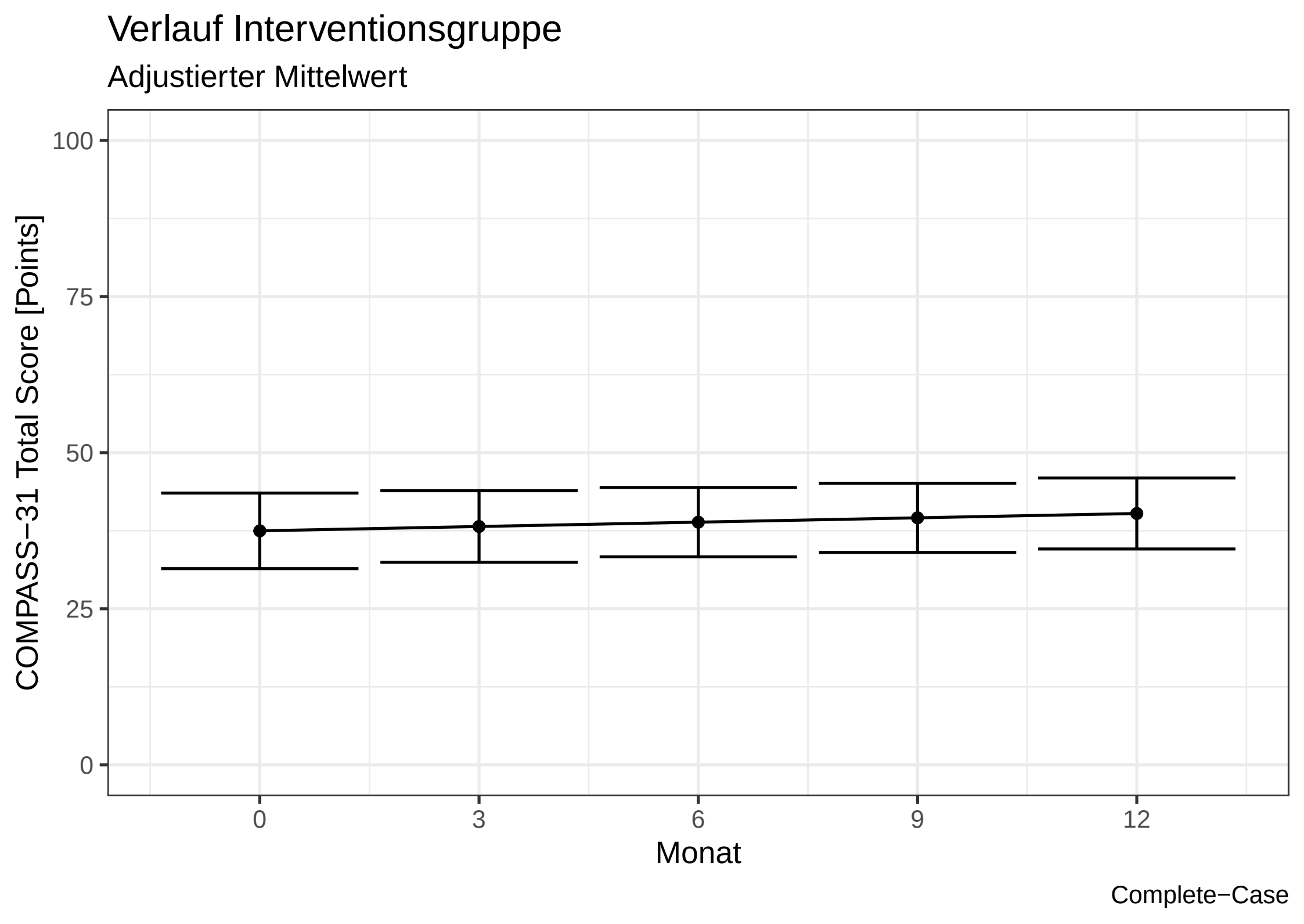

Verlauf Interventionsgruppe
Adjustierter Mittelwert
COMPASS−31 Bladder Score [Points]
1.50
1.25
1.00
0.75
0
3
6
9
12
Monat
Complete−Case

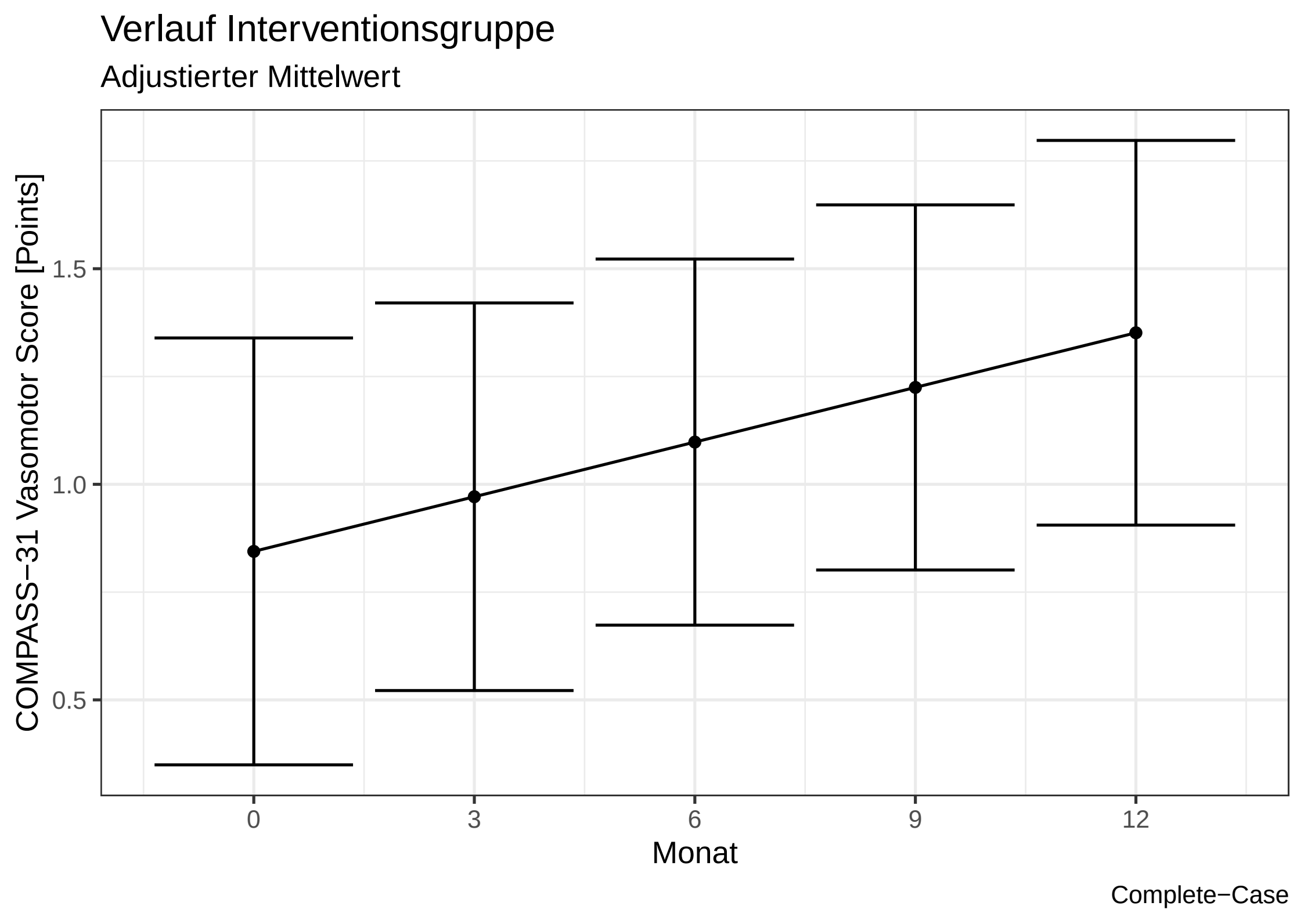

Verlauf Interventionsgruppe
Adjustierter Mittelwert
COMPASS−31 Gastrointestinal Tract Score [Points]
9.0
8.5
8.0
7.5
7.0
0
3
6
9
12
Monat
Complete−Case

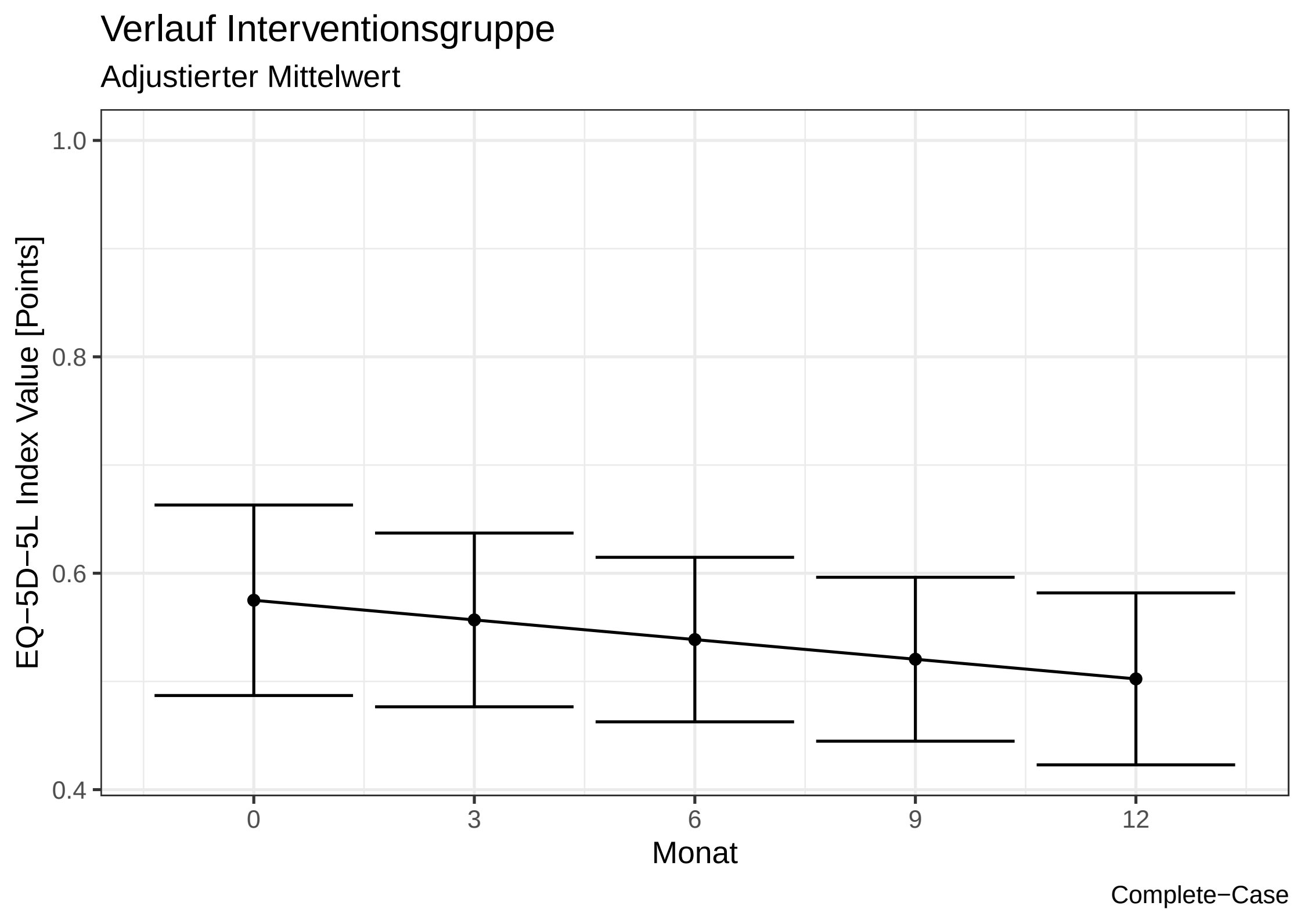

Verlauf Interventionsgruppe
Adjustierter Mittelwert
COMPASS−31 Orthostasis Score [Points]
23
21
19
17
0
3
6
9
12
Monat
Complete−Case

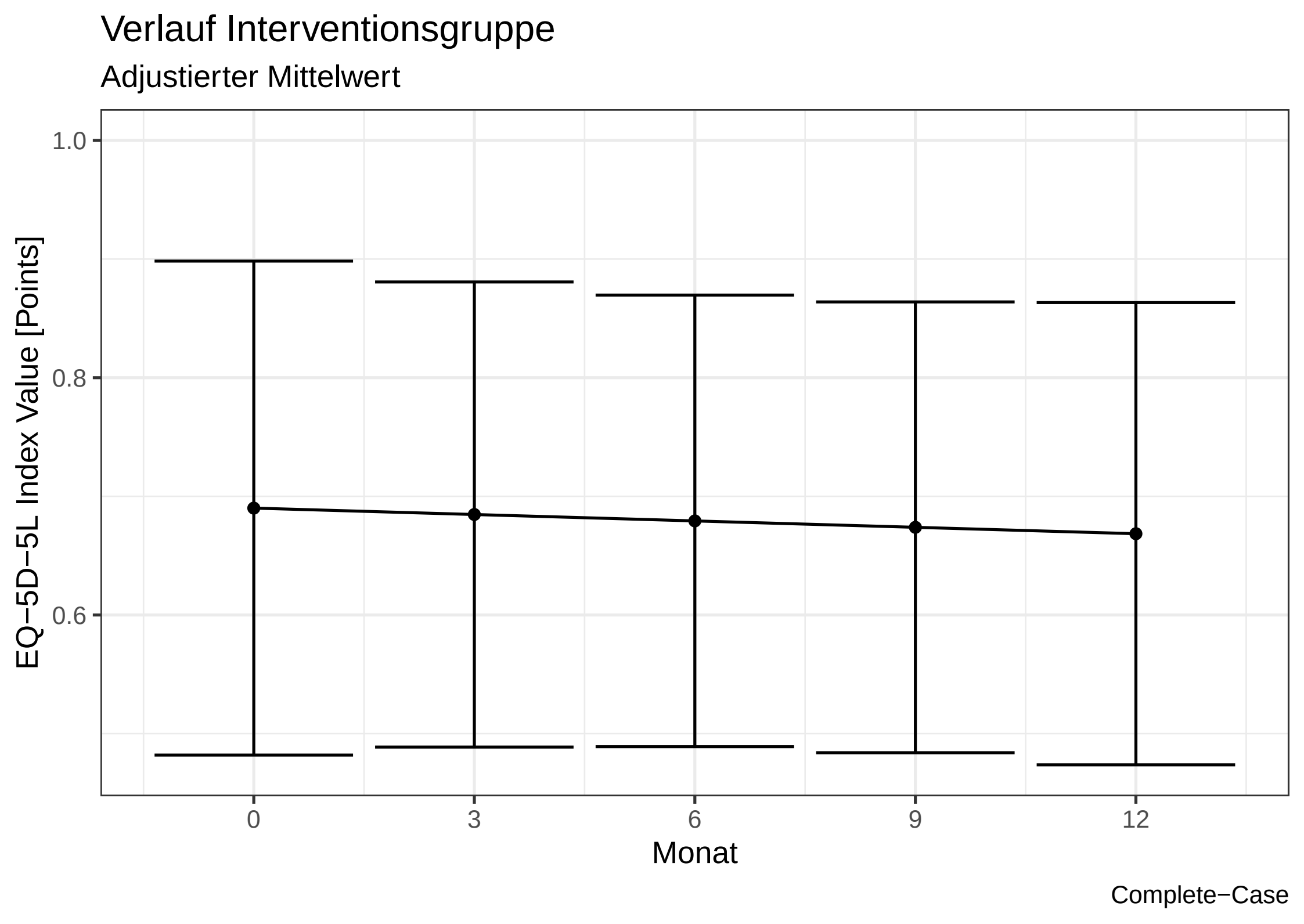

Verlauf Interventionsgruppe
Adjustierter Mittelwert
COMPASS−31 Pupillomotor Score [Points]
2.8
2.6
2.4
2.2
0
3
6
9
12
Monat
Complete−Case

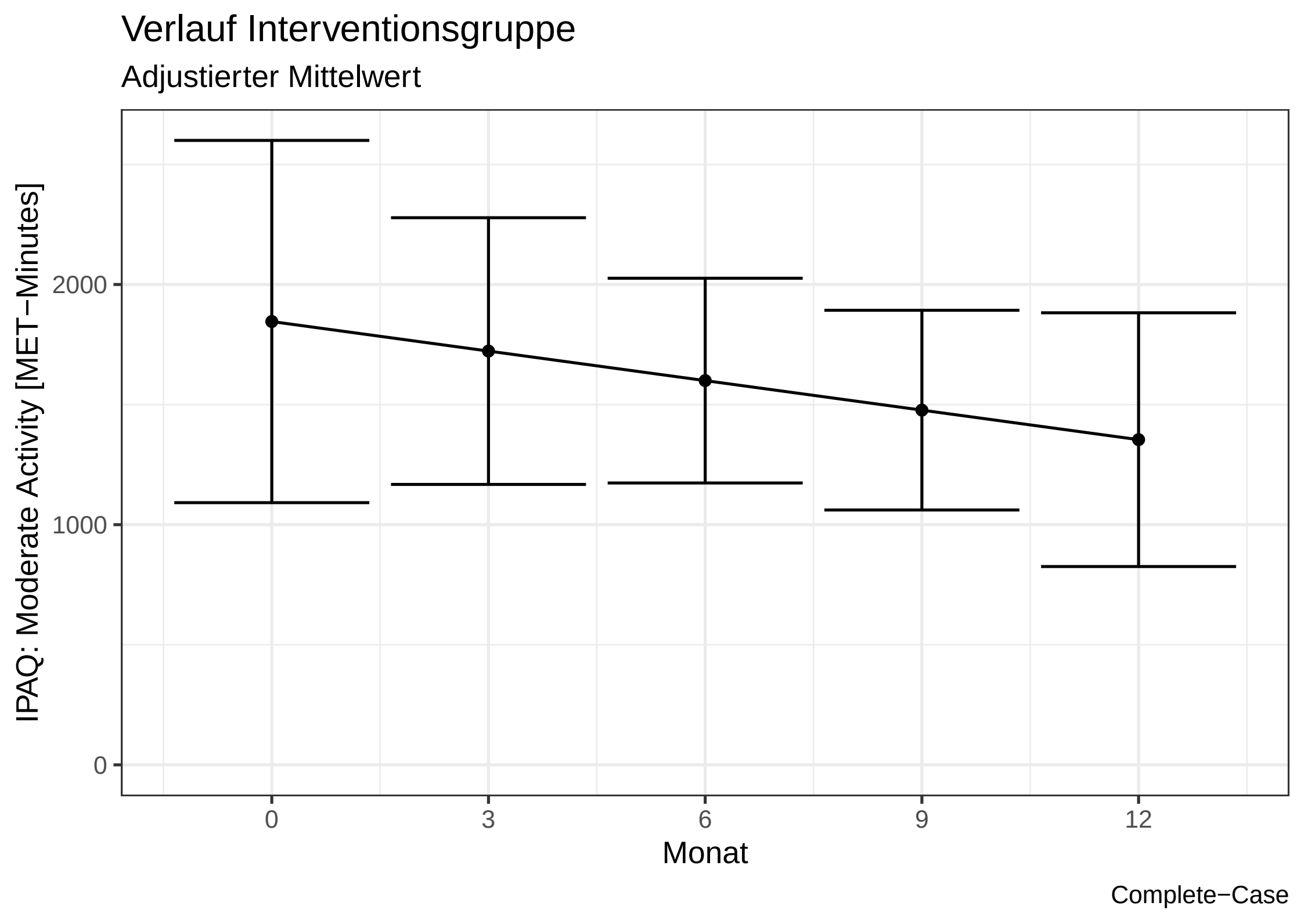

Verlauf Interventionsgruppe
Adjustierter Mittelwert
COMPASS−31 Secretomotor Score [Points]
6
5
4
0
3
6
9
12
Monat
Complete−Case

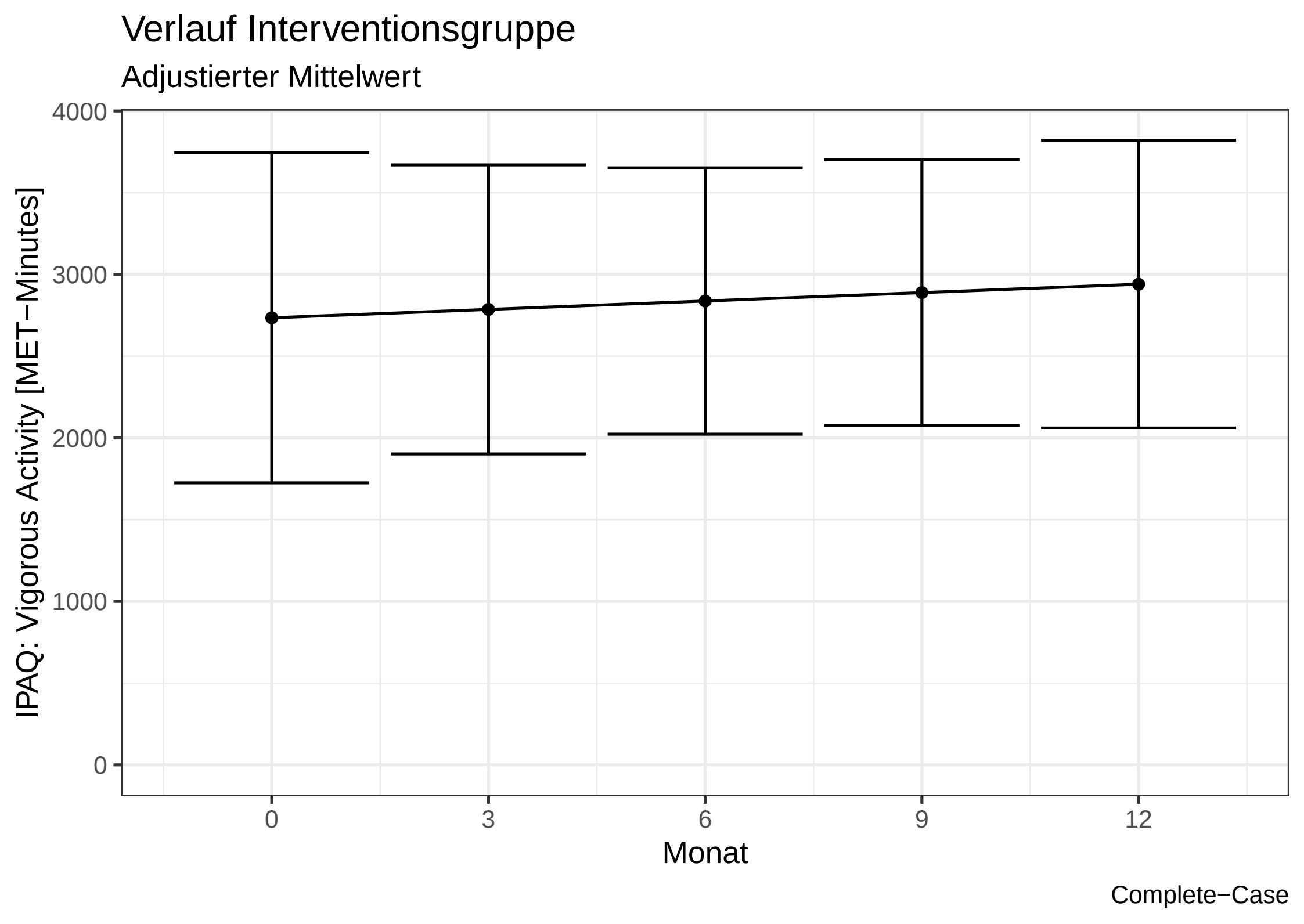

Verlauf Interventionsgruppe
Adjustierter Mittelwert
COMPASS−31 Total Score [Points]
100
75
50
25
0
0
3
6
9
12
Monat
Complete−Case

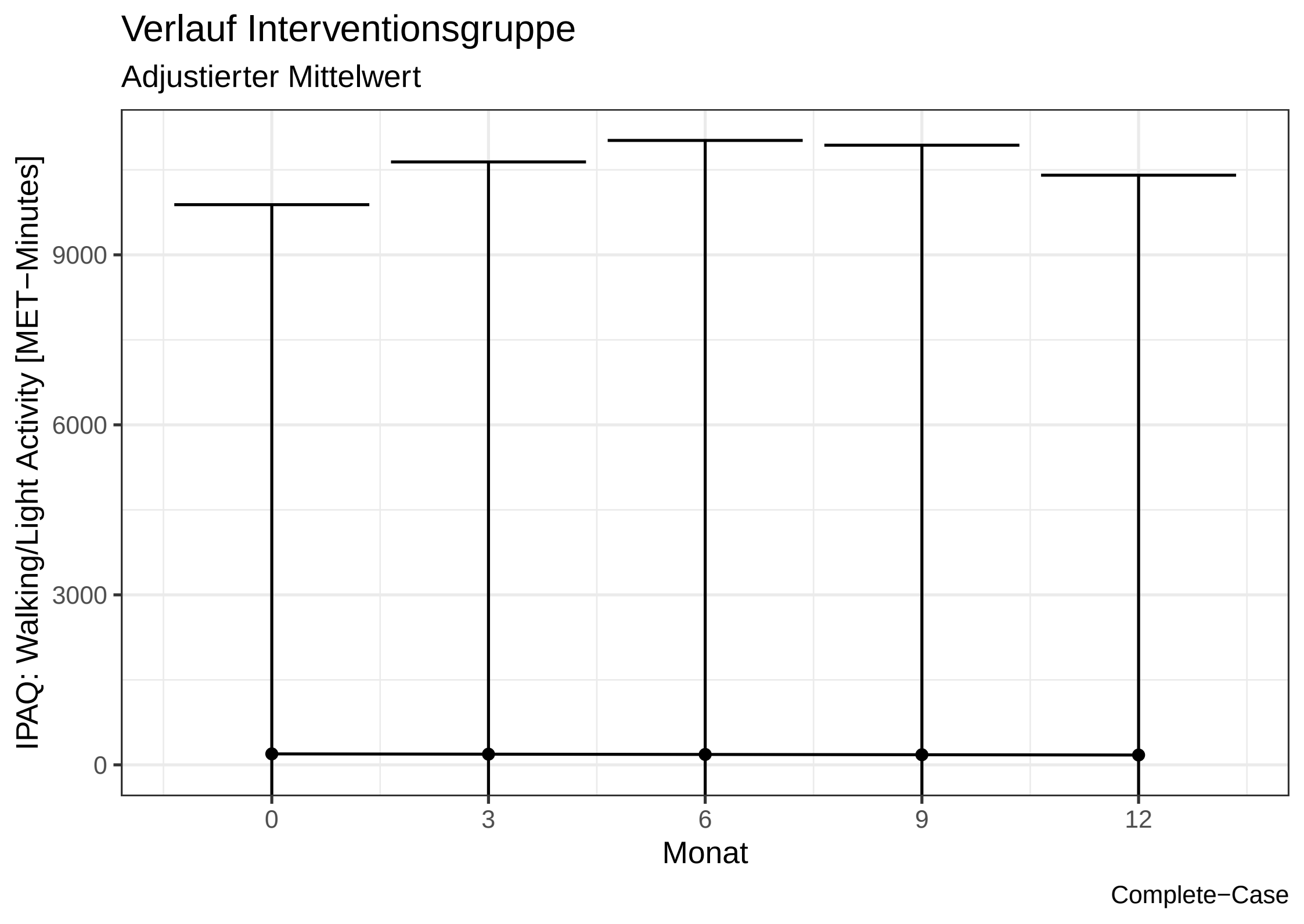

Verlauf Interventionsgruppe
Adjustierter Mittelwert
COMPASS−31 Vasomotor Score [Points]
1.5
1.0
0.5
0
3
6
9
12
Monat
Complete−Case

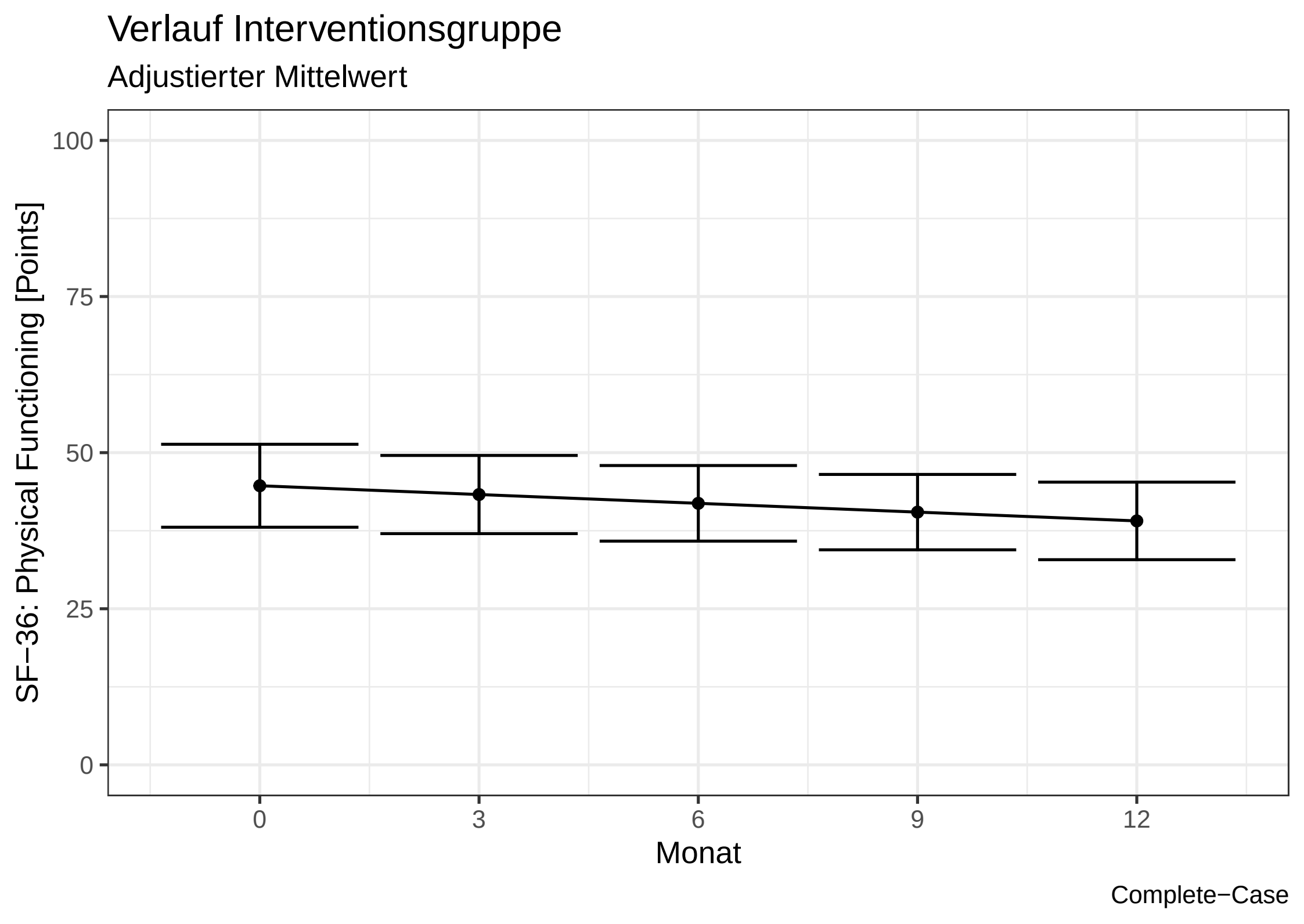

Verlauf Interventionsgruppe
Adjustierter Mittelwert
EQ−5D−5L Index Value [Points]
1.0
0.8
0.6
0.4
0
3
6
9
12
Monat
Complete−Case


Verlauf Interventionsgruppe
Adjustierter Mittelwert
1.0
0.8
0.6
EQ−5D−5L Index Value [Points]
0
3
6
9
12
Monat
Complete−Case


Verlauf Interventionsgruppe
Adjustierter Mittelwert
IPAQ: Moderate Activity [MET−Minutes]
2000
1000
0
0
3
6
9
12
Monat
Complete−Case

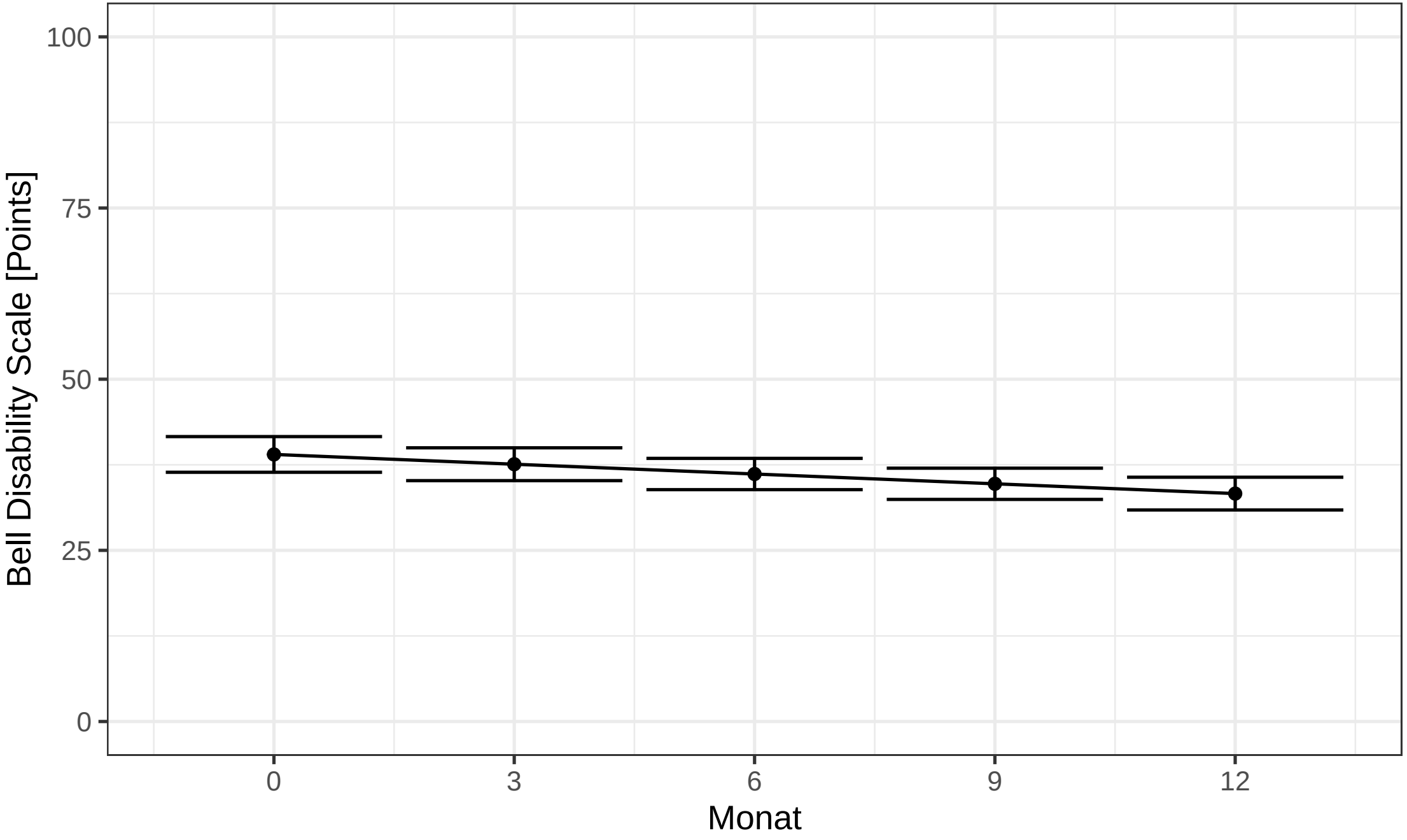

Verlauf Interventionsgruppe
Adjustierter Mittelwert
IPAQ: Vigorous Activity [MET−Minutes]
4000
3000
2000
1000
0
0
3
6
9
12
Monat
Complete−Case


Verlauf Interventionsgruppe
Adjustierter Mittelwert
IPAQ: Walking/Light Activity [MET−Minutes]
9000
6000
3000
0
0
3
6
9
12
Monat
Complete−Case


Verlauf Interventionsgruppe
Adjustierter Mittelwert
SF−36: Physical Functioning [Points]
100
75
50
25
0
0
3
6
9
12
Monat
Complete−Case

Verlauf Interventionsgruppe

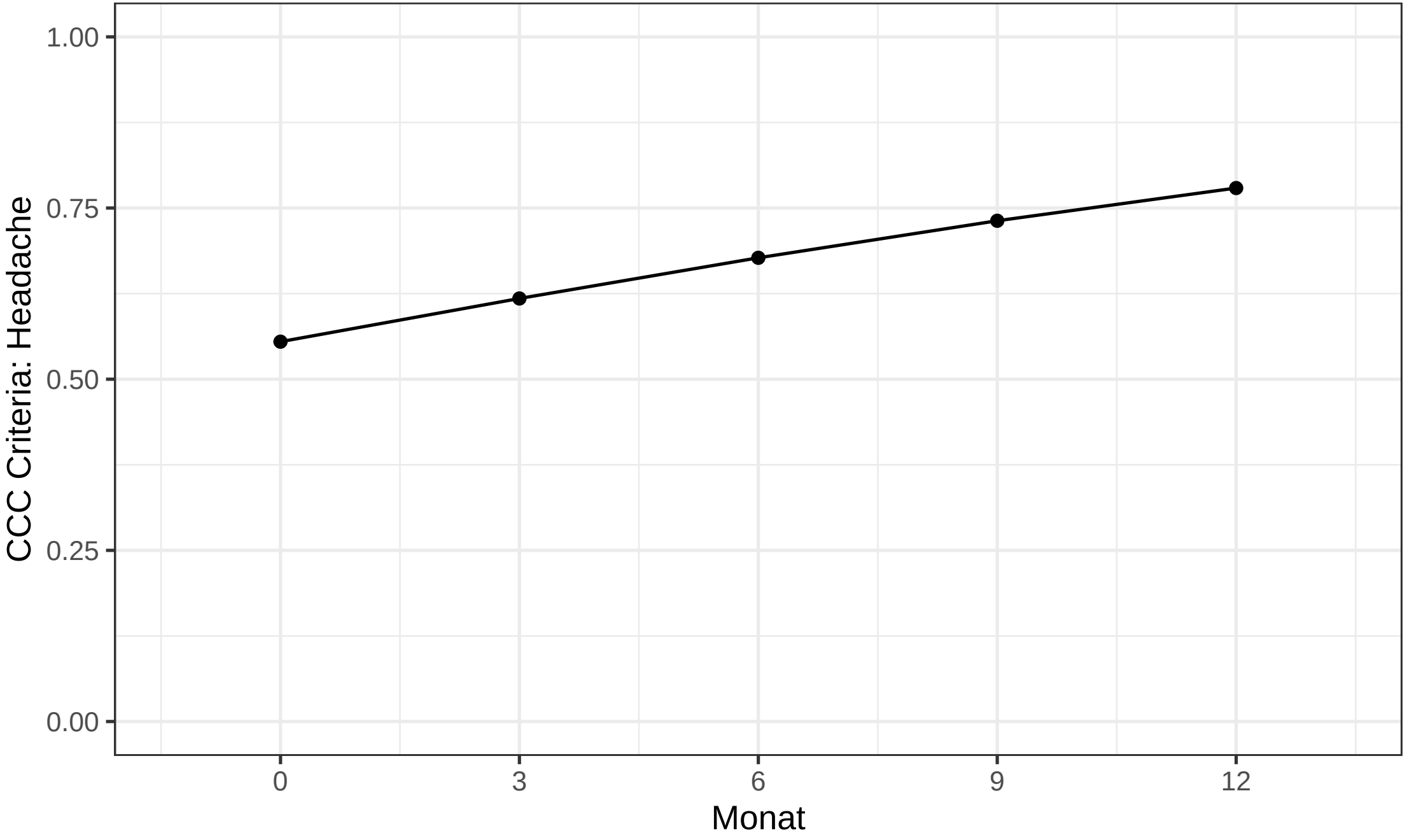
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

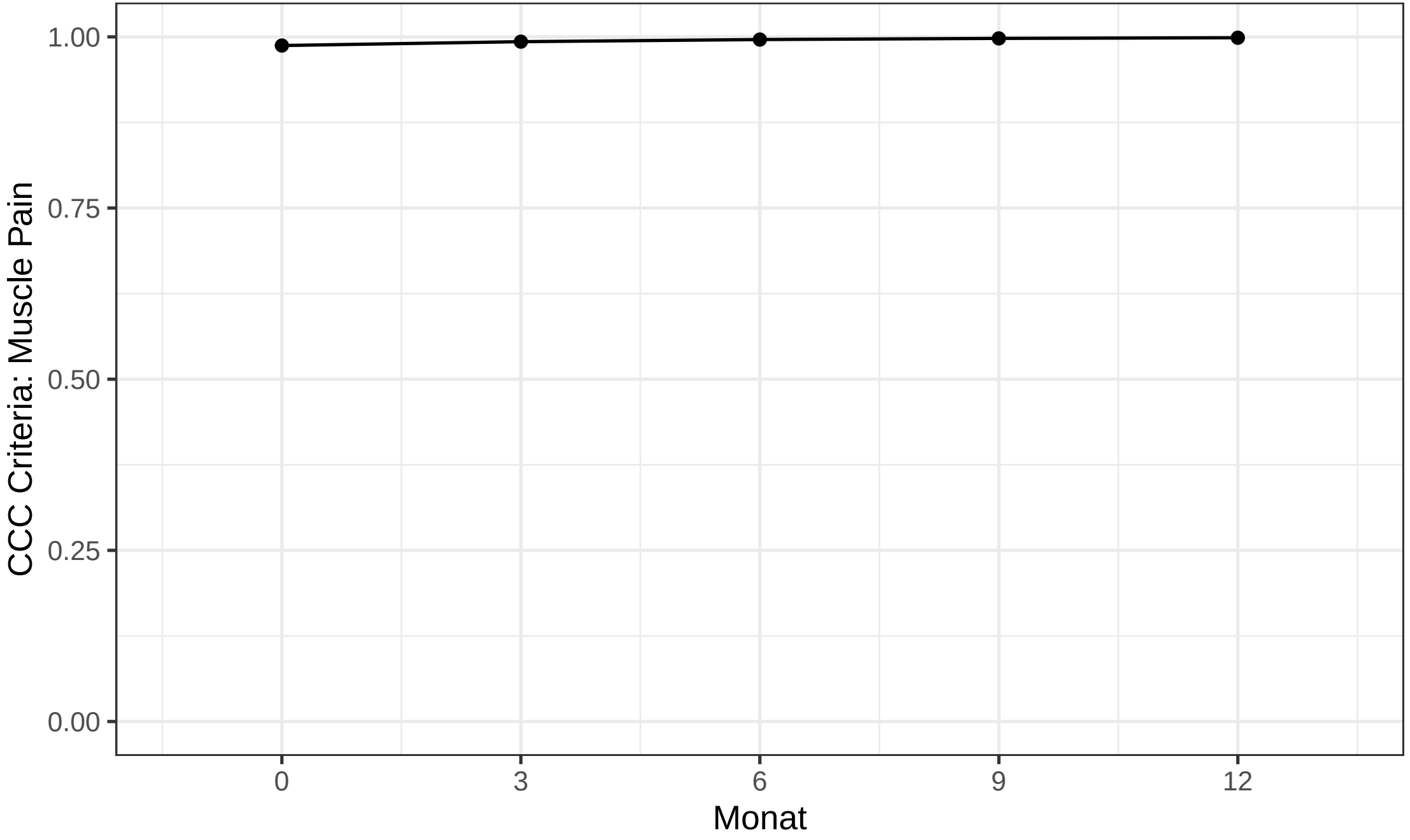
Adjustierter Mittelwert



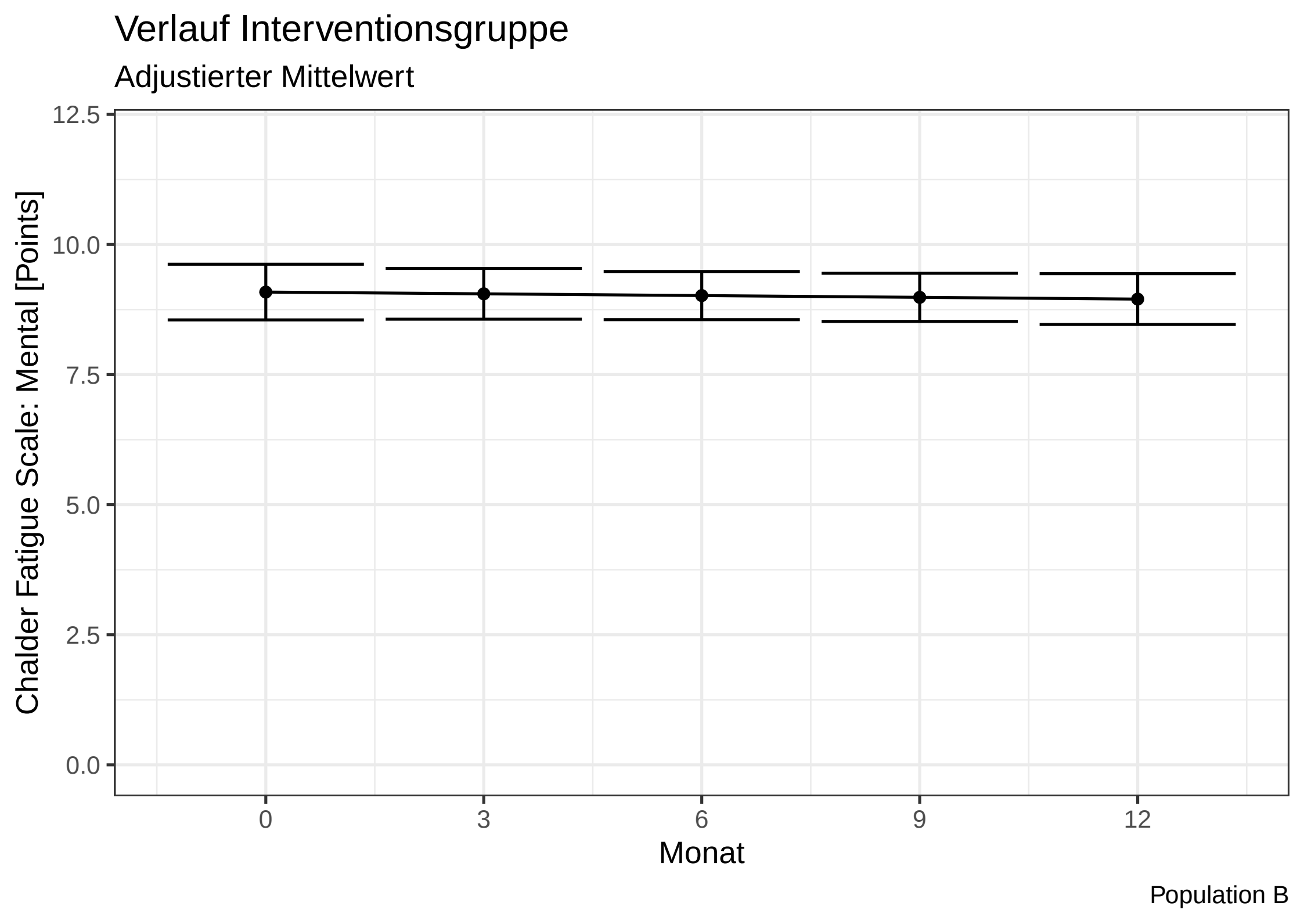
Population B

Verlauf Interventionsgruppe

Adjustierter Mittelwert

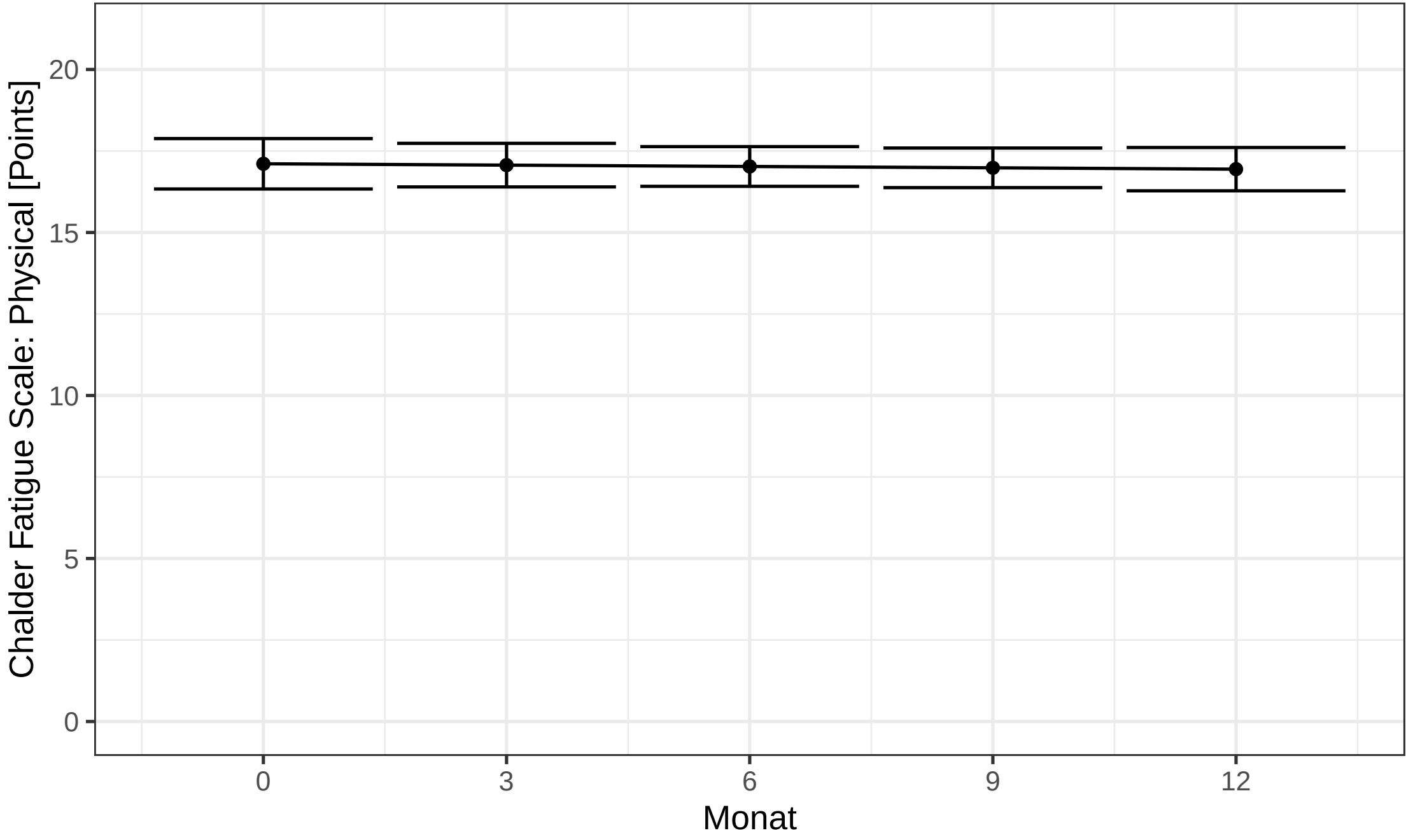


Population B


Verlauf Interventionsgruppe
Adjustierter Mittelwert
Chalder Fatigue Scale: Mental [Points]
12.5
10.0
7.5
5.0
2.5
0.0
0
3
6
9
12
Monat
Population B

Verlauf Interventionsgruppe

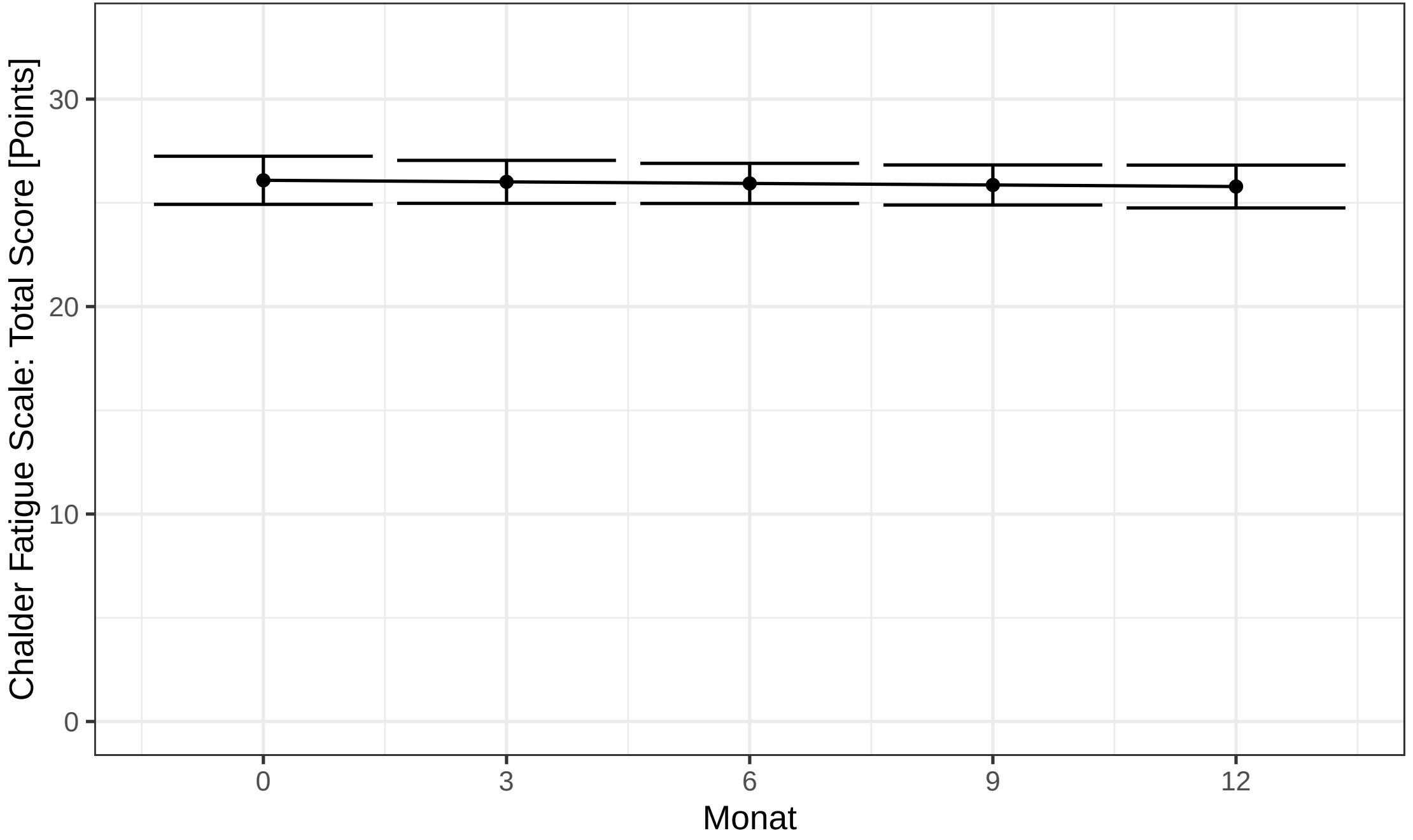
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

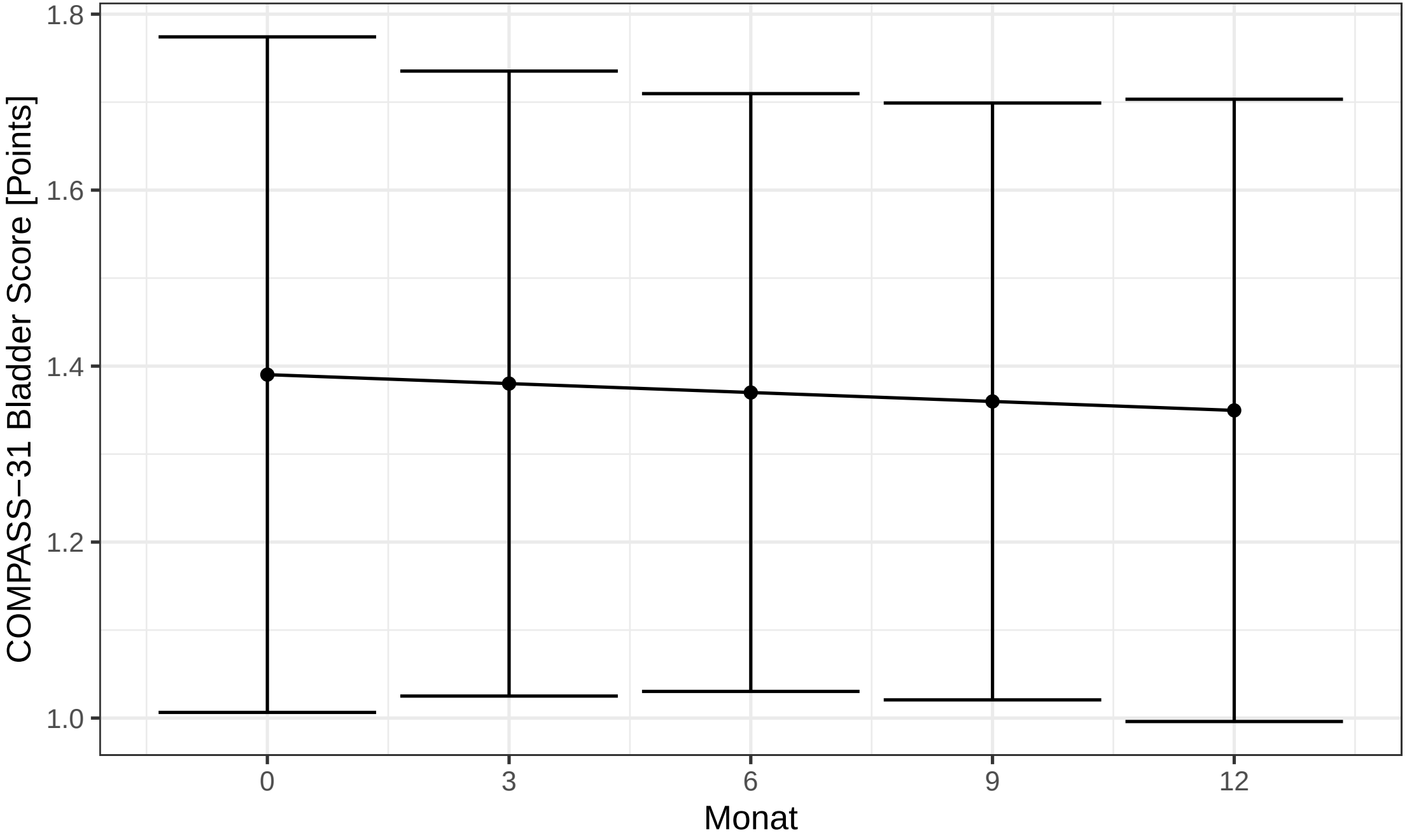
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

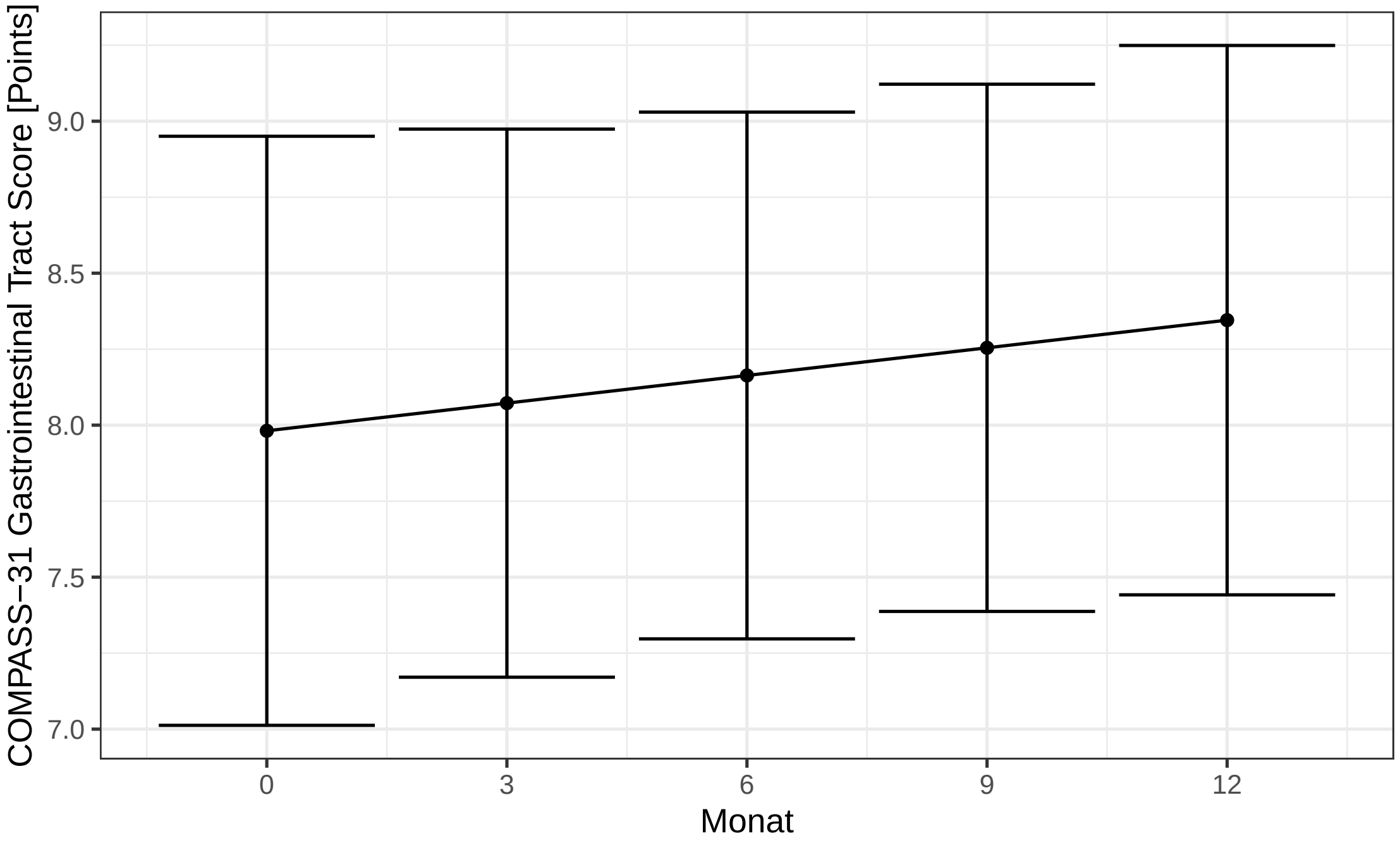
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

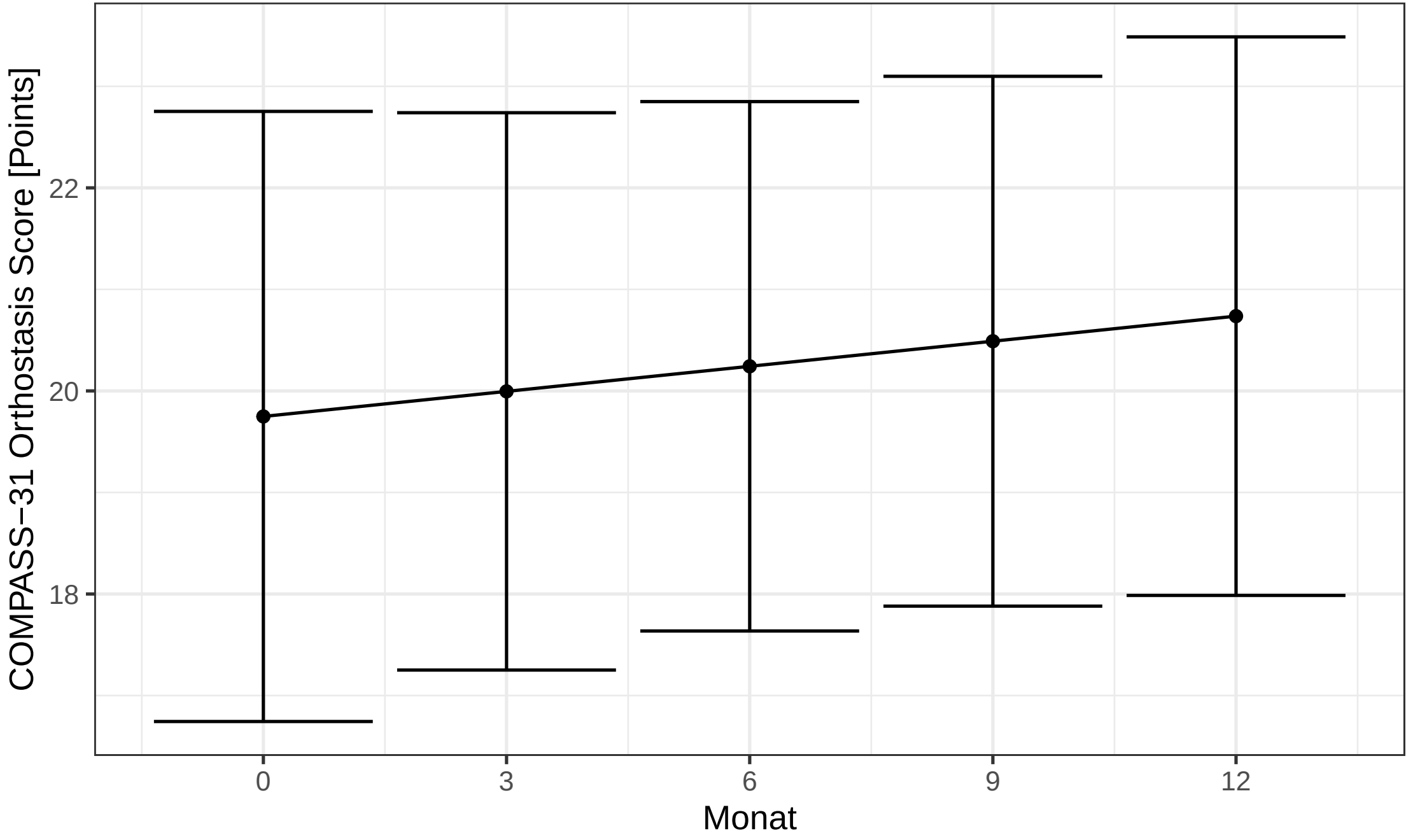
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

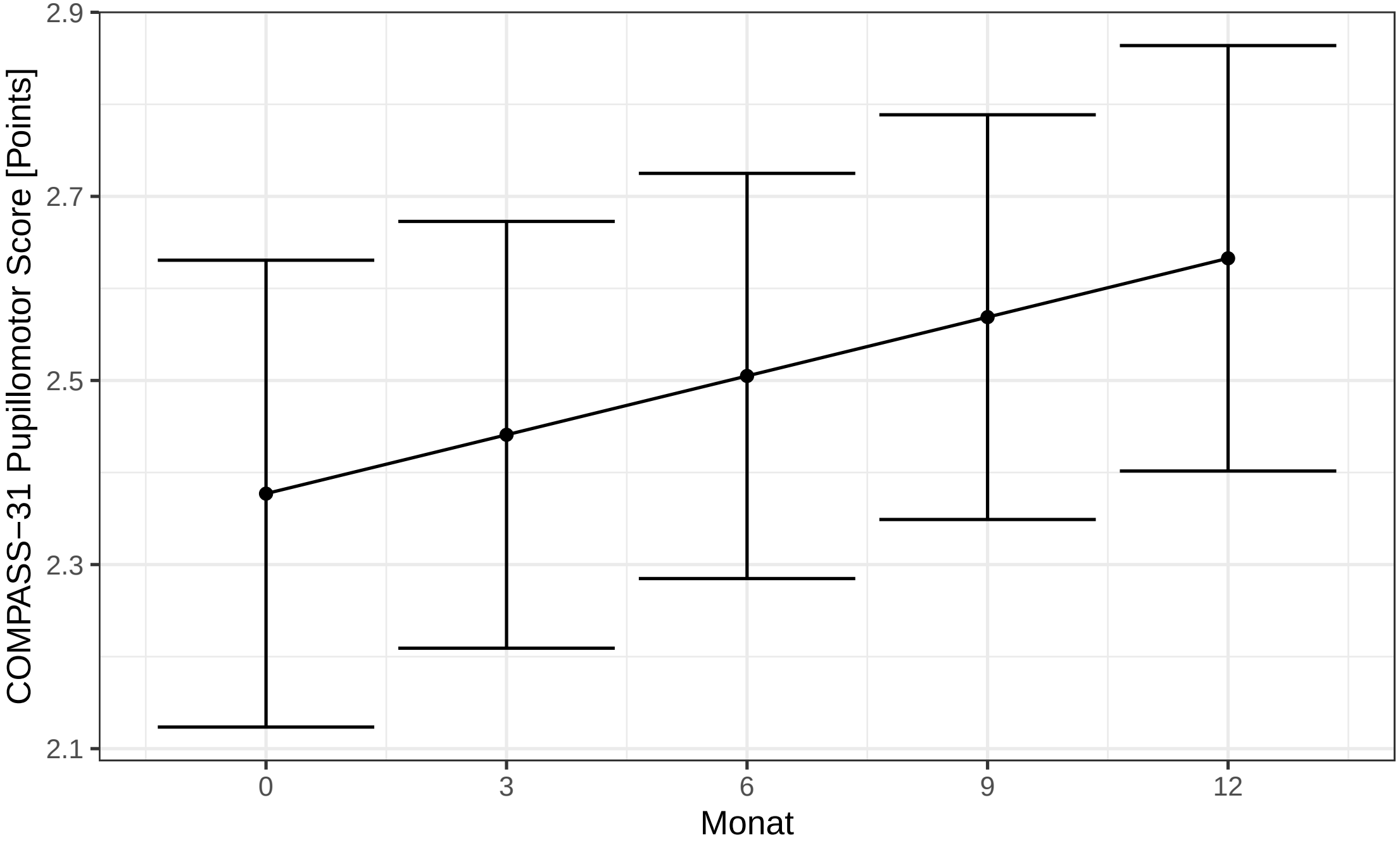
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

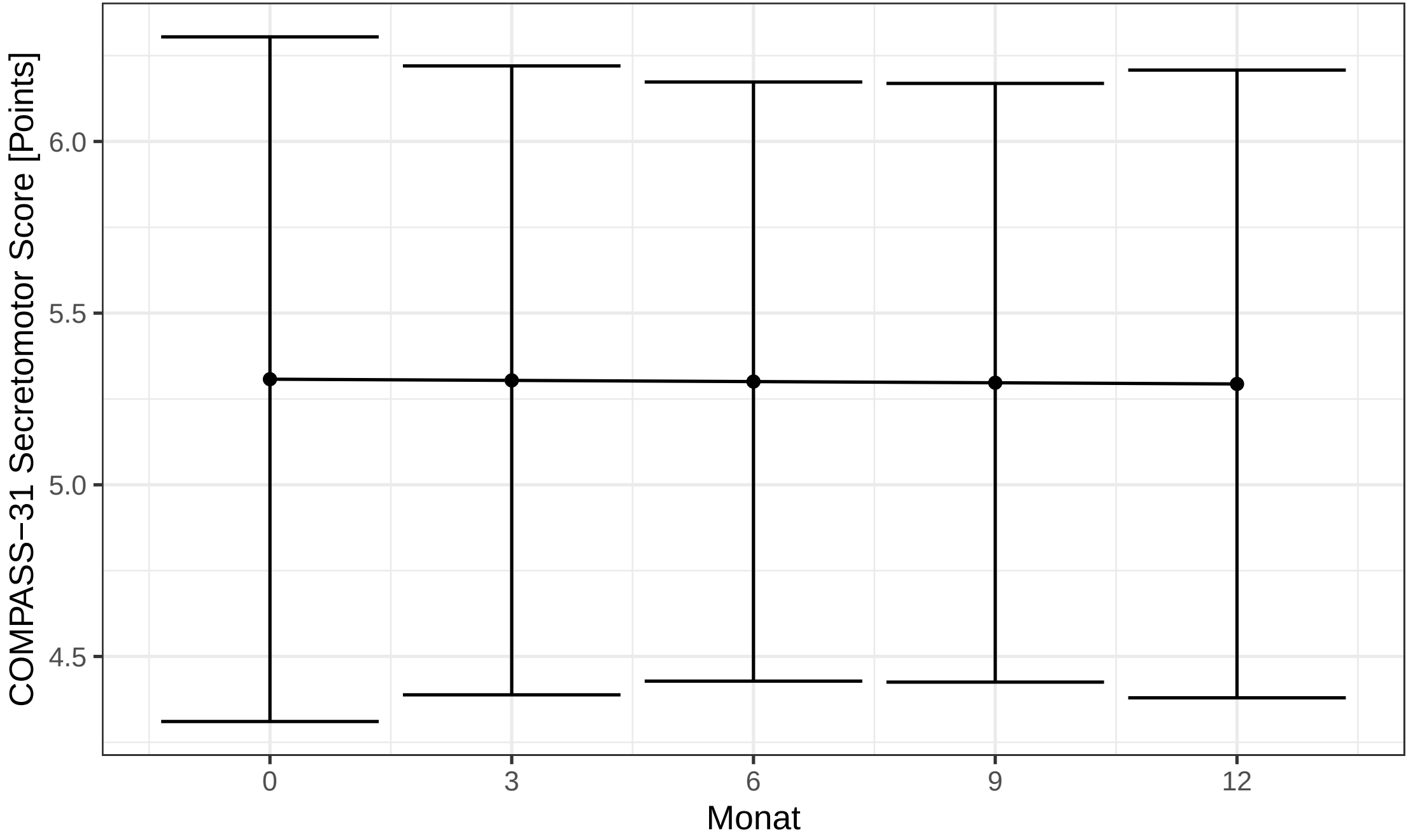
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

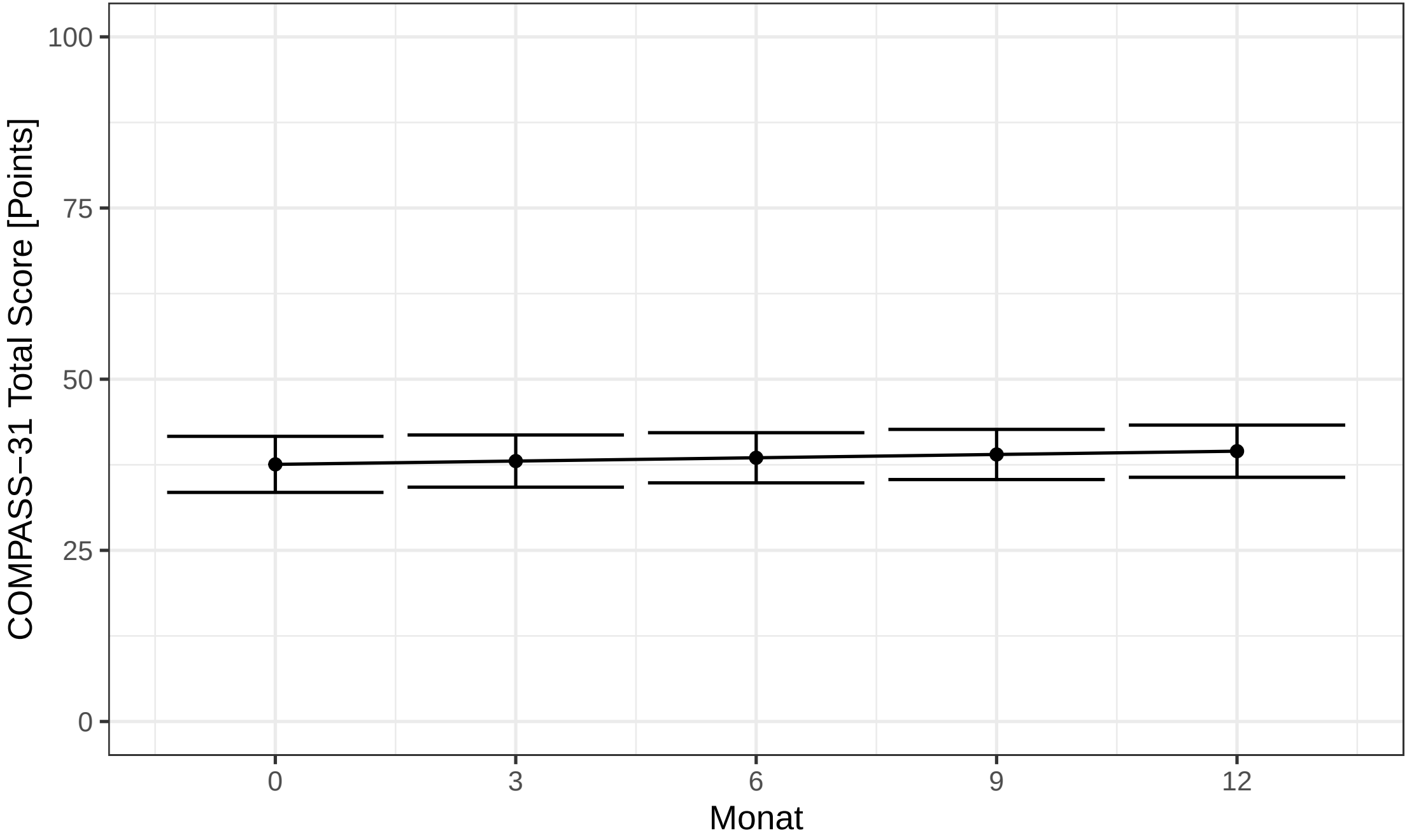
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

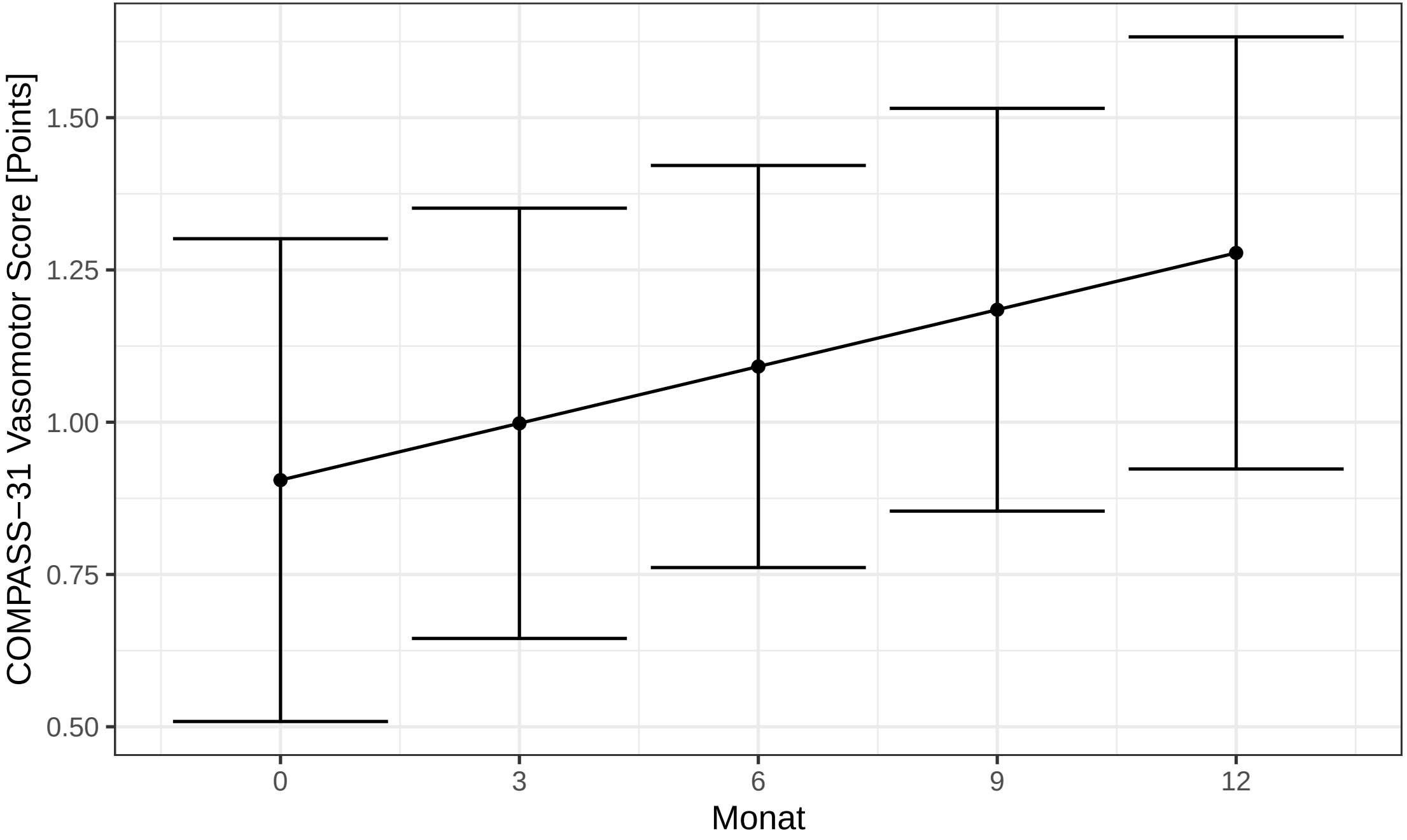
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

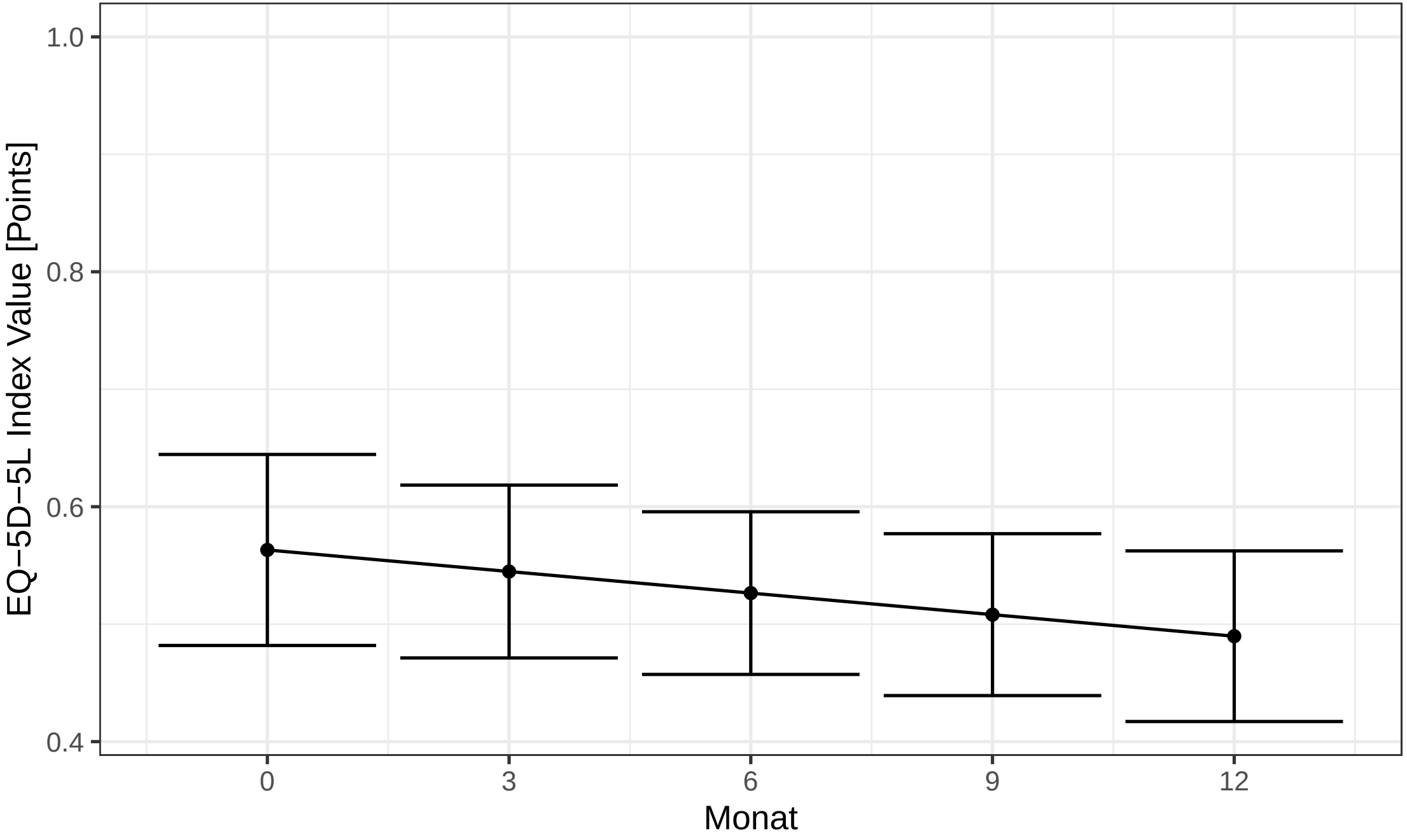
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

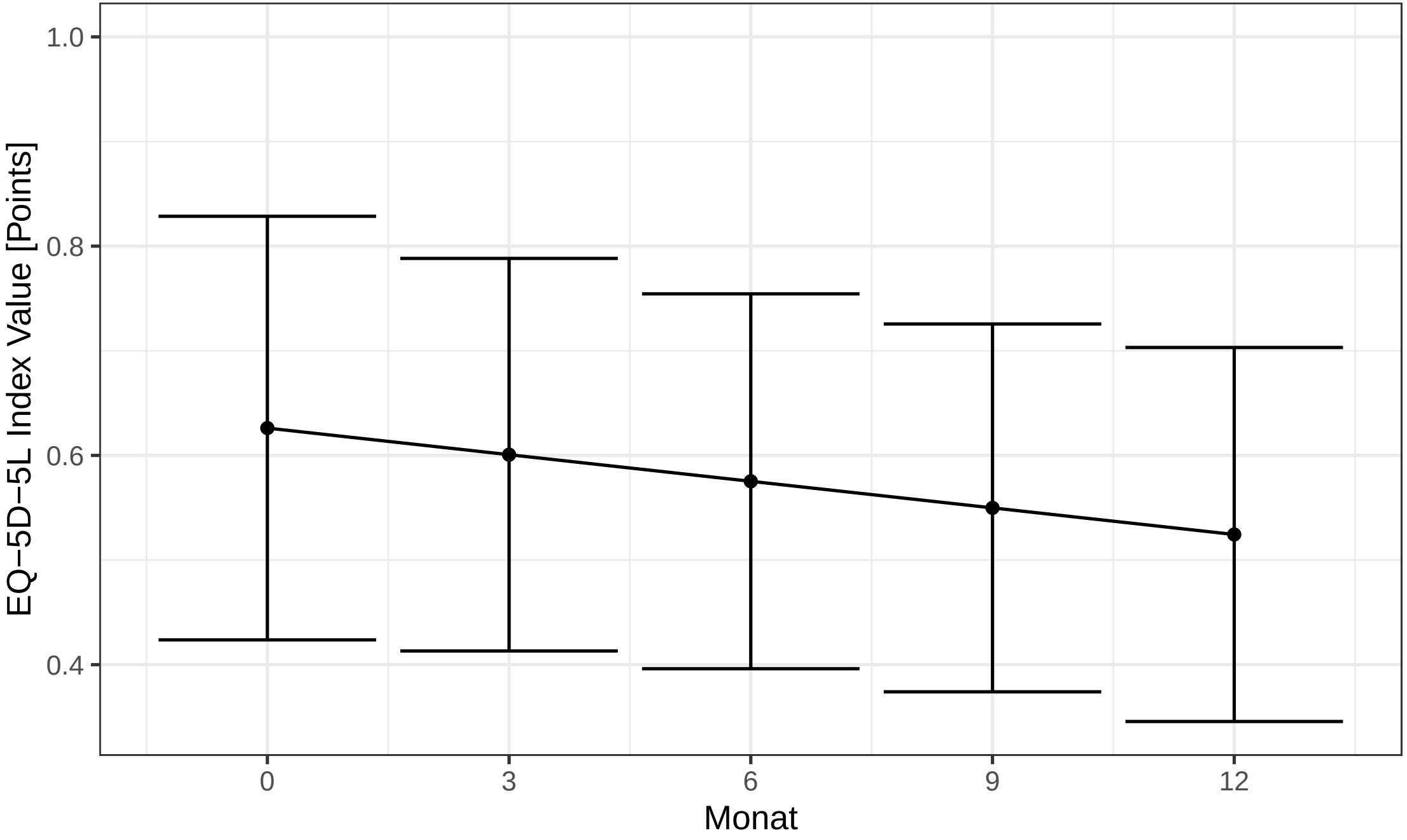
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

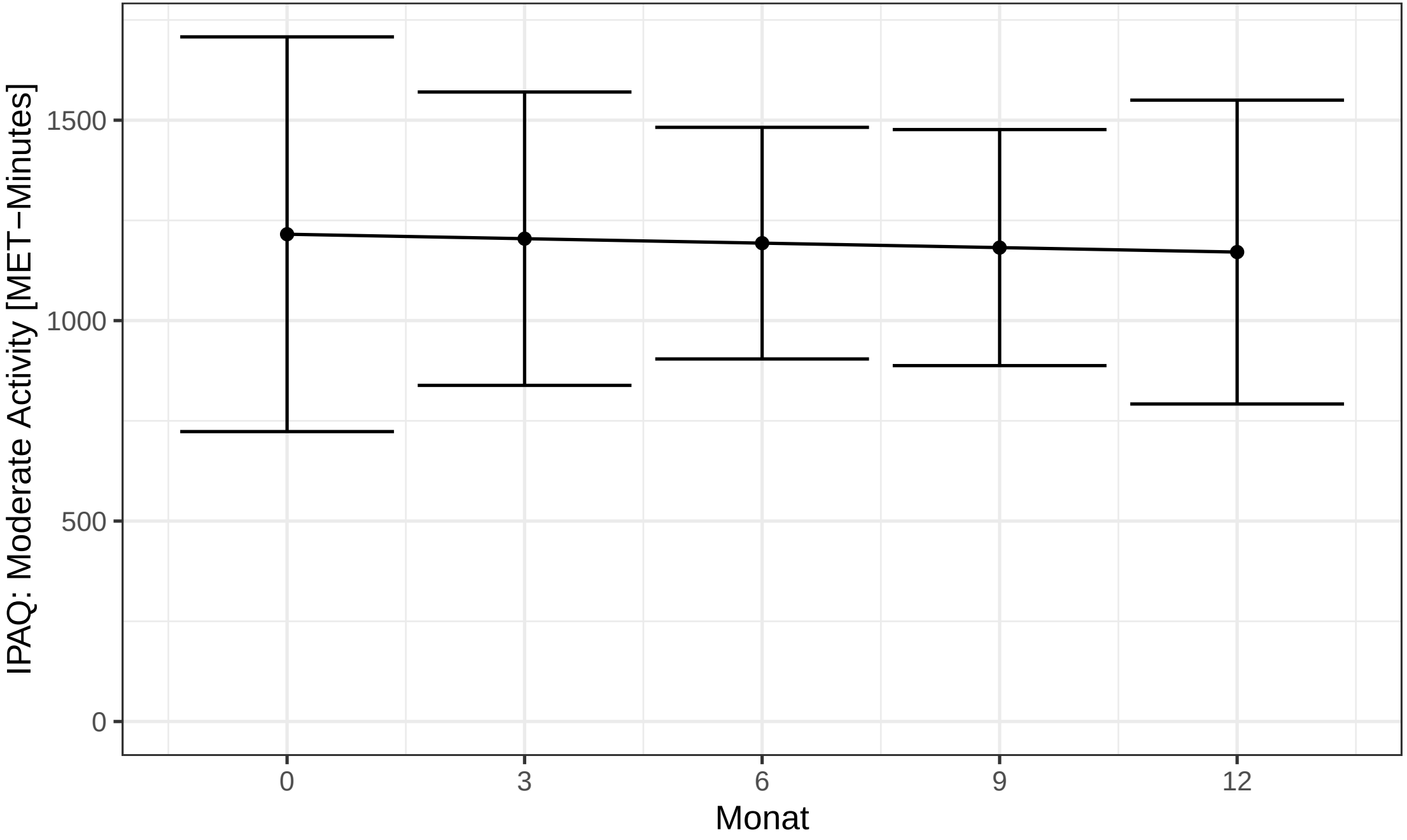
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

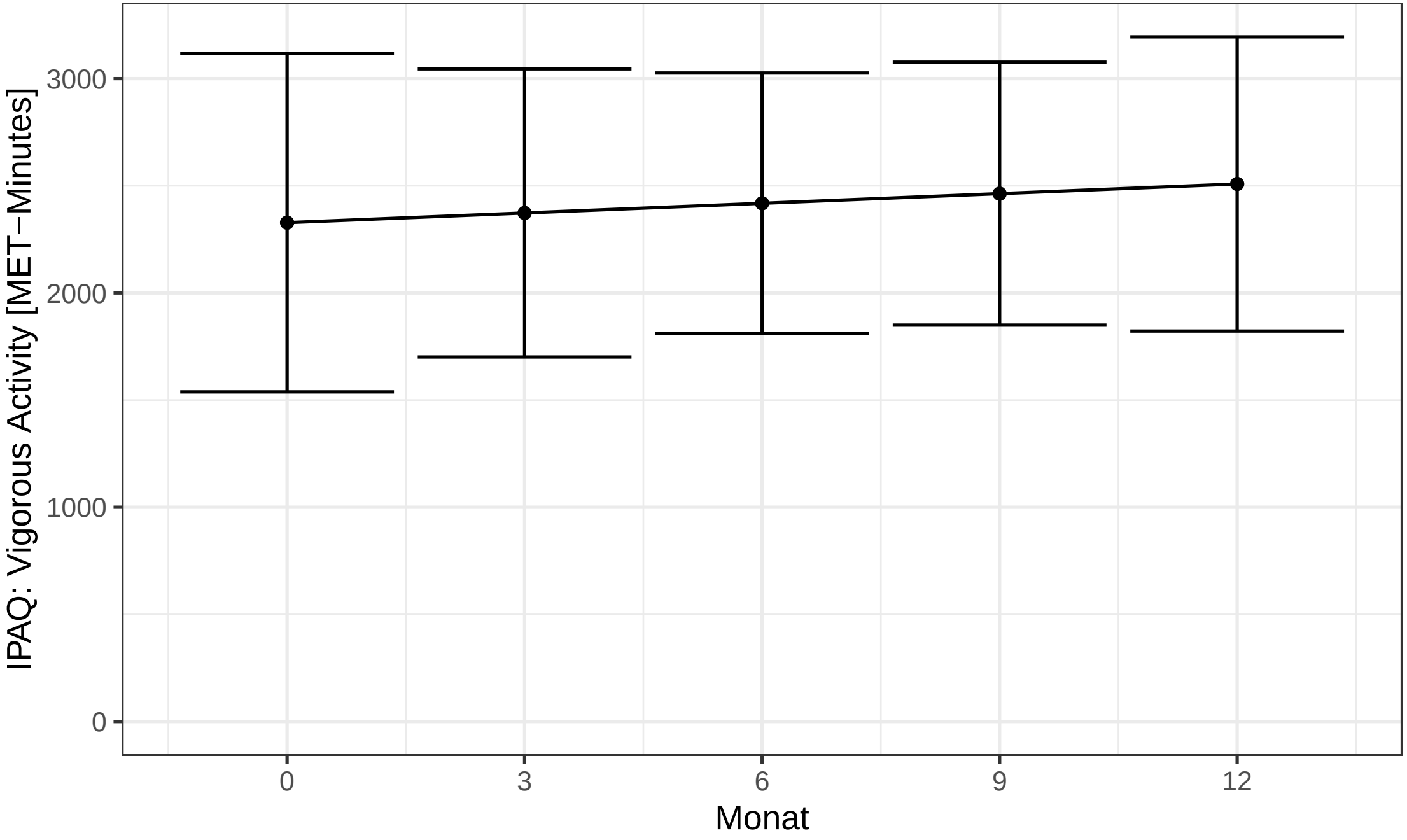
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

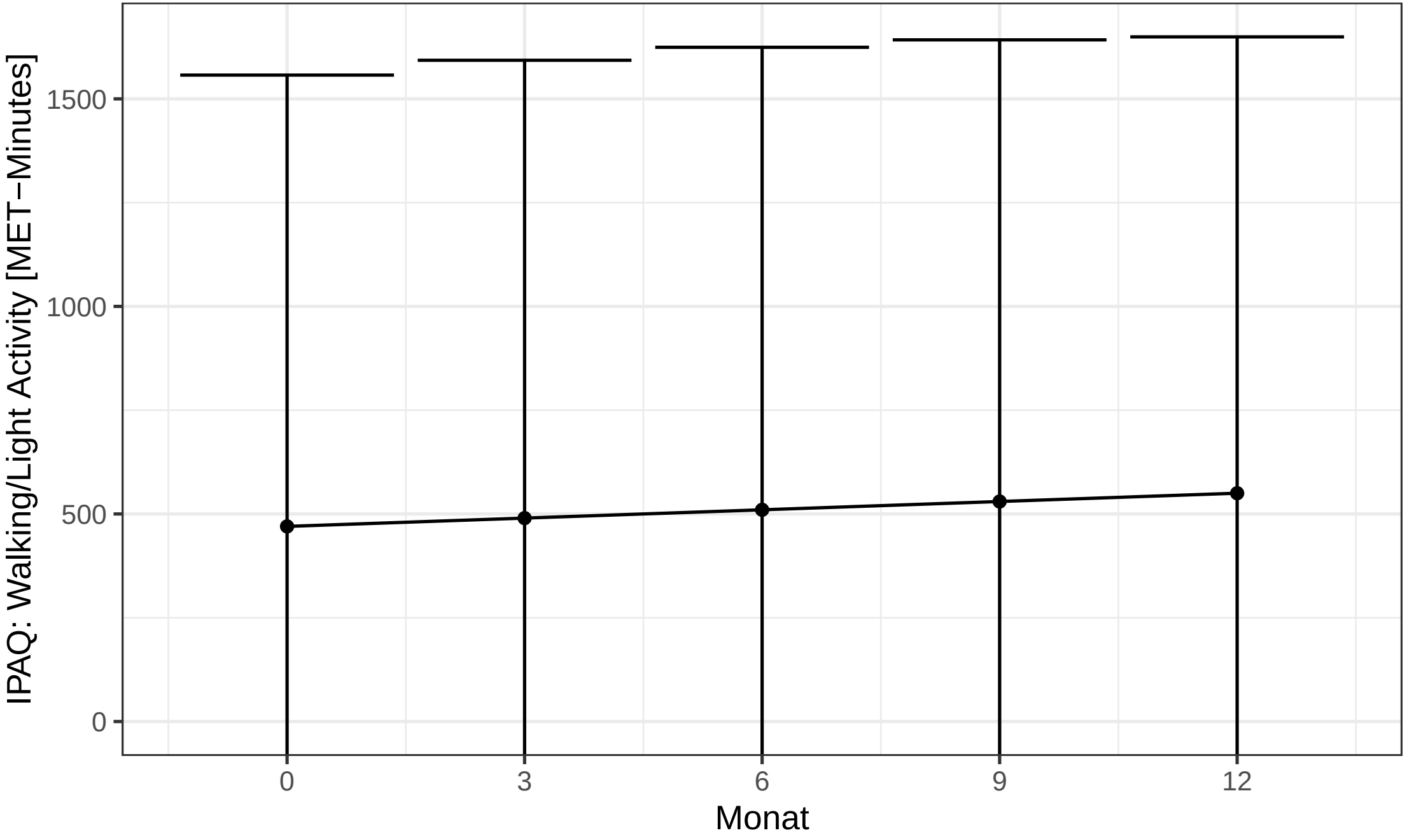
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

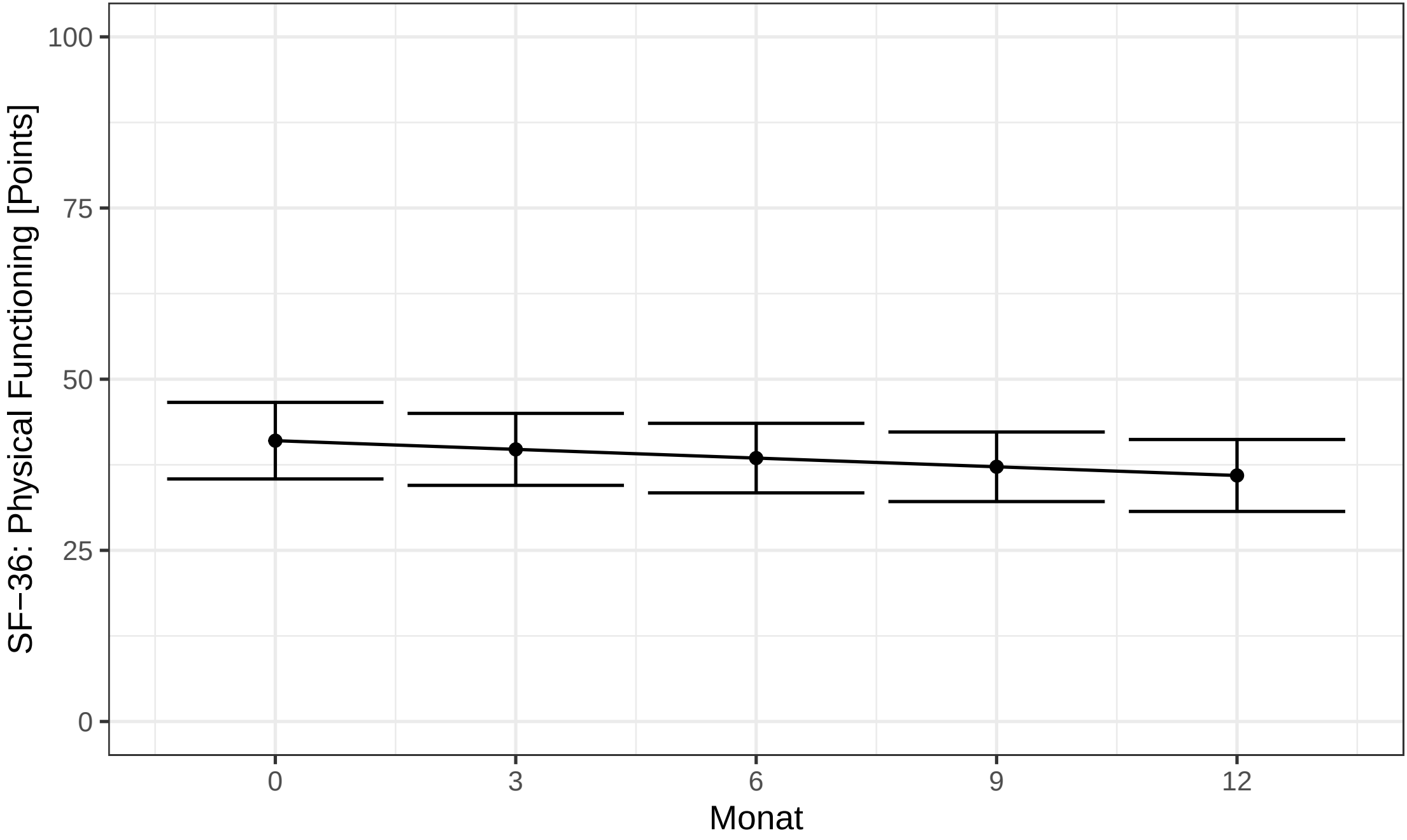
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

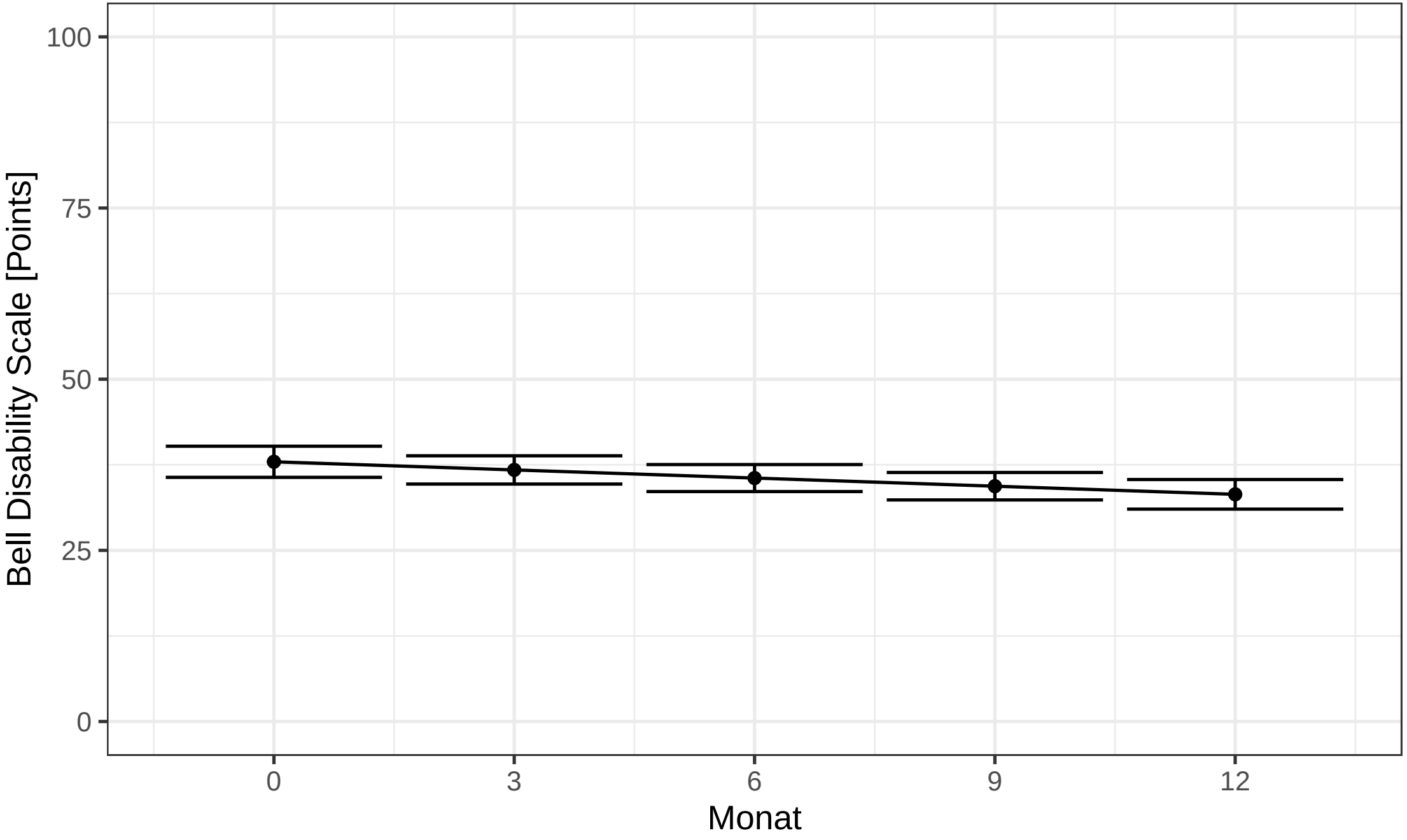
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

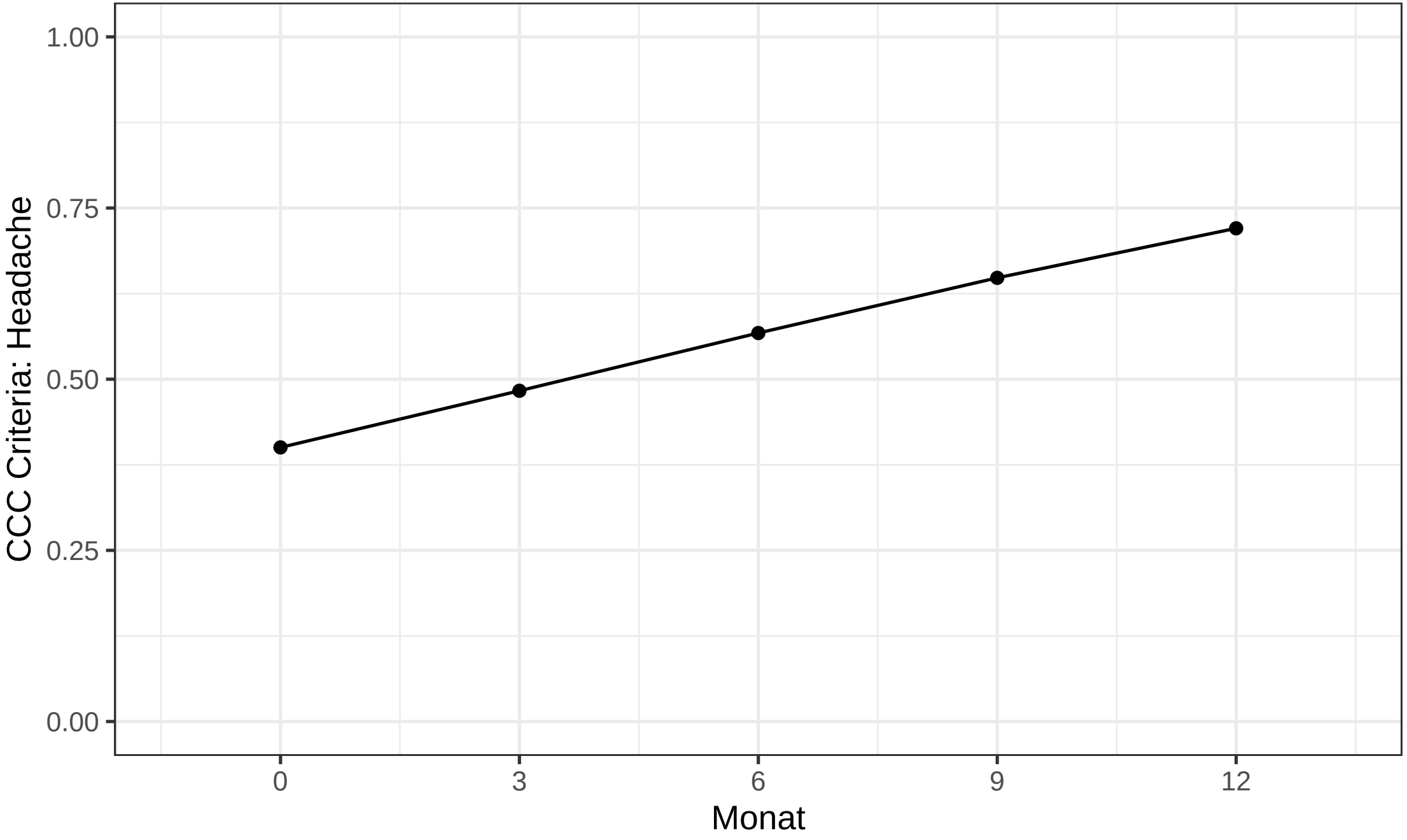
Adjustierter Mittelwert



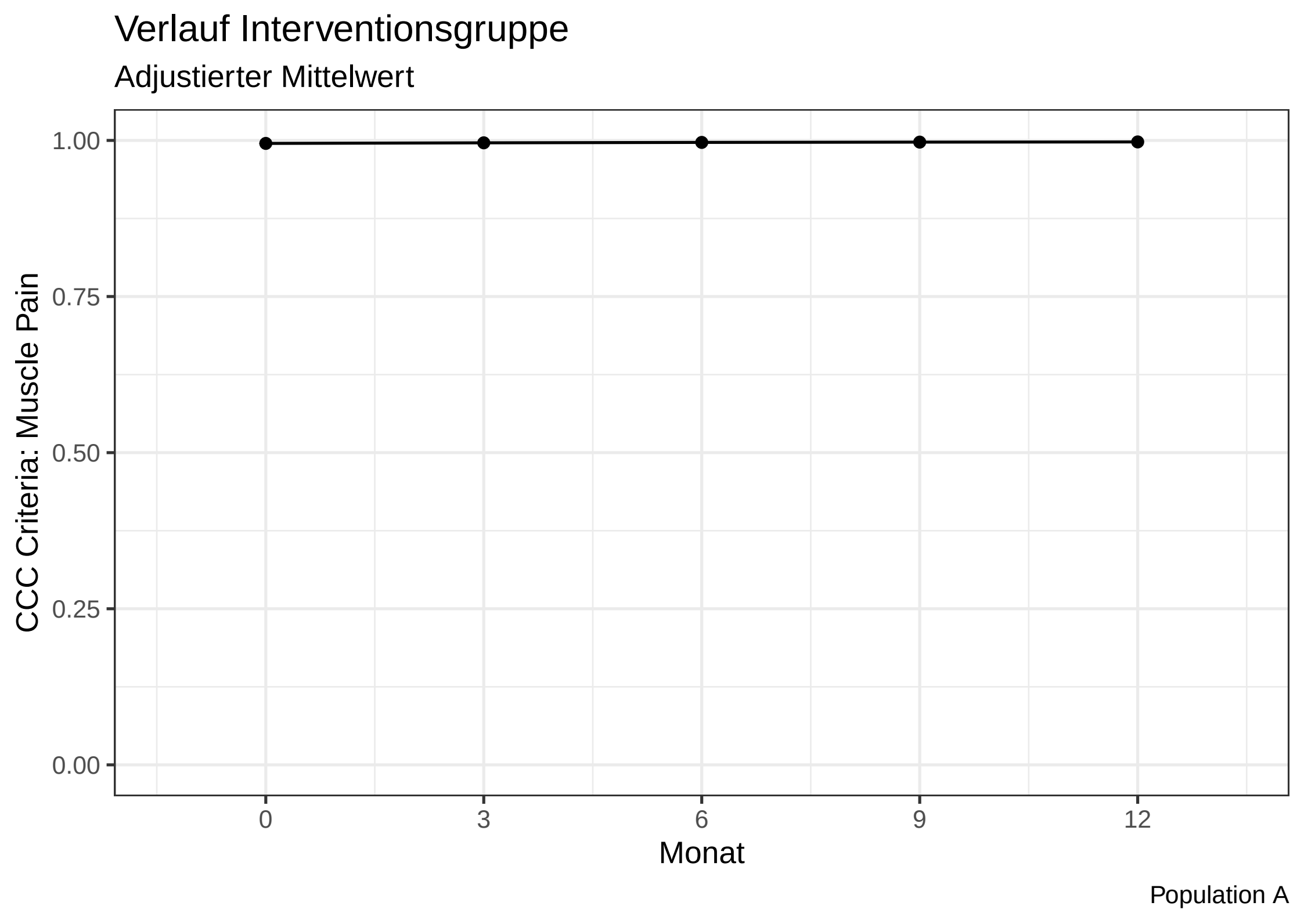
Population A

Verlauf Interventionsgruppe

Adjustierter Mittelwert

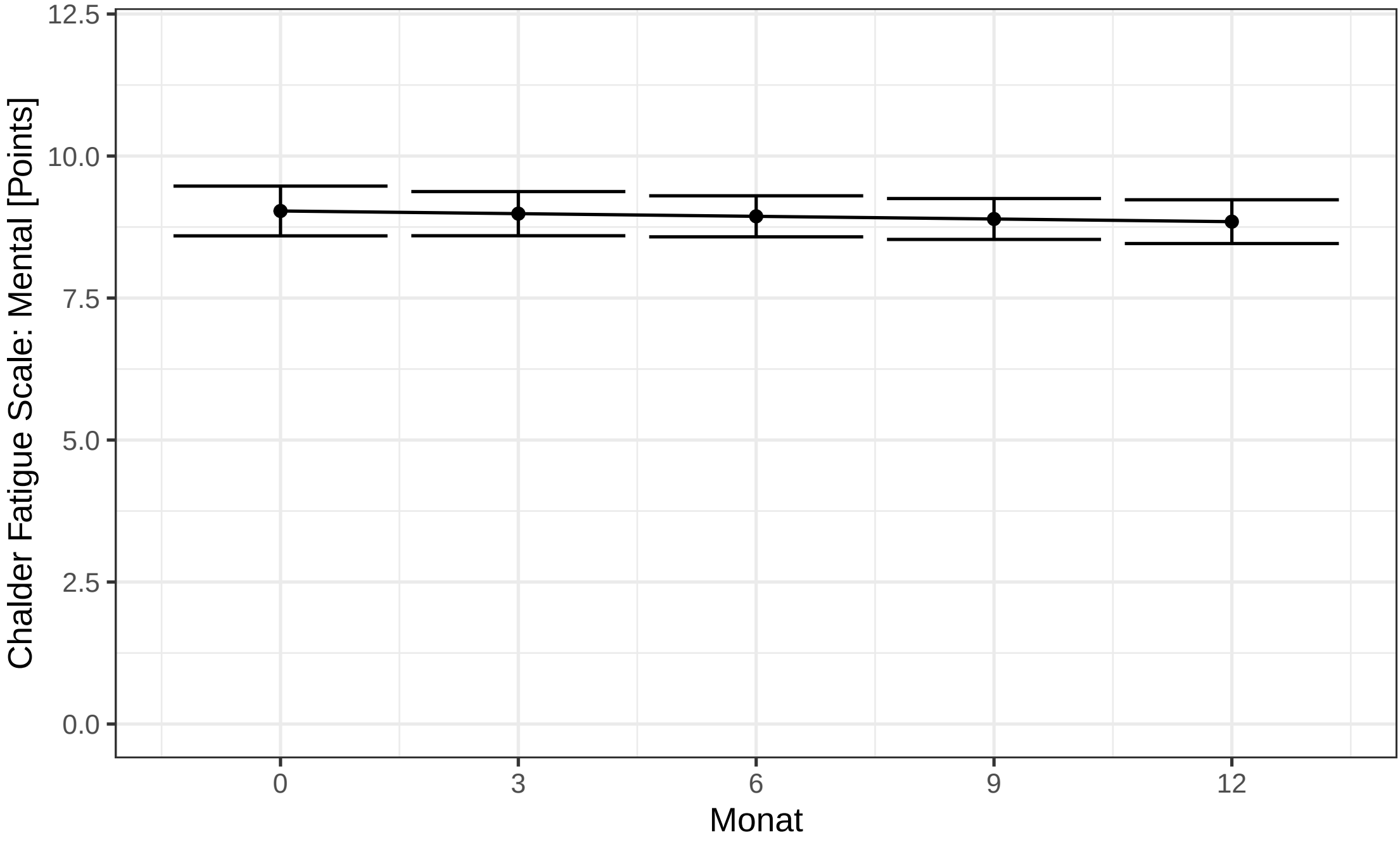


Population A


Verlauf Interventionsgruppe
Adjustierter Mittelwert
CCC Criteria: Muscle Pain
1.00
0.75
0.50
0.25
0.00
0
3
6
9
12
Monat
Population A

Verlauf Interventionsgruppe

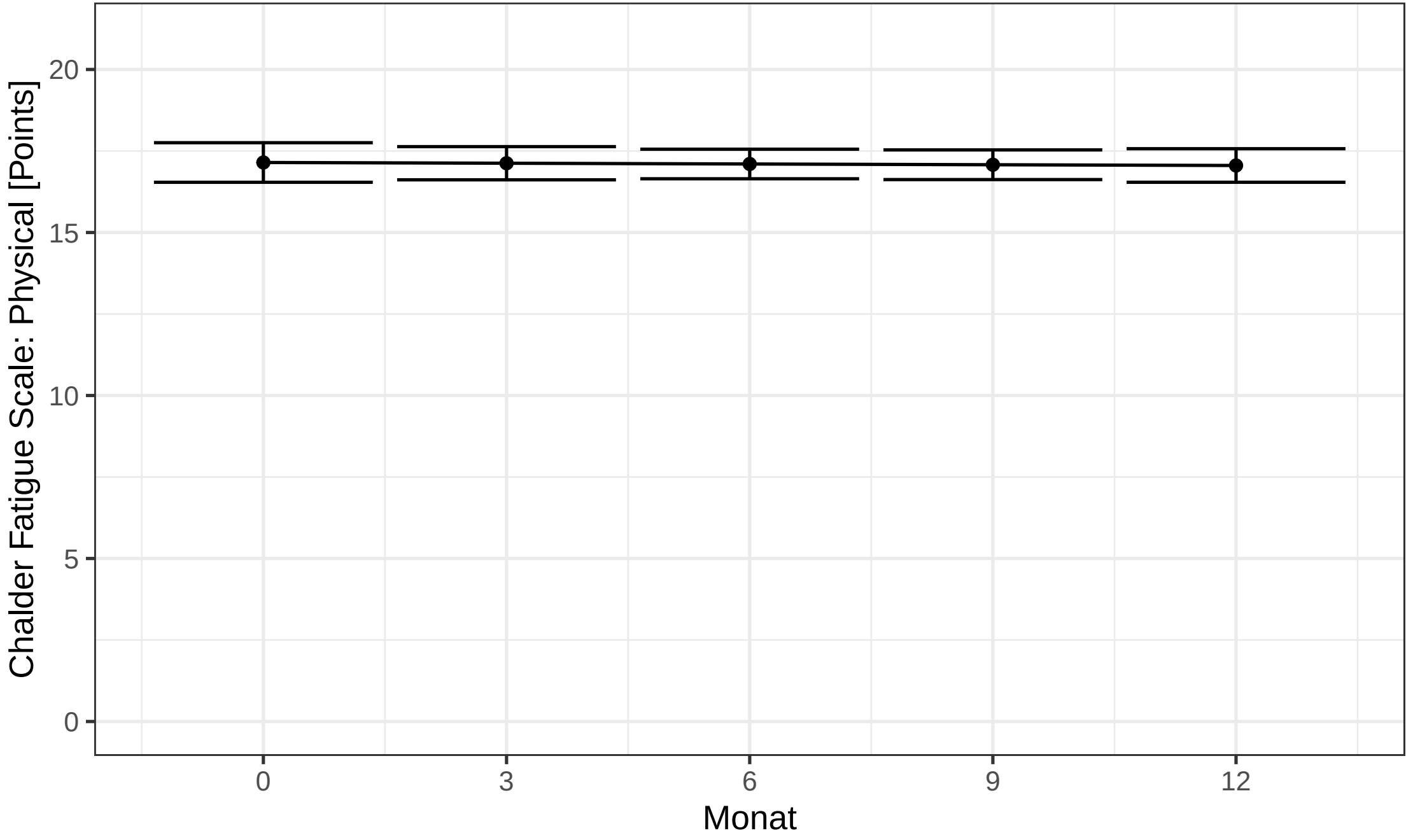
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

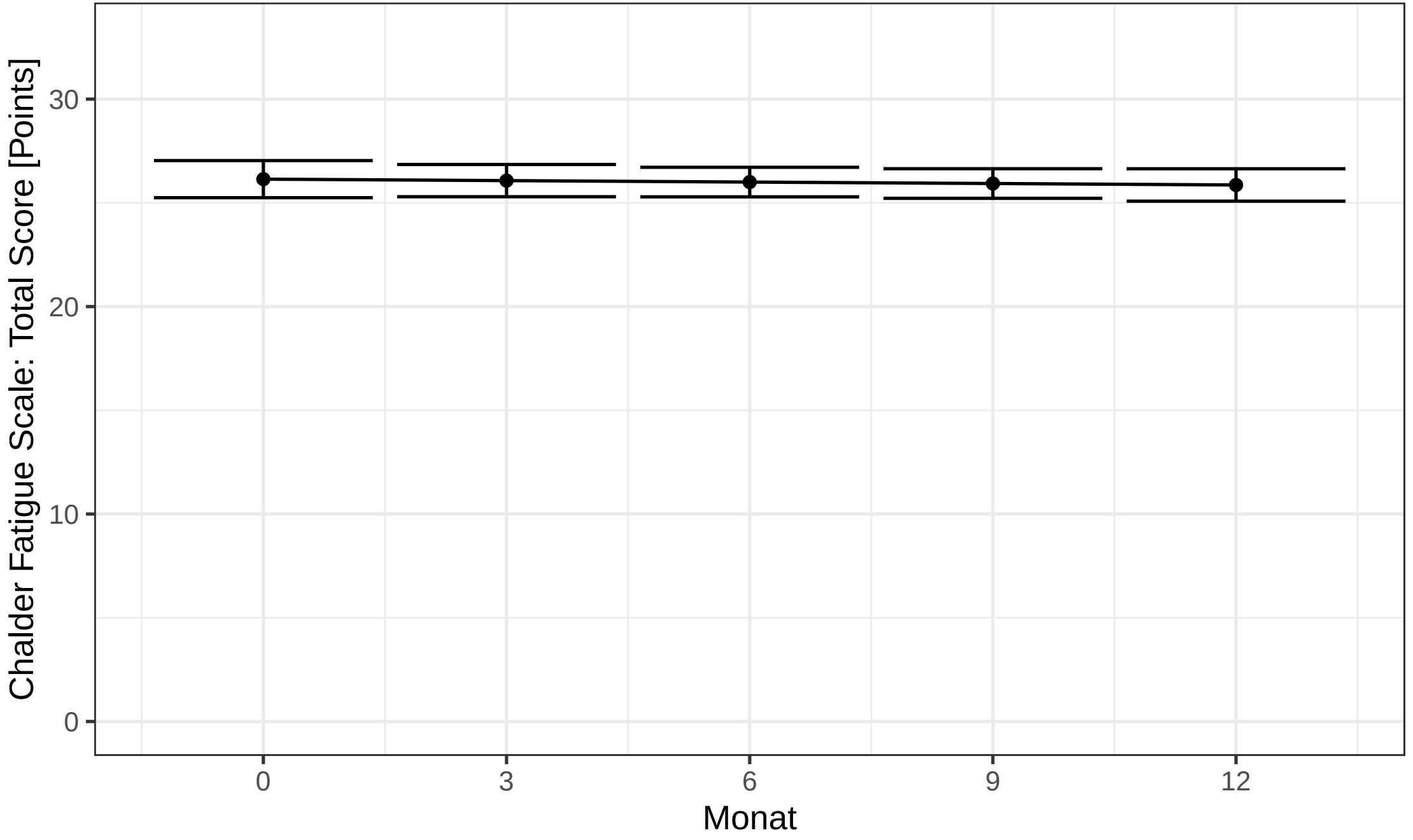
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

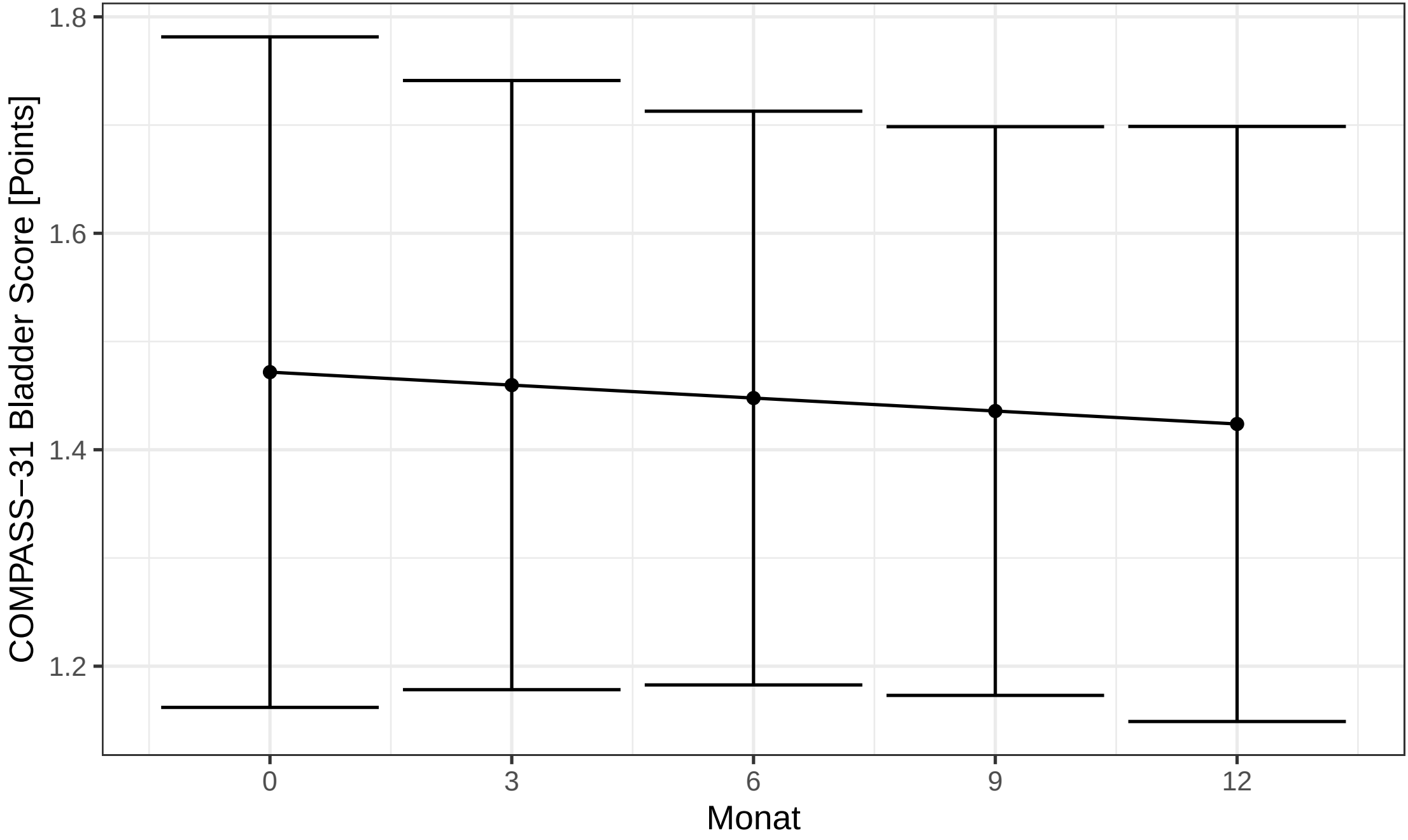
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

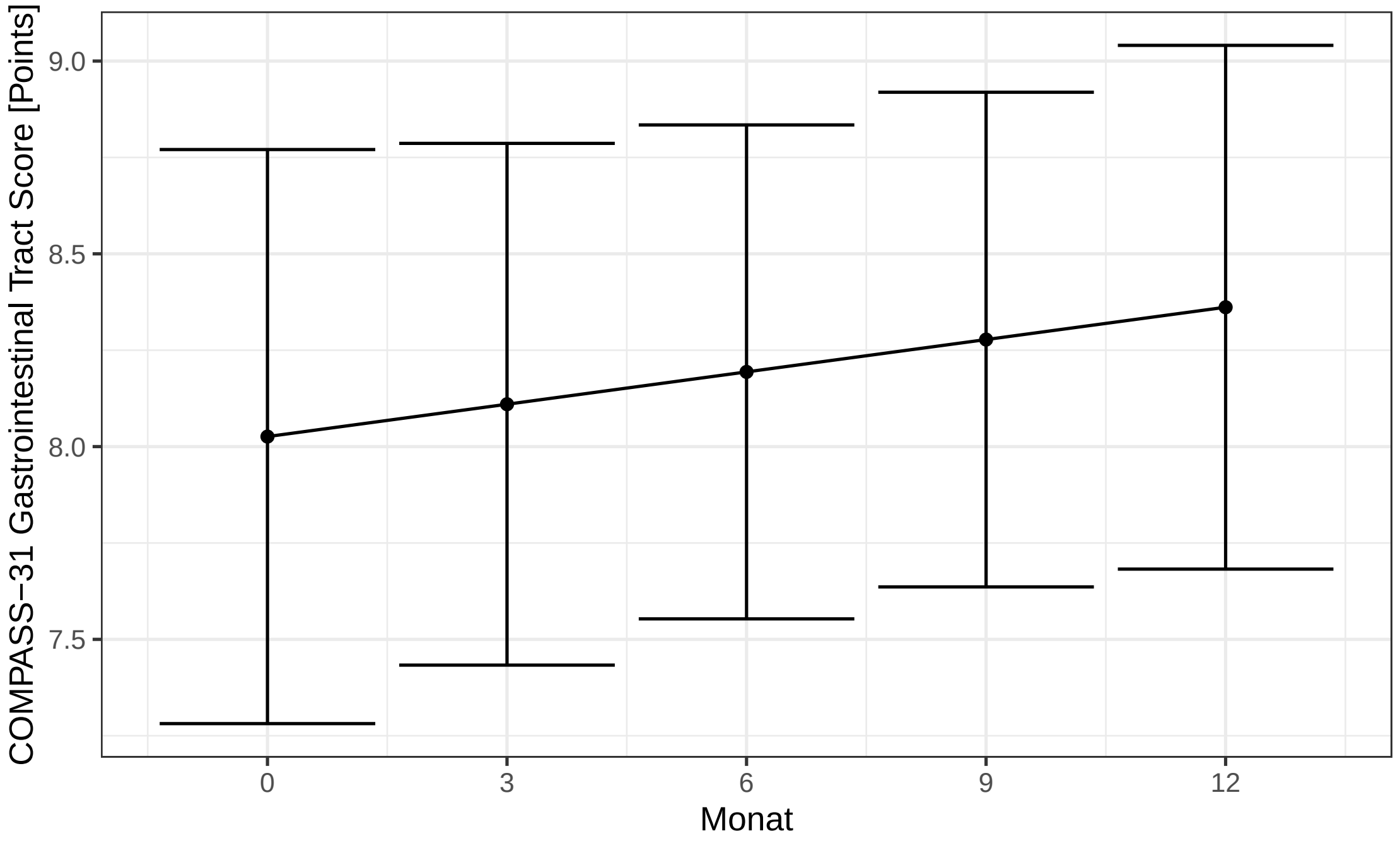
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

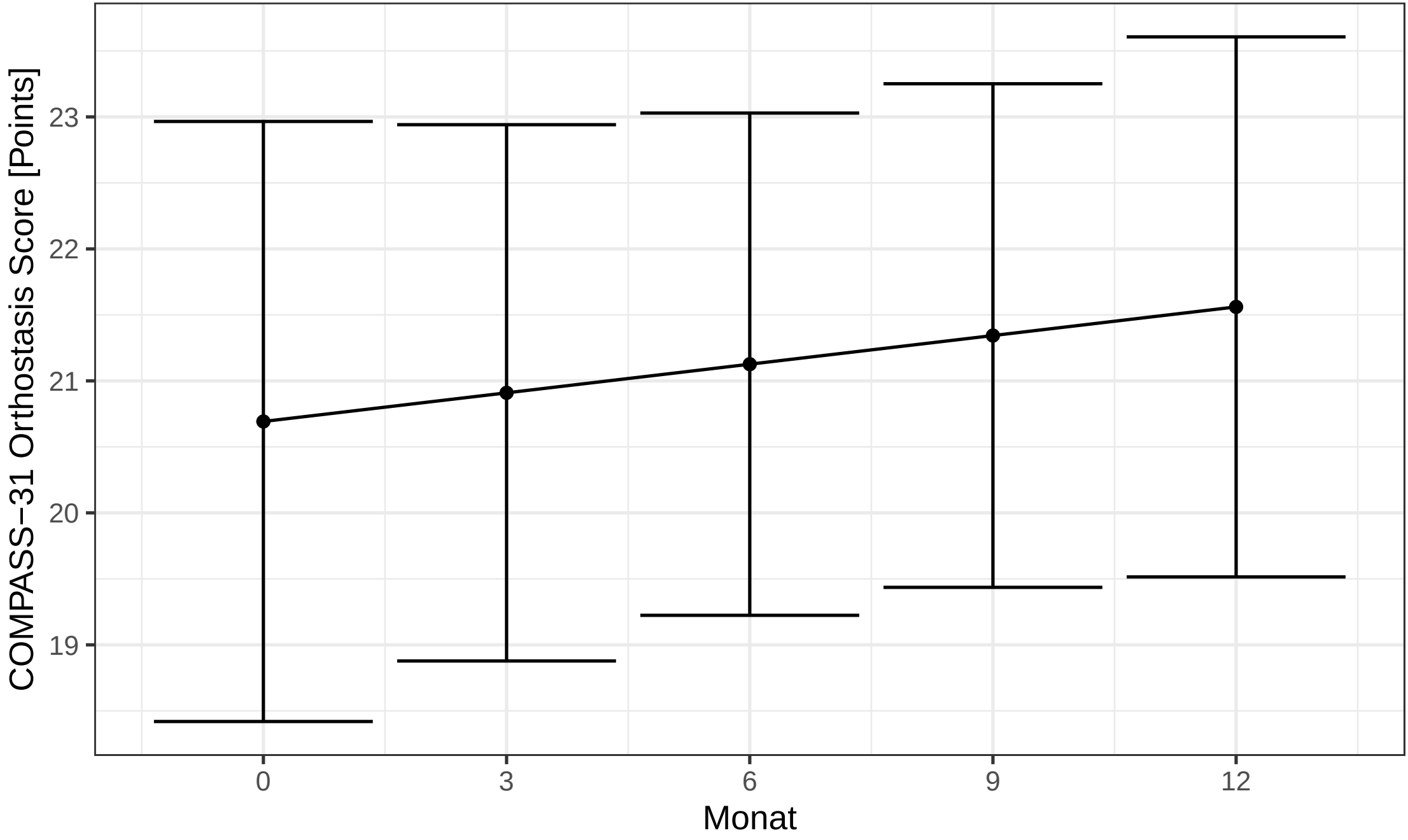
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

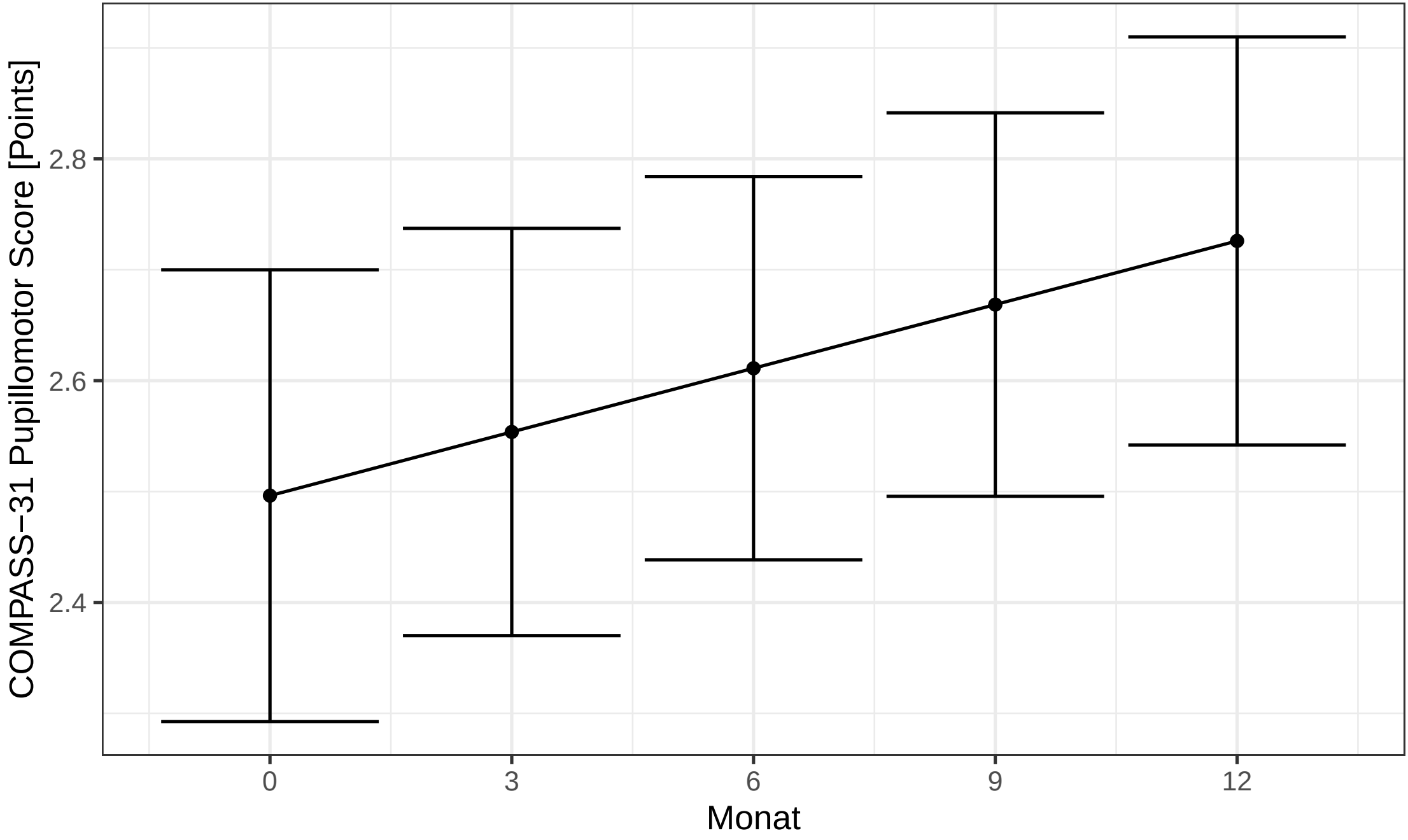
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

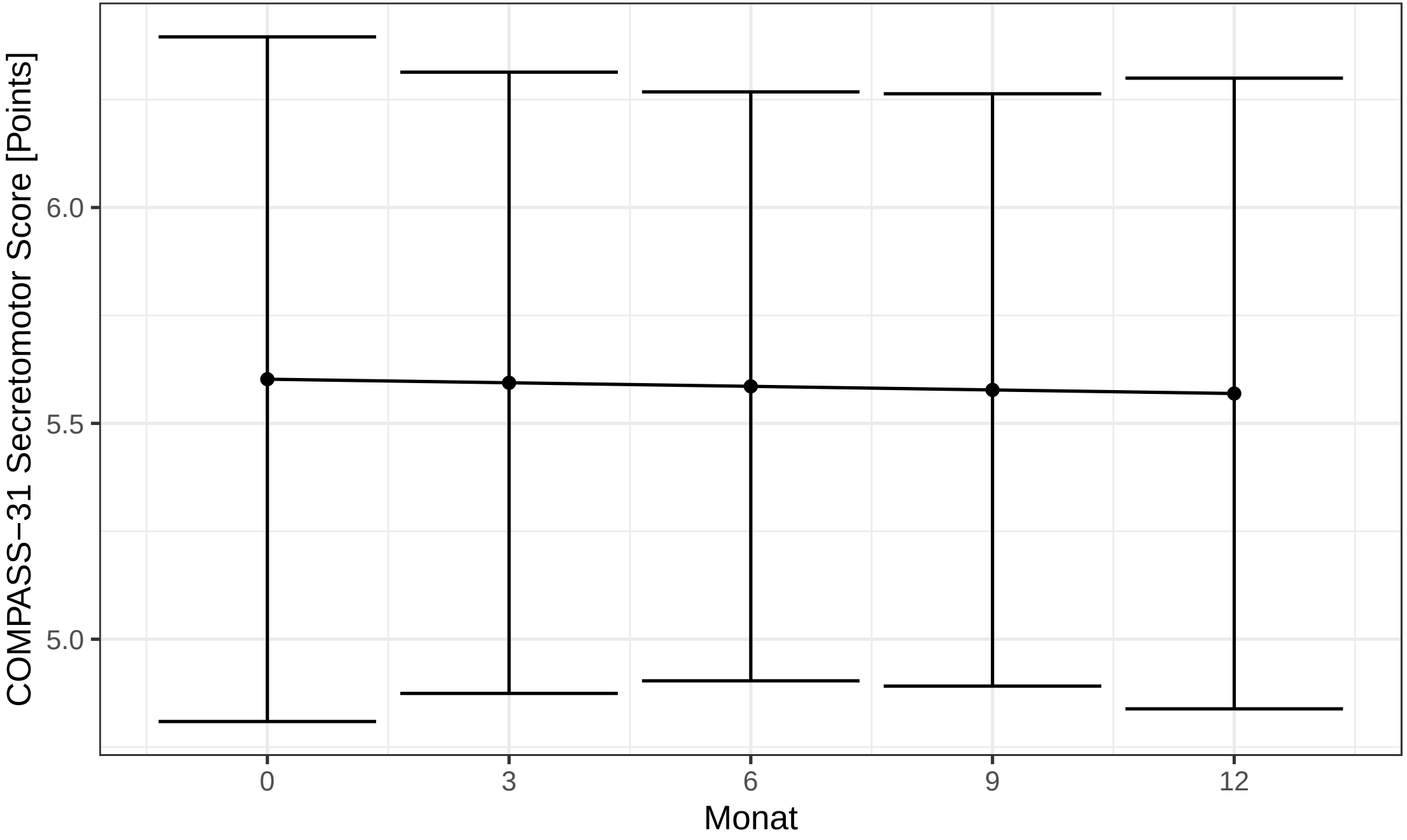
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

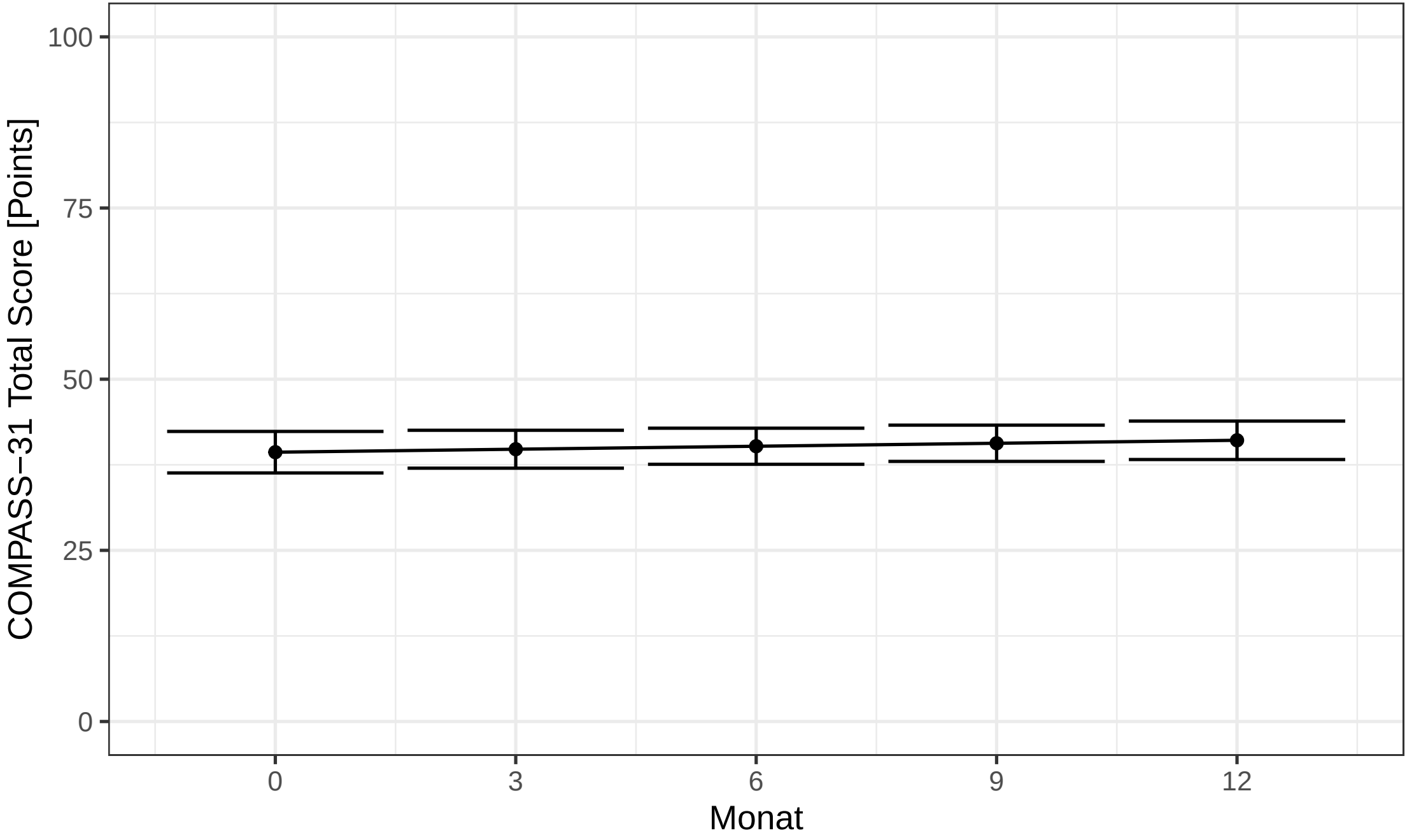
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

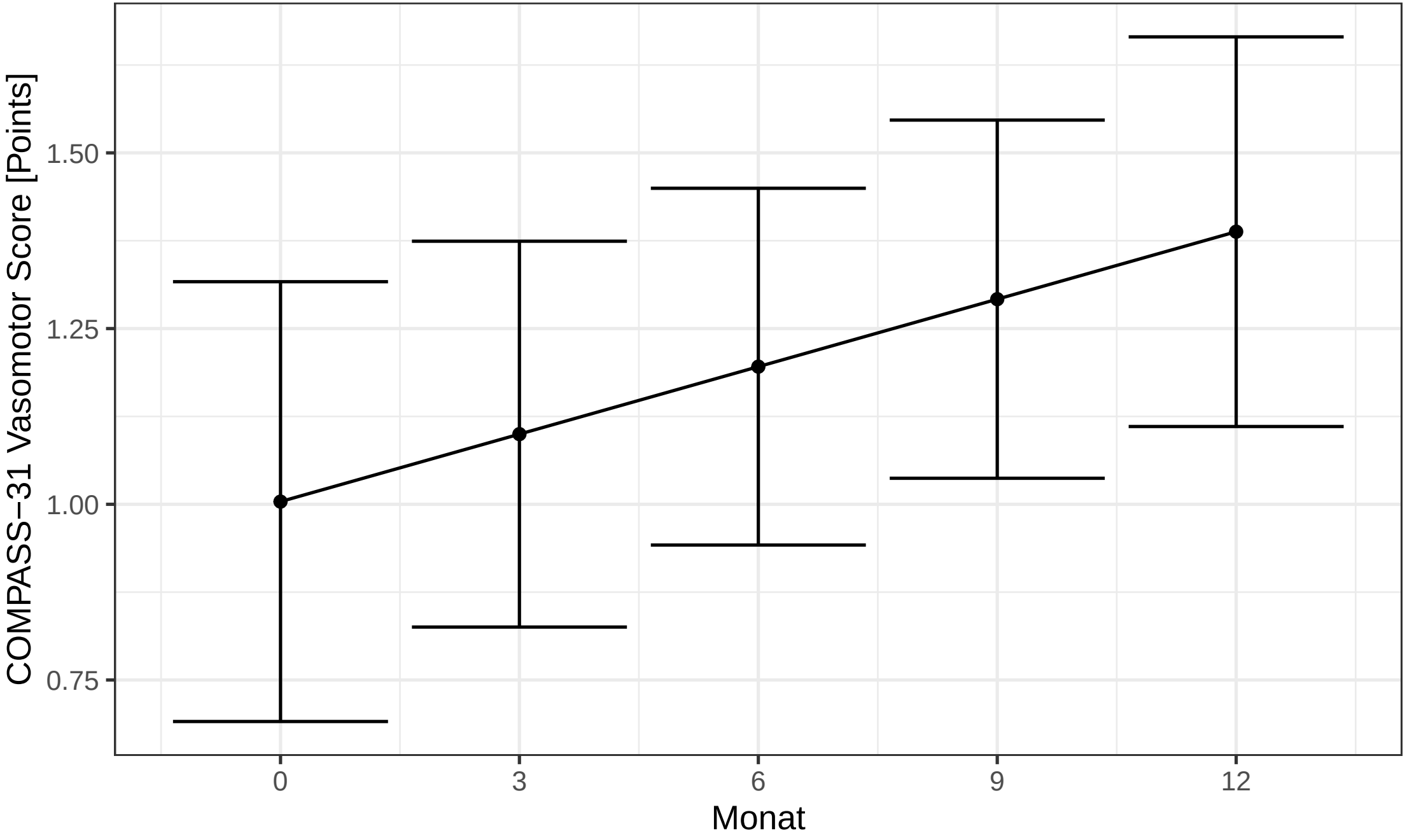
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

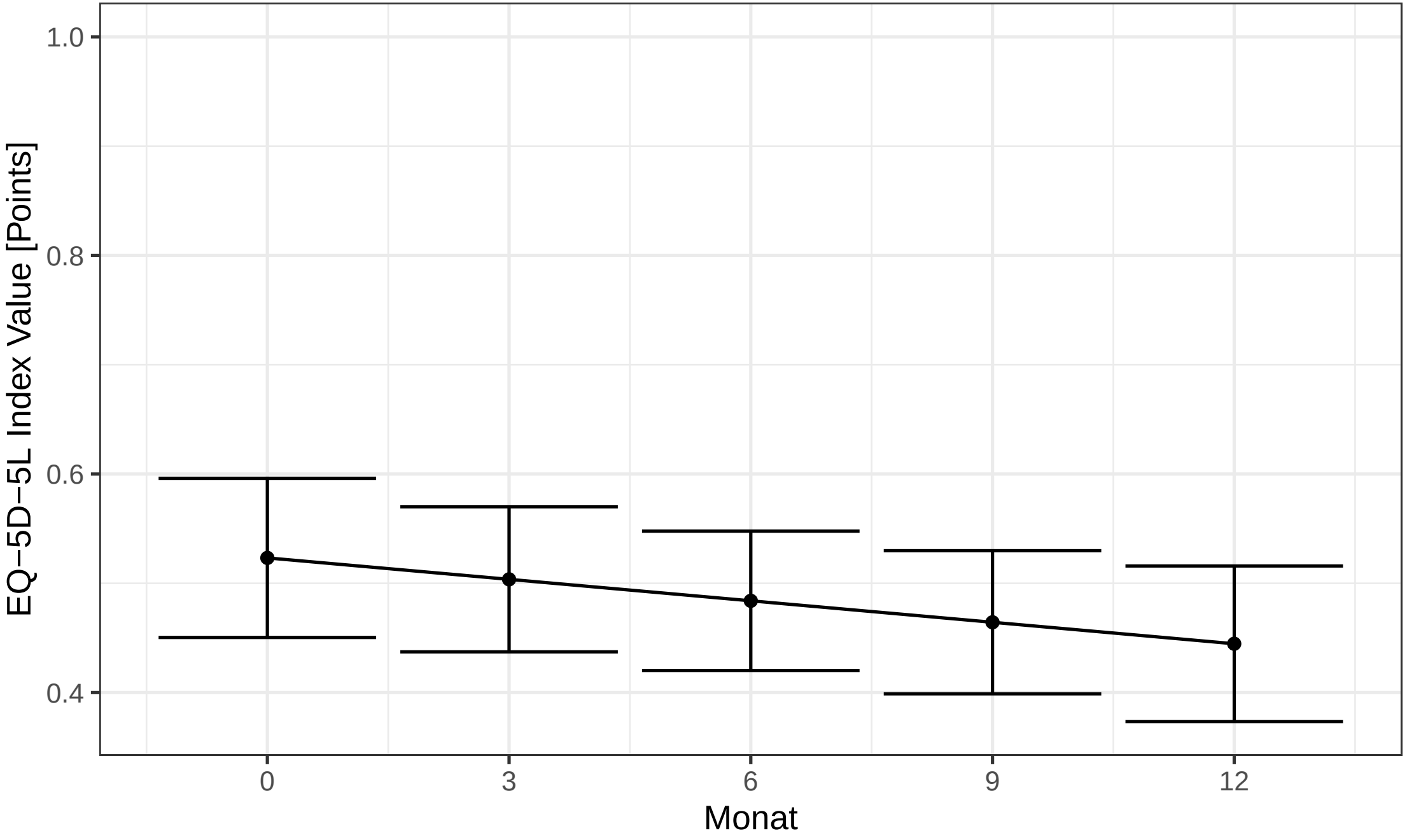
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

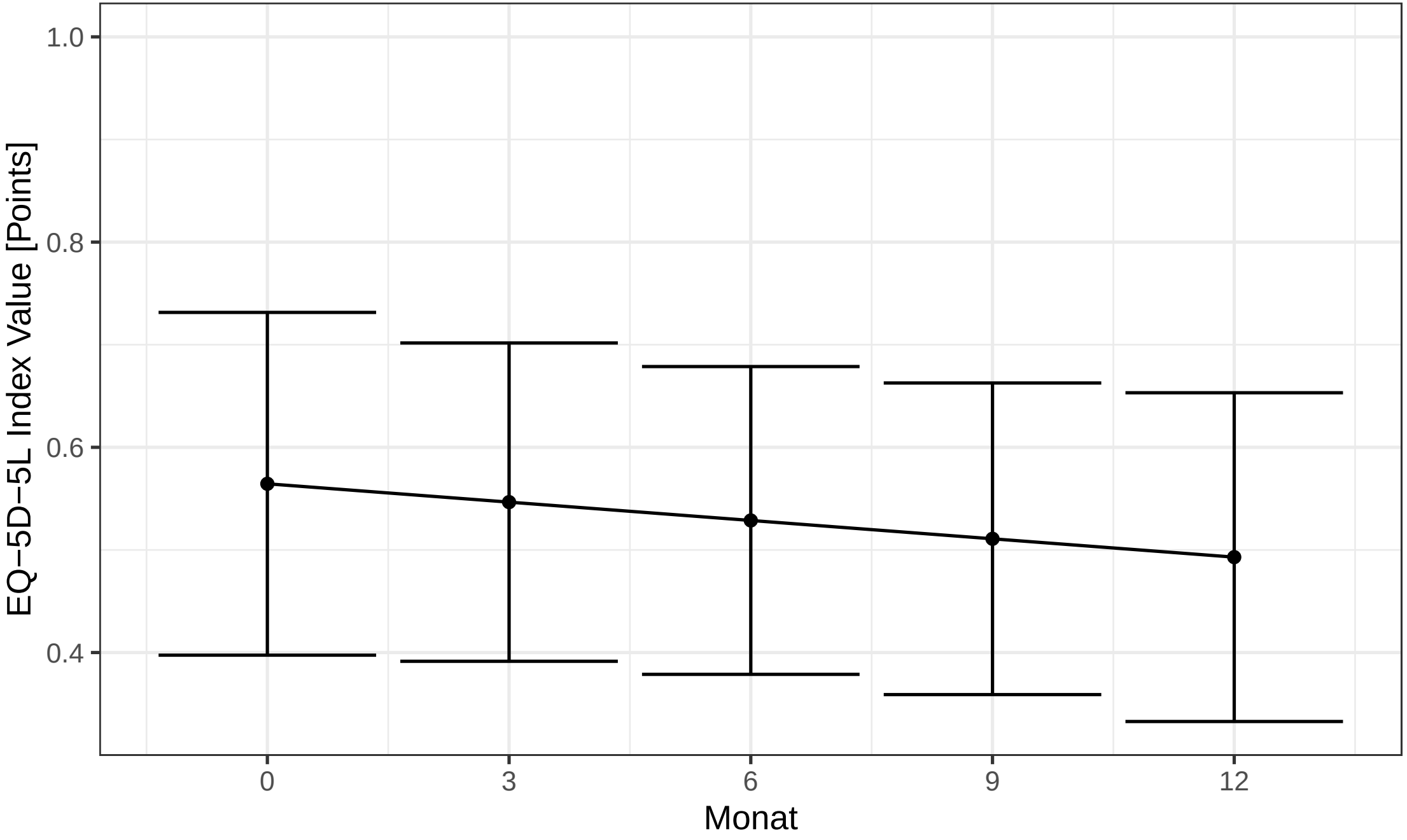
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

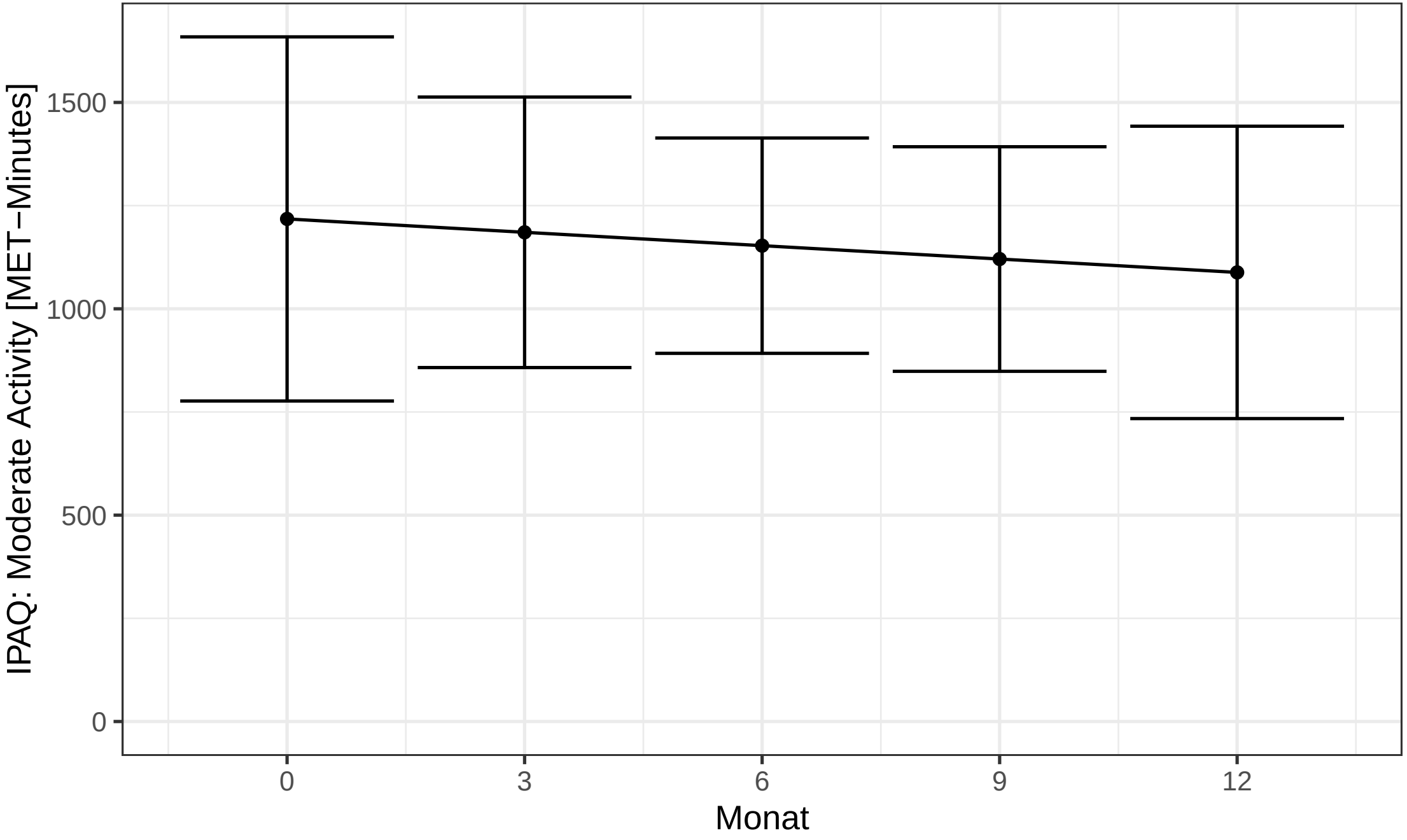
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

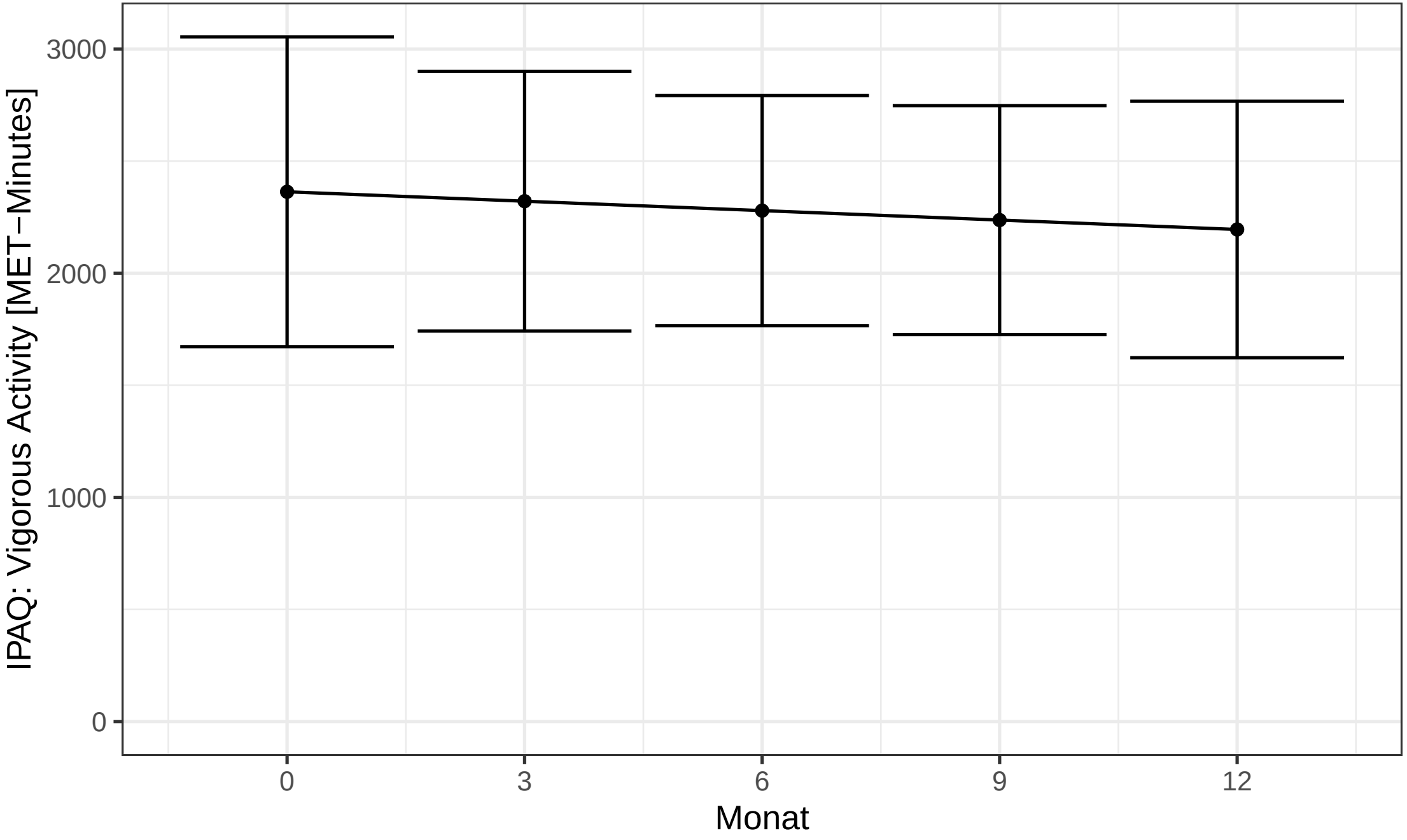
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

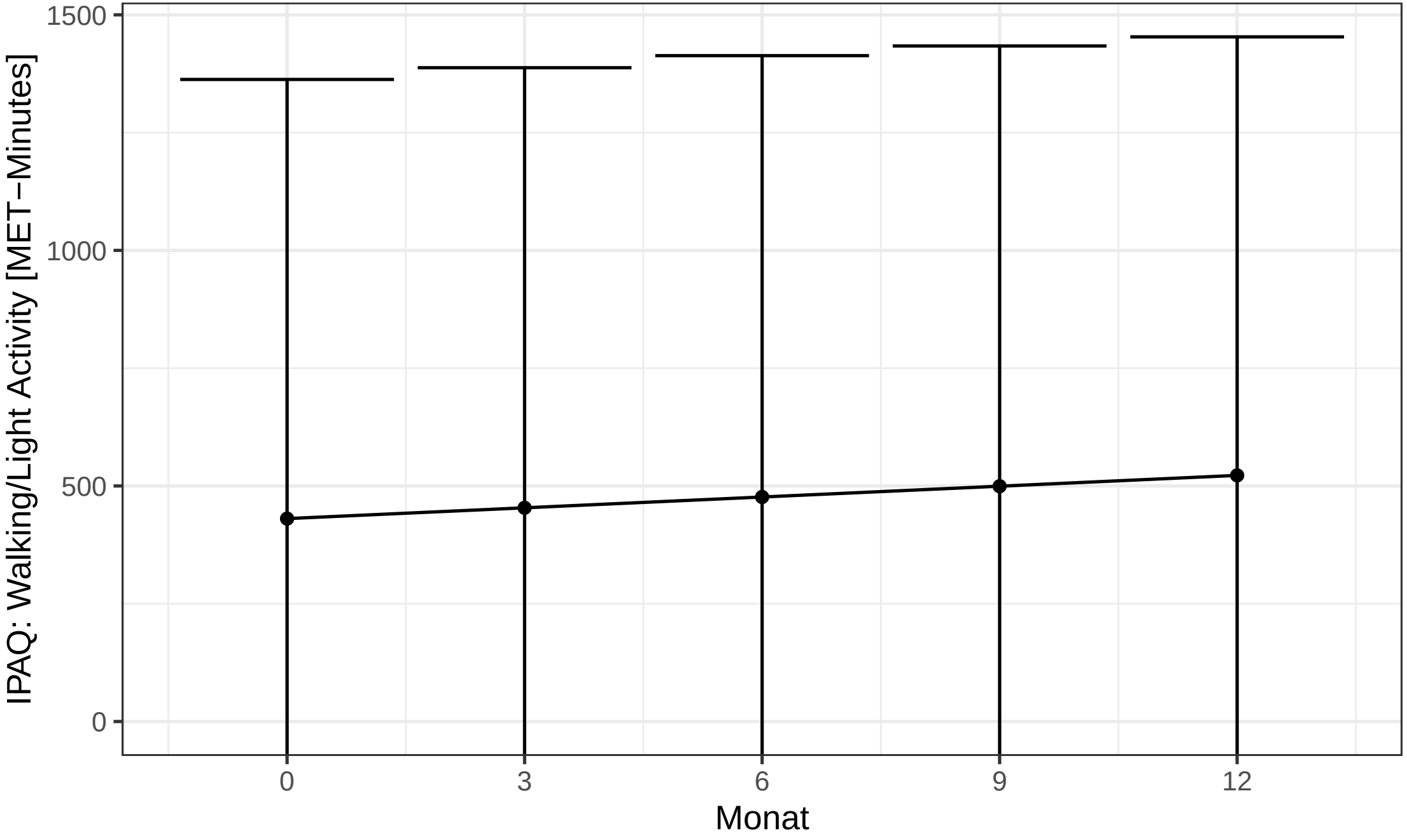
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

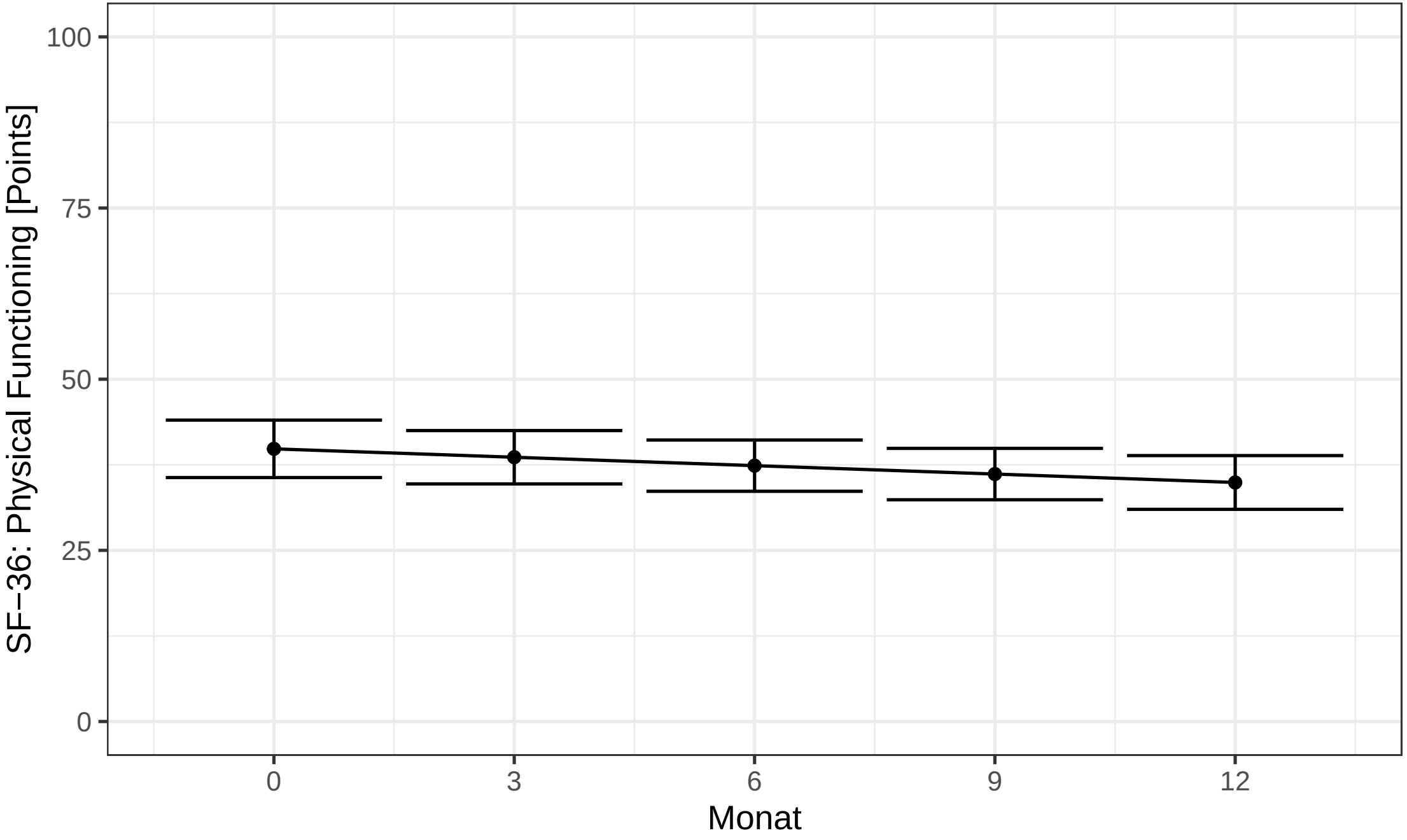
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

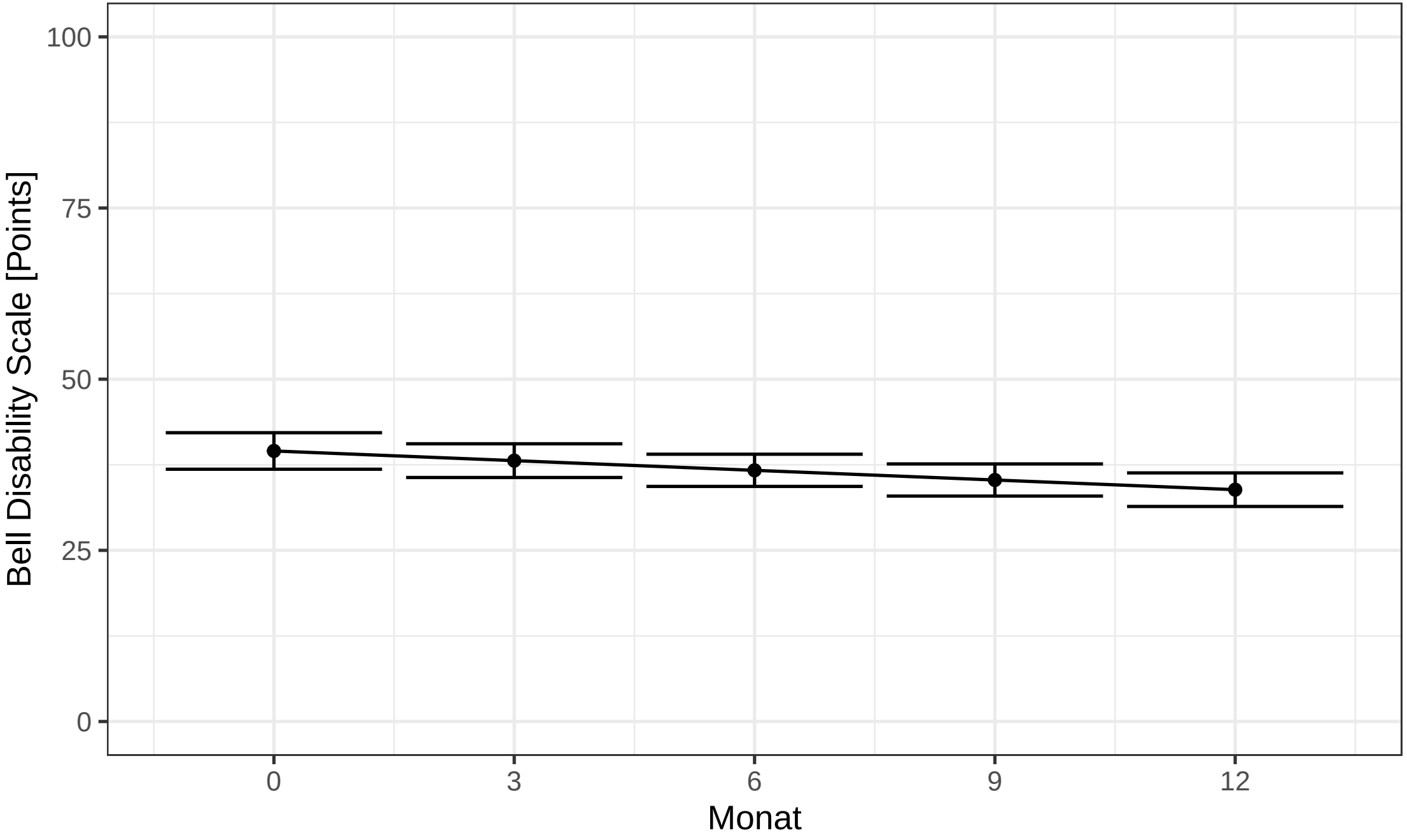
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

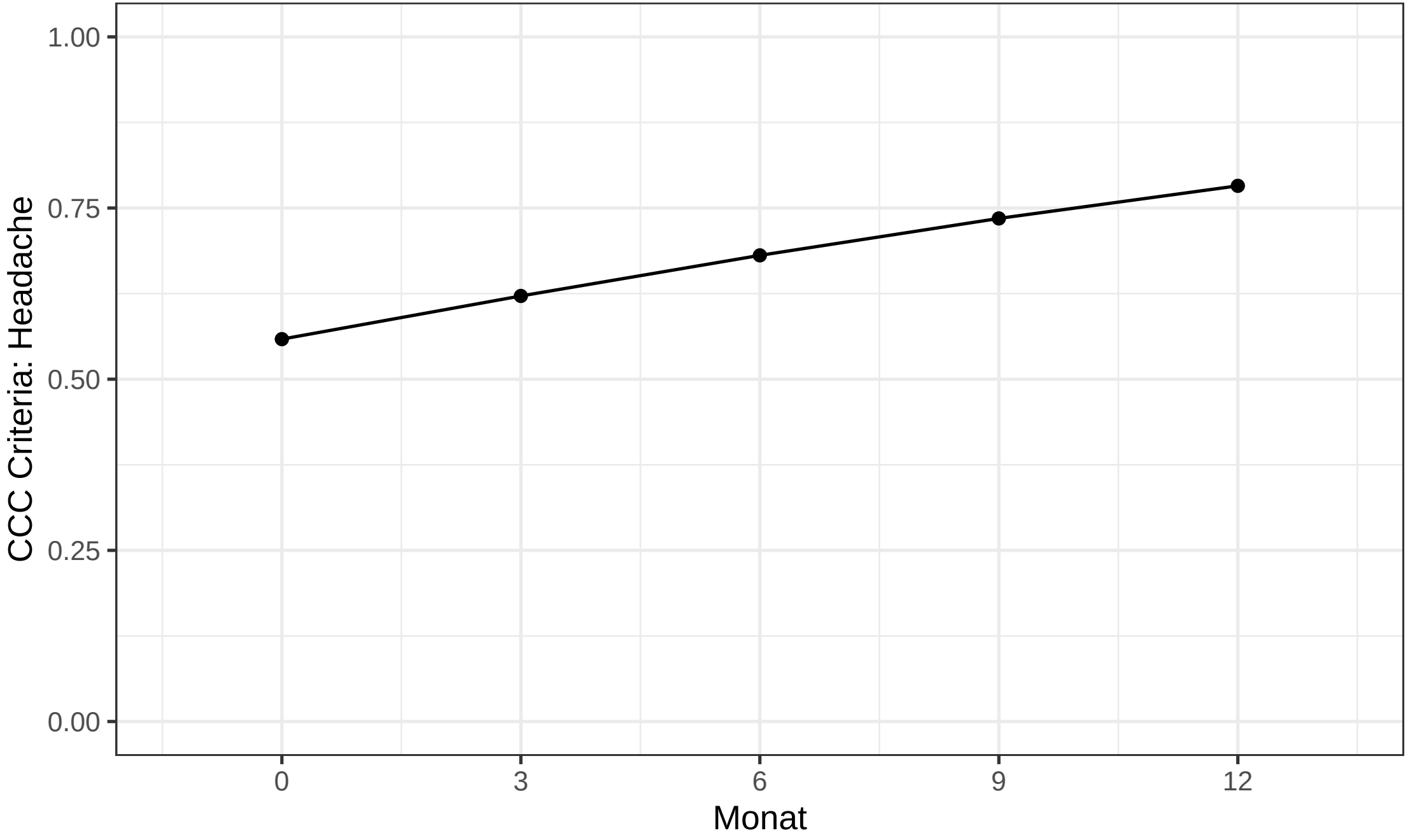
Adjustierter Mittelwert



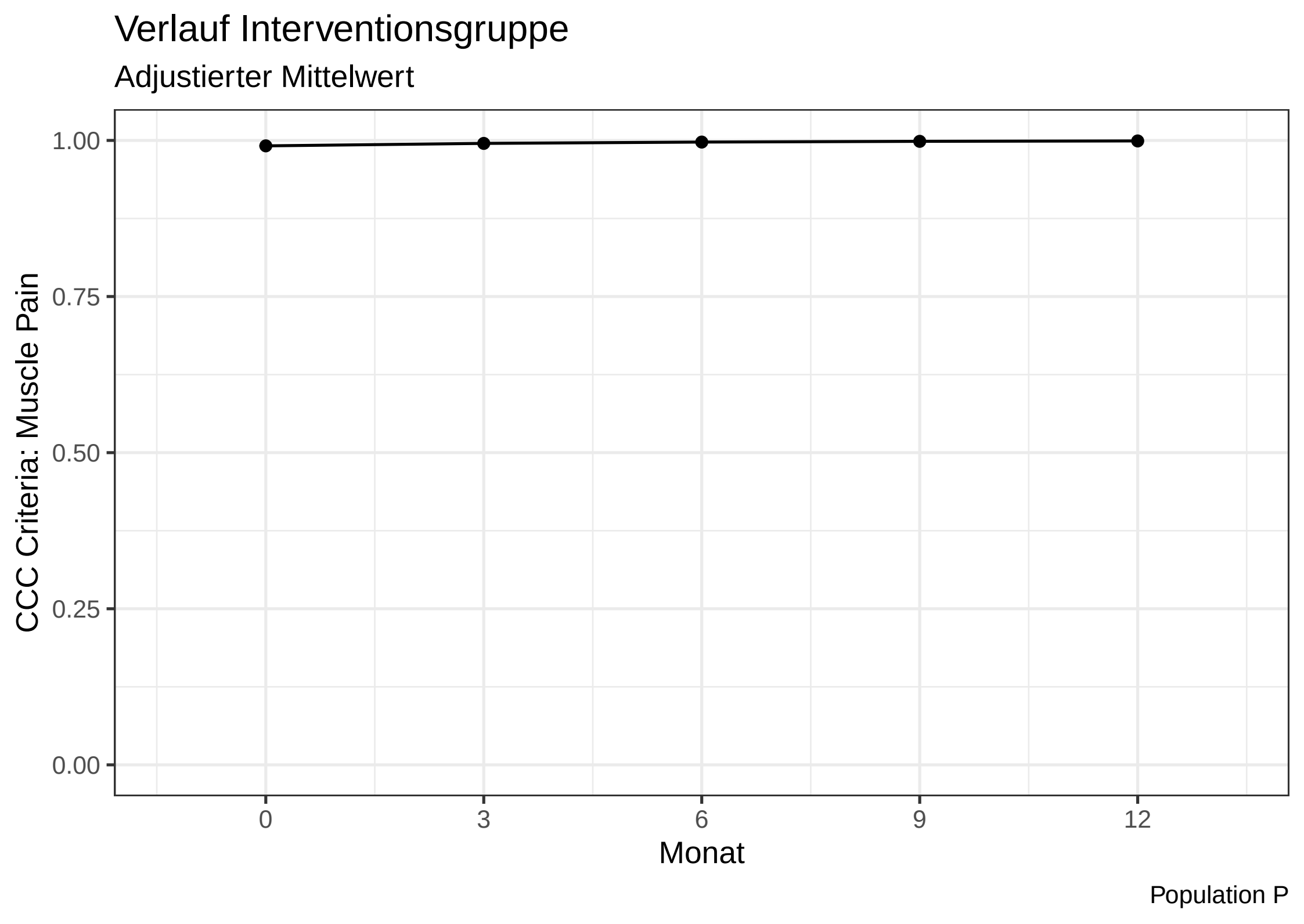
Population P

Verlauf Interventionsgruppe

Adjustierter Mittelwert

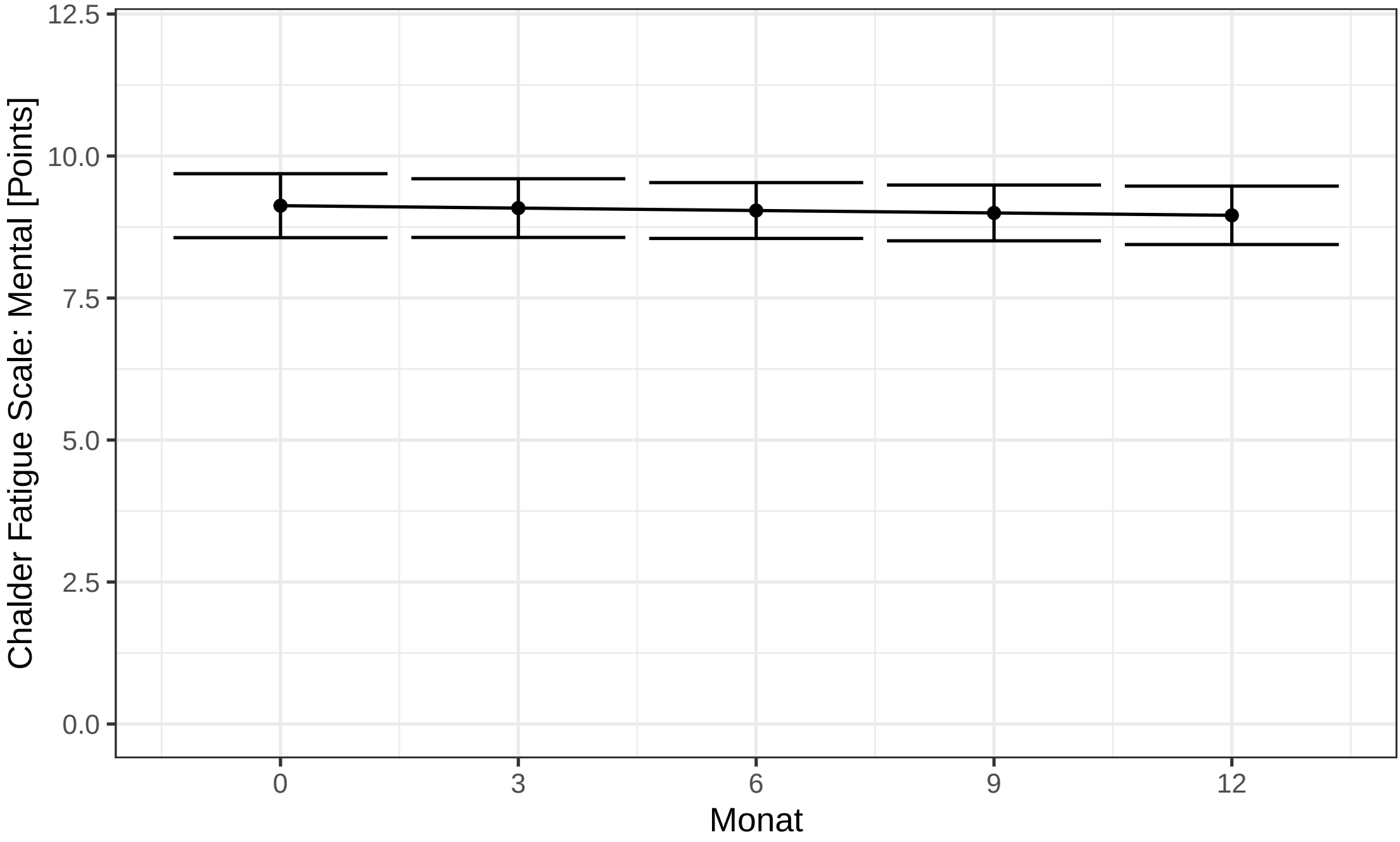


Population P


Verlauf Interventionsgruppe
Adjustierter Mittelwert
CCC Criteria: Muscle Pain
1.00
0.75
0.50
0.25
0.00
0
3
6
9
12
Monat
Population P

Verlauf Interventionsgruppe

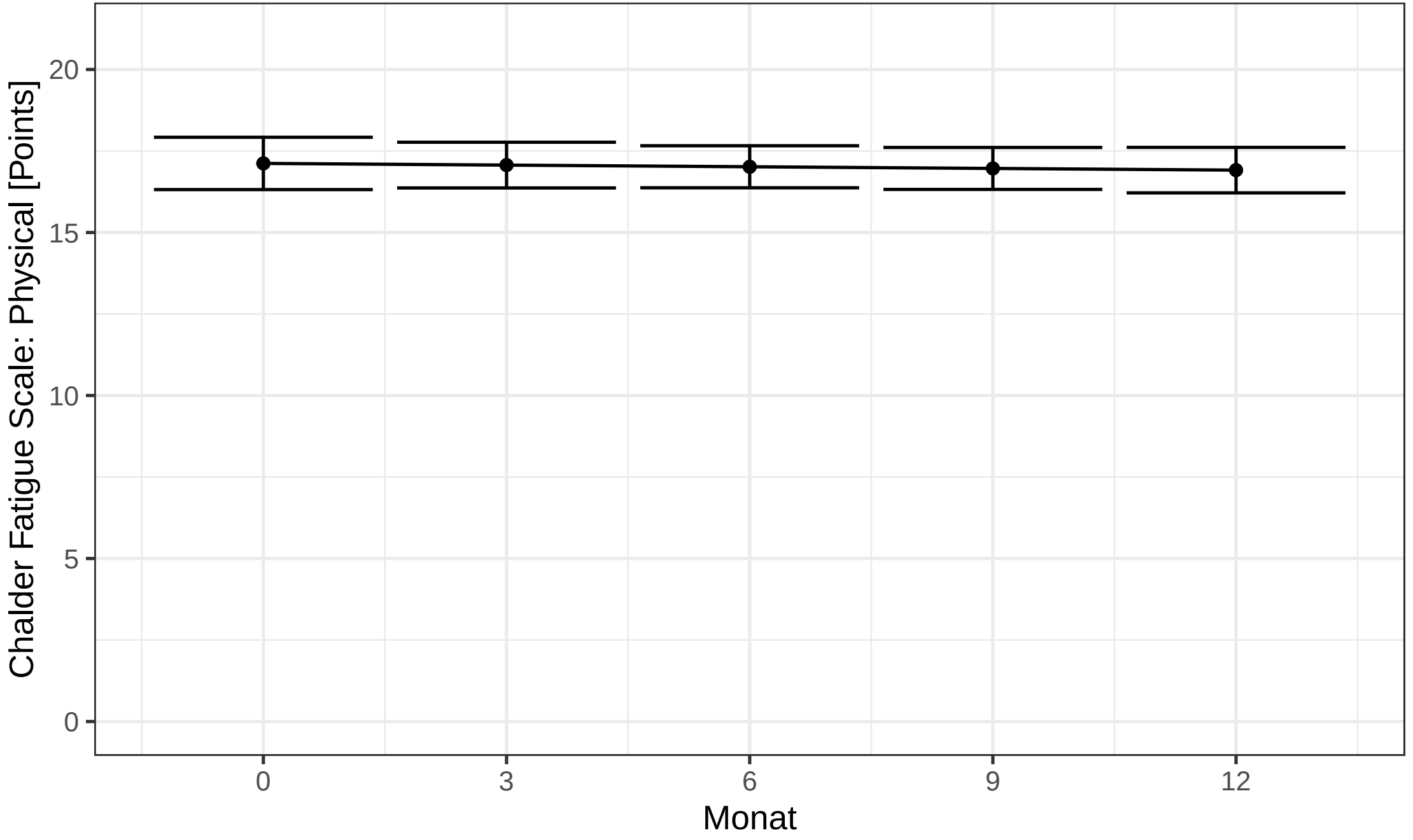
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

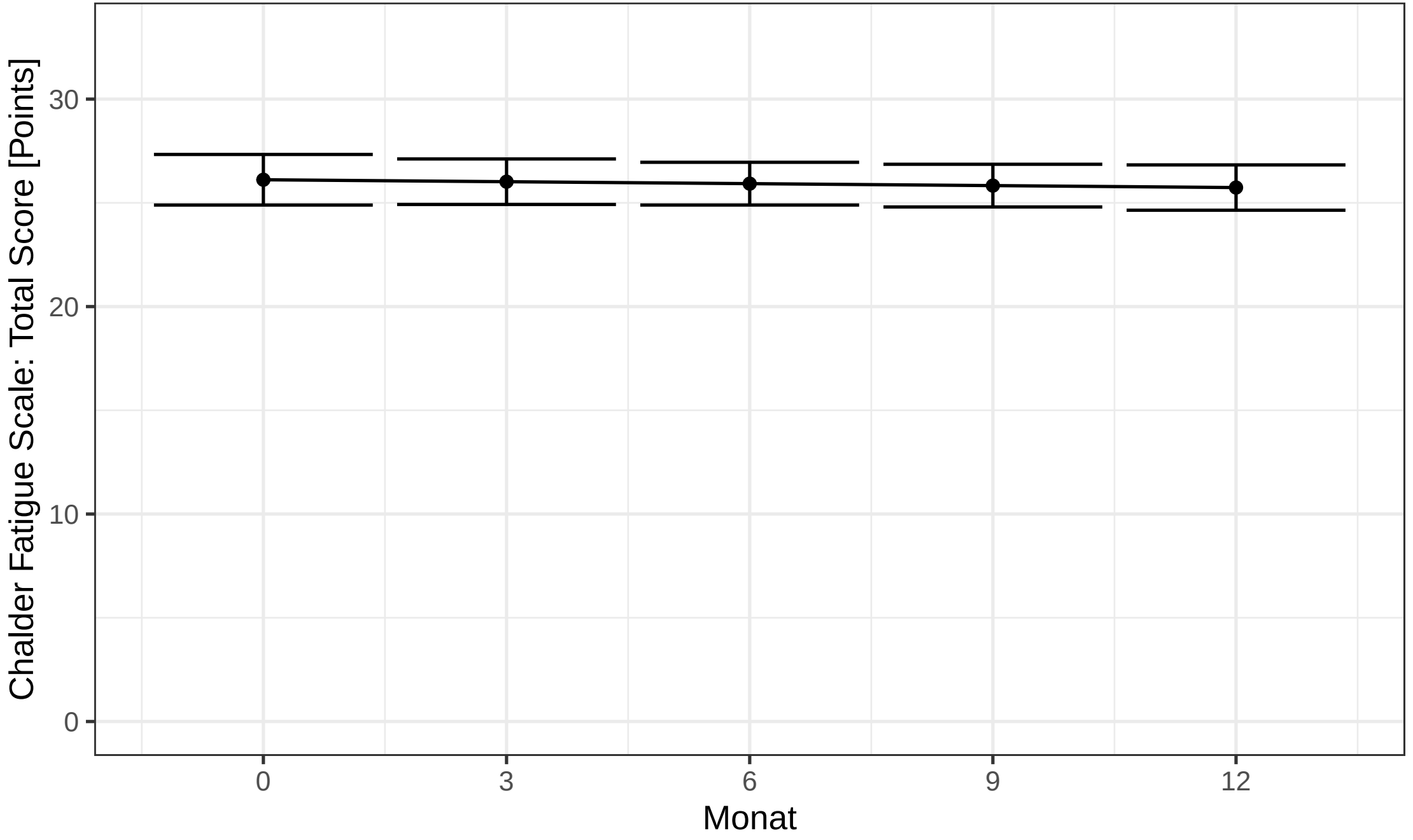
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

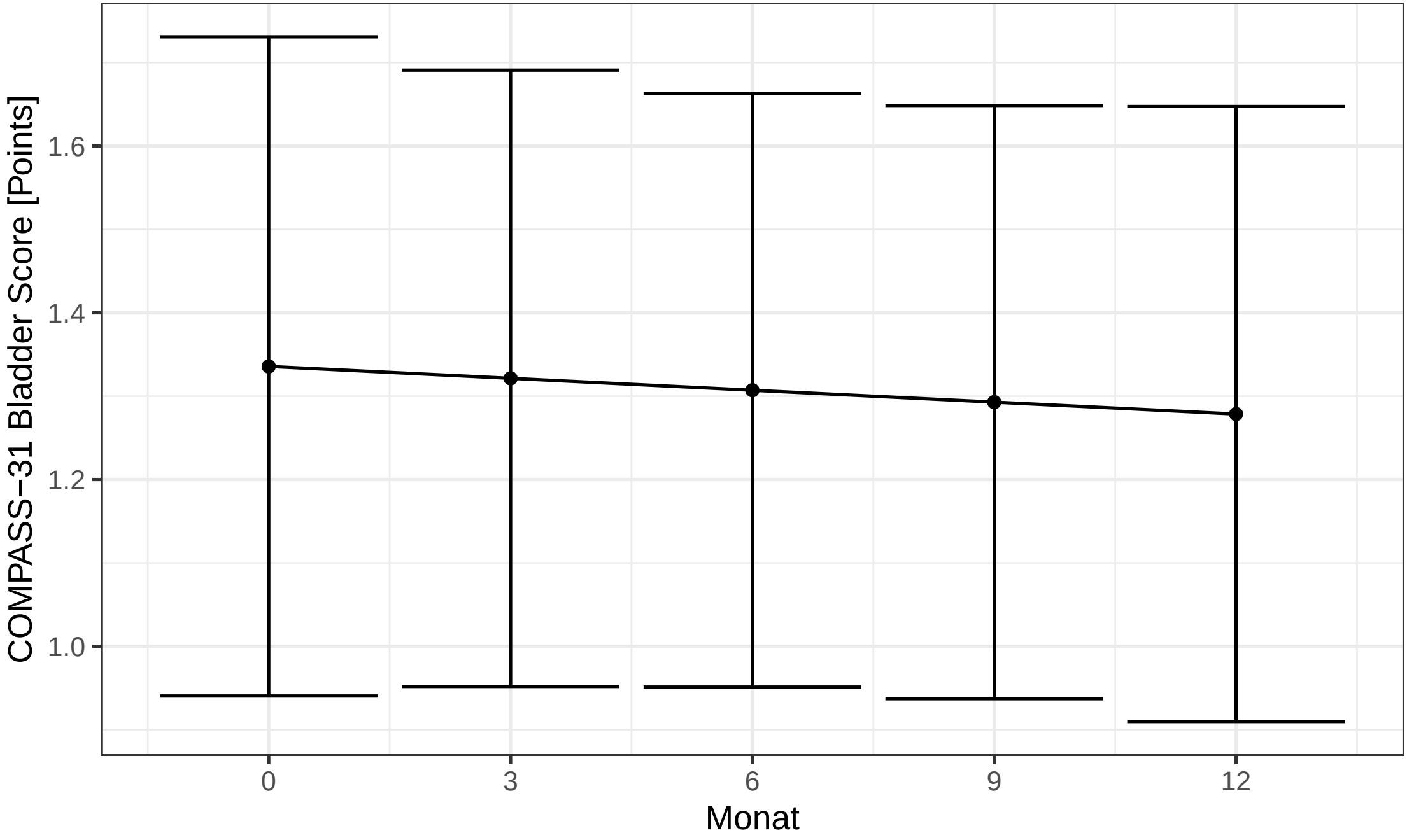
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

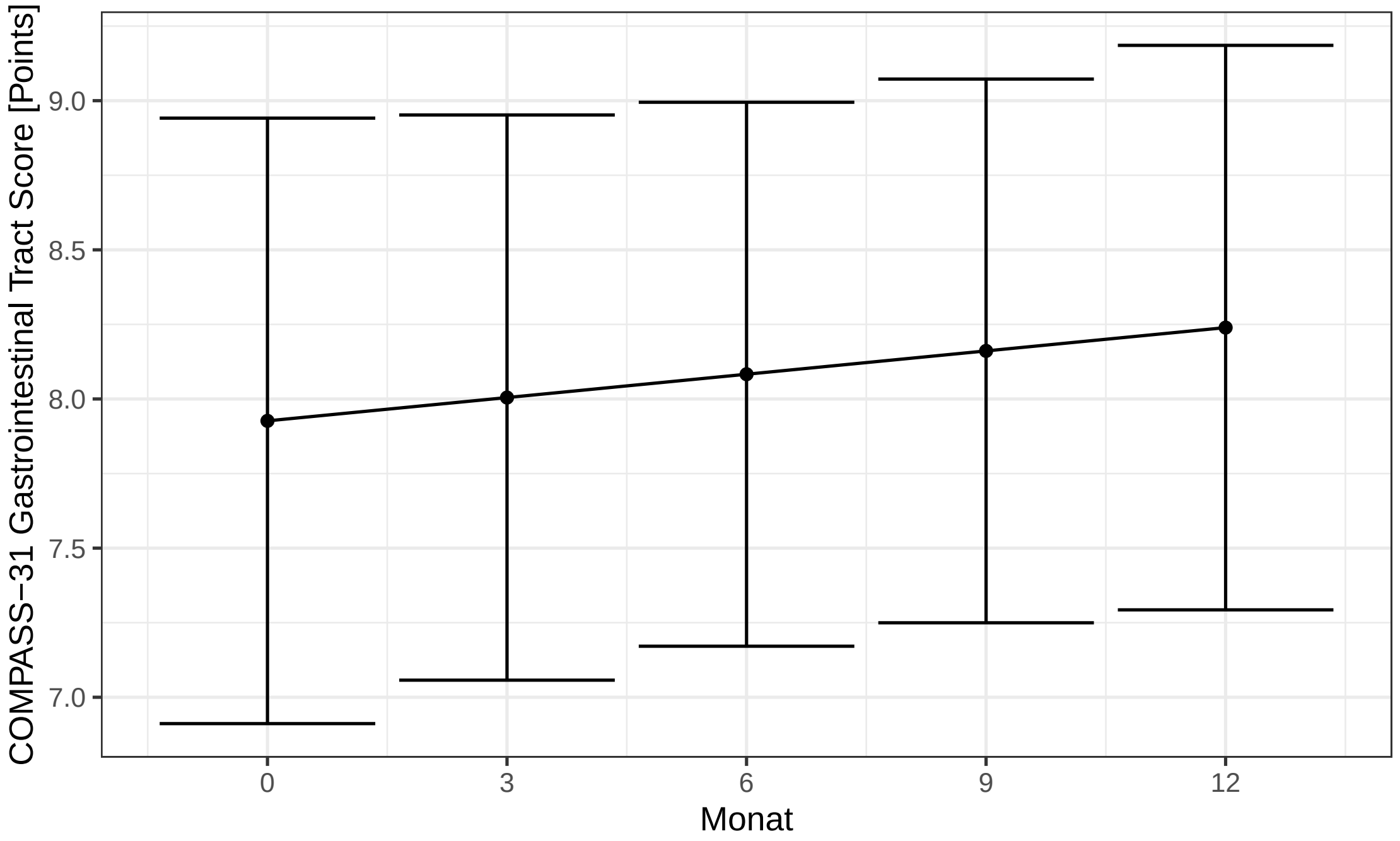
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

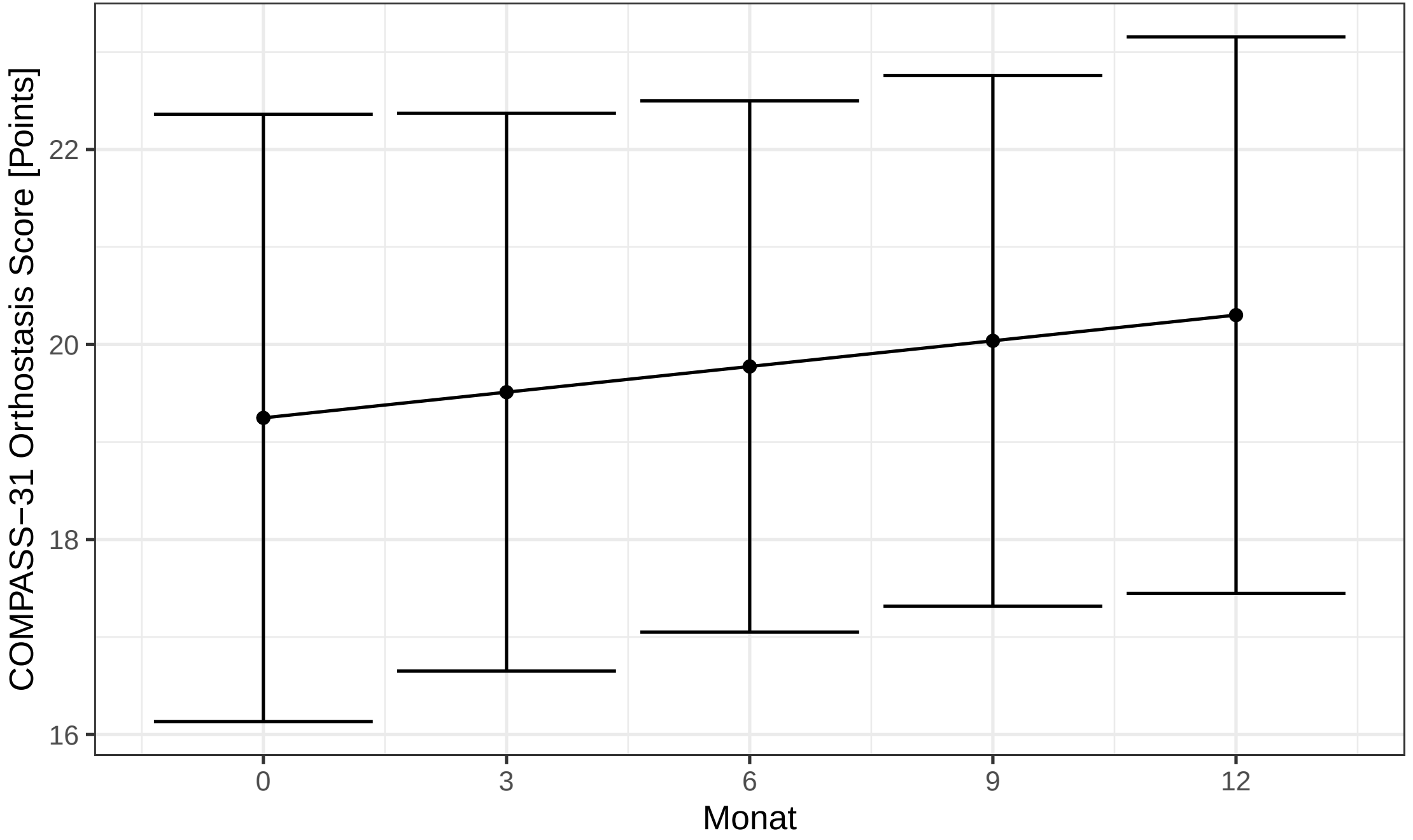
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

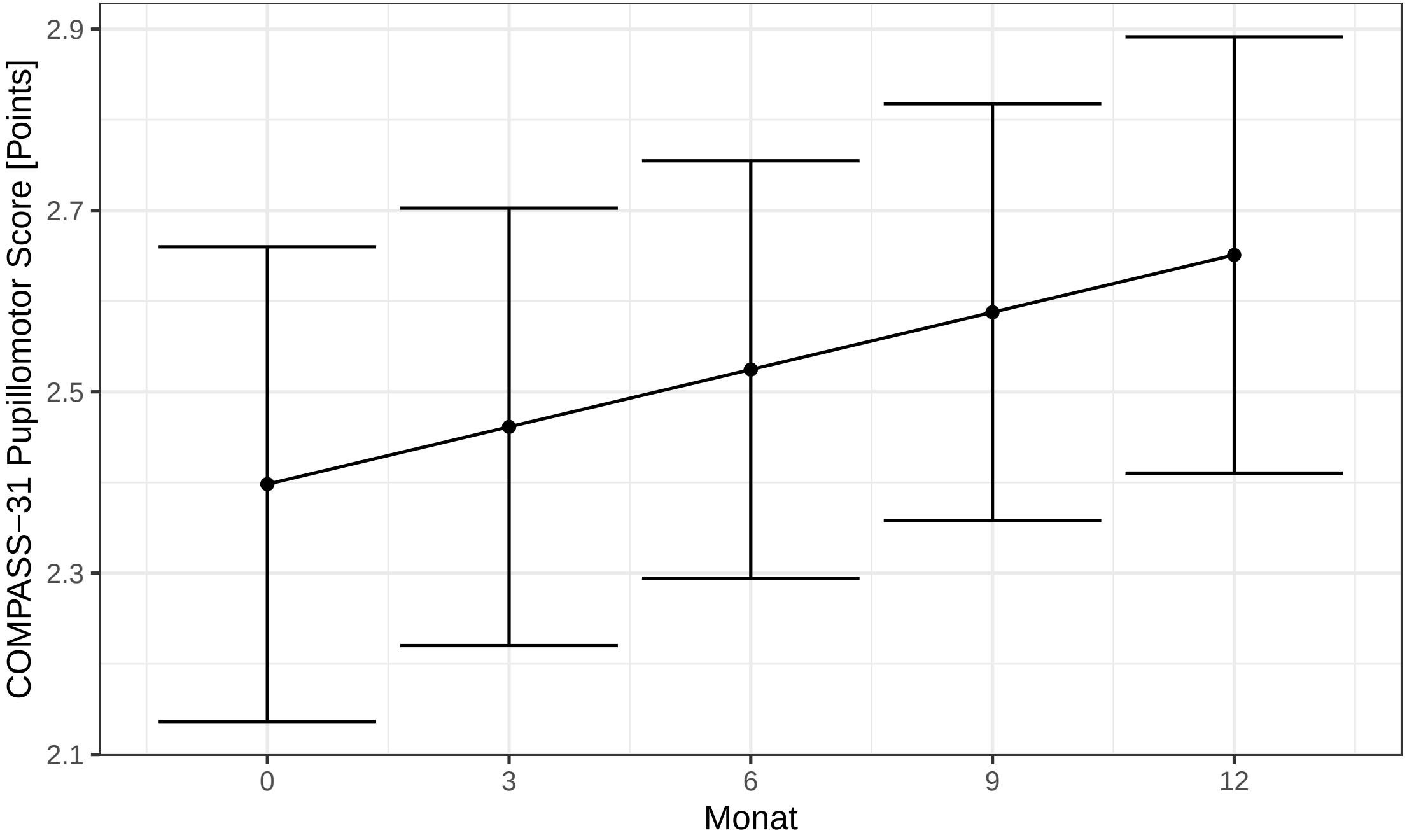
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

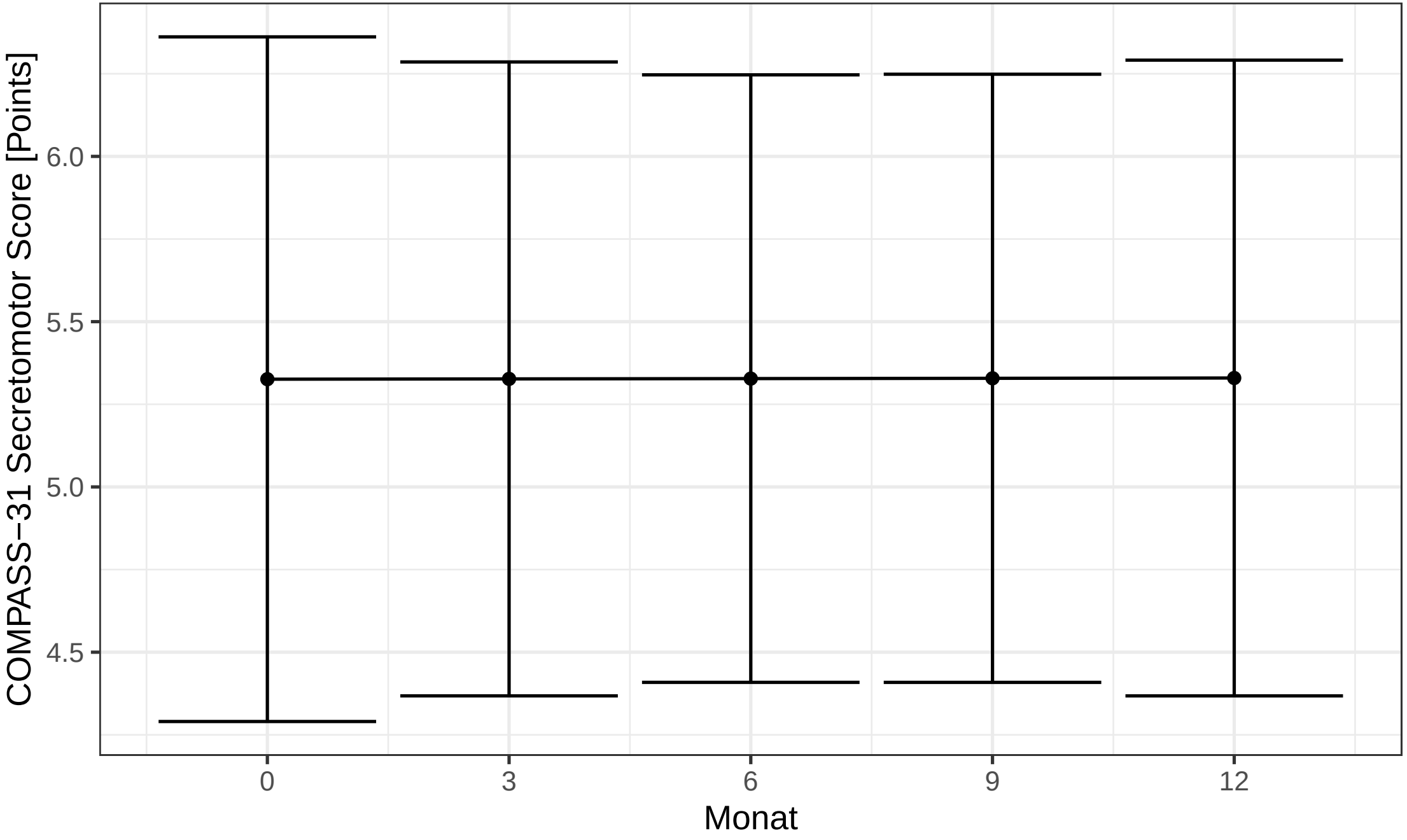
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

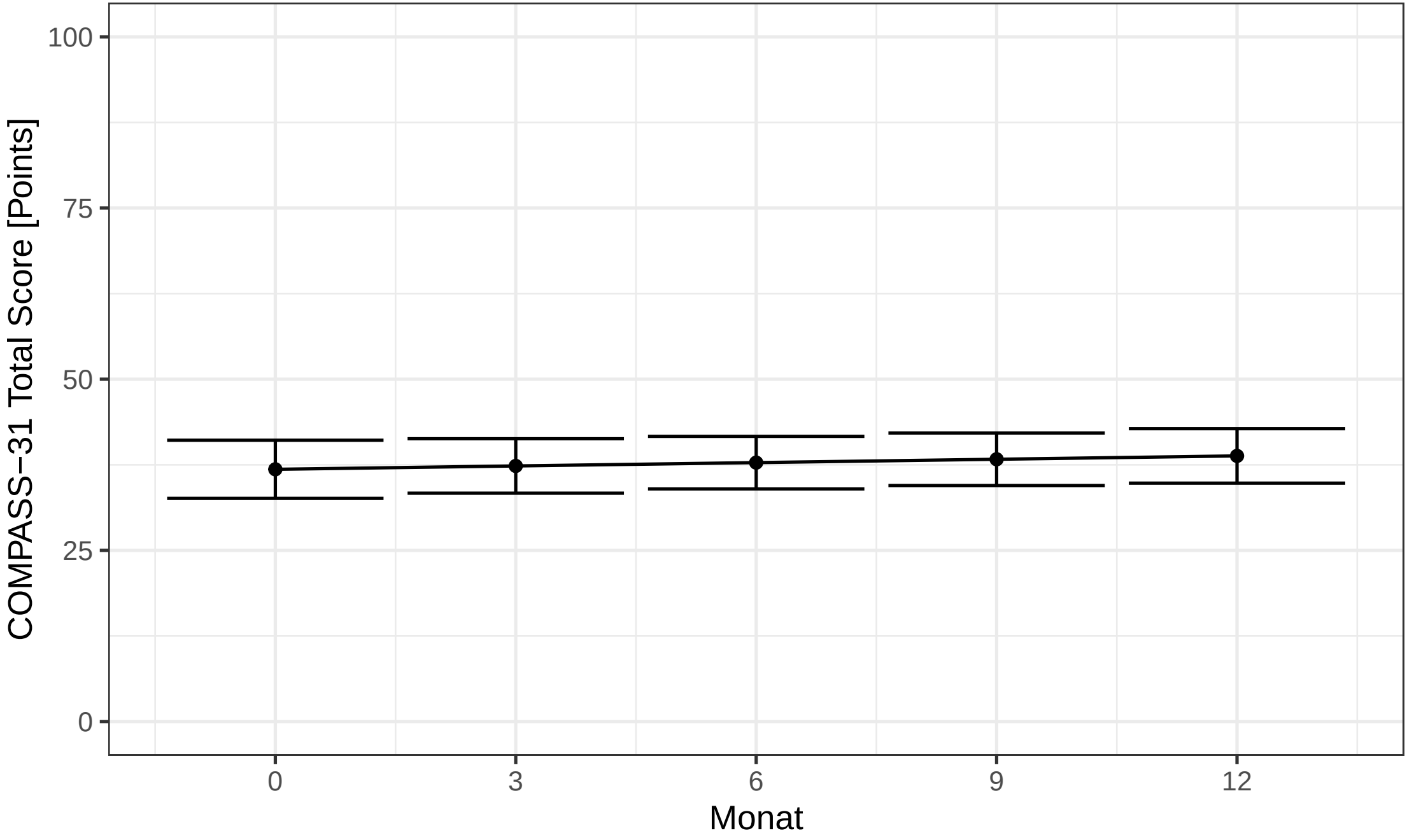
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

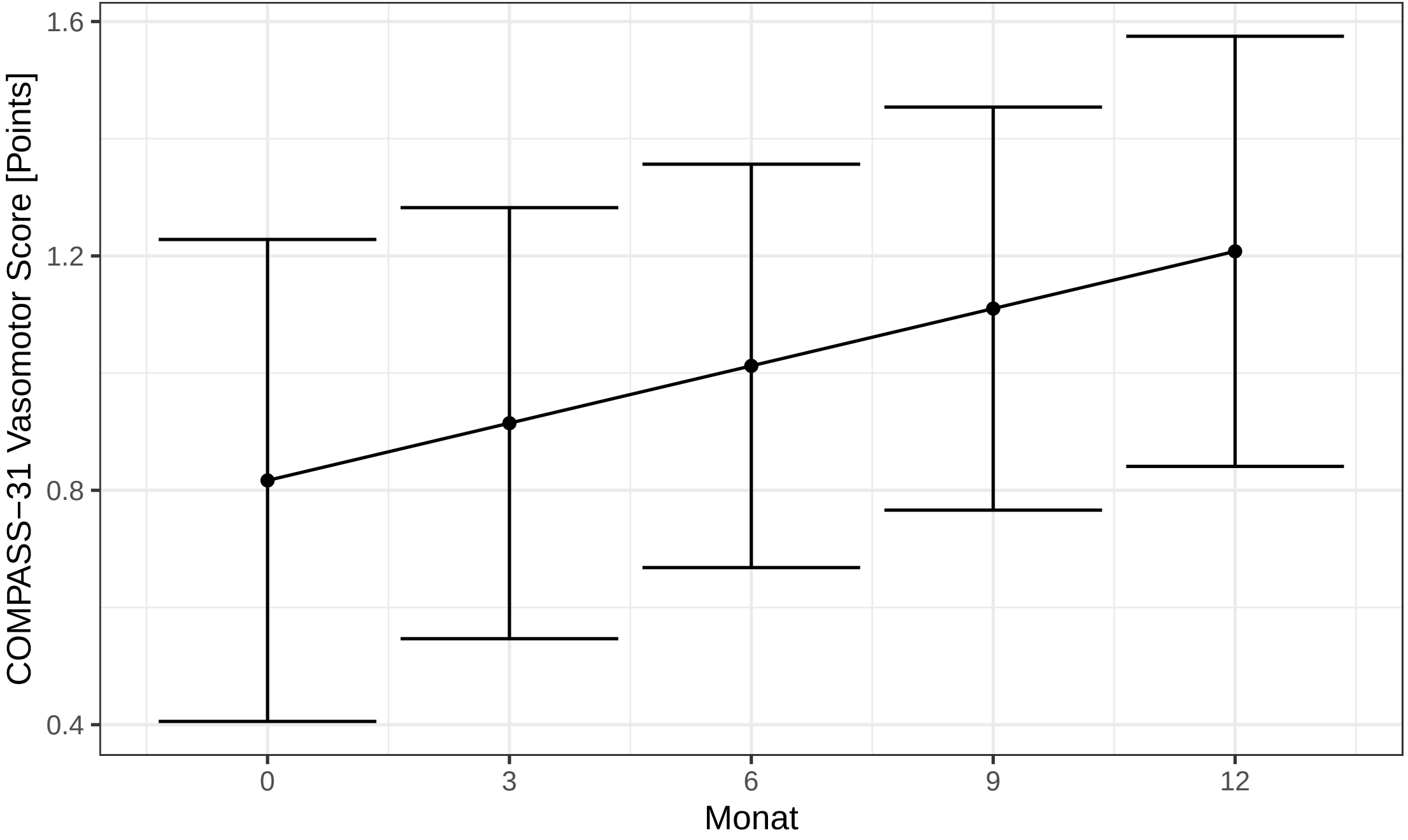
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

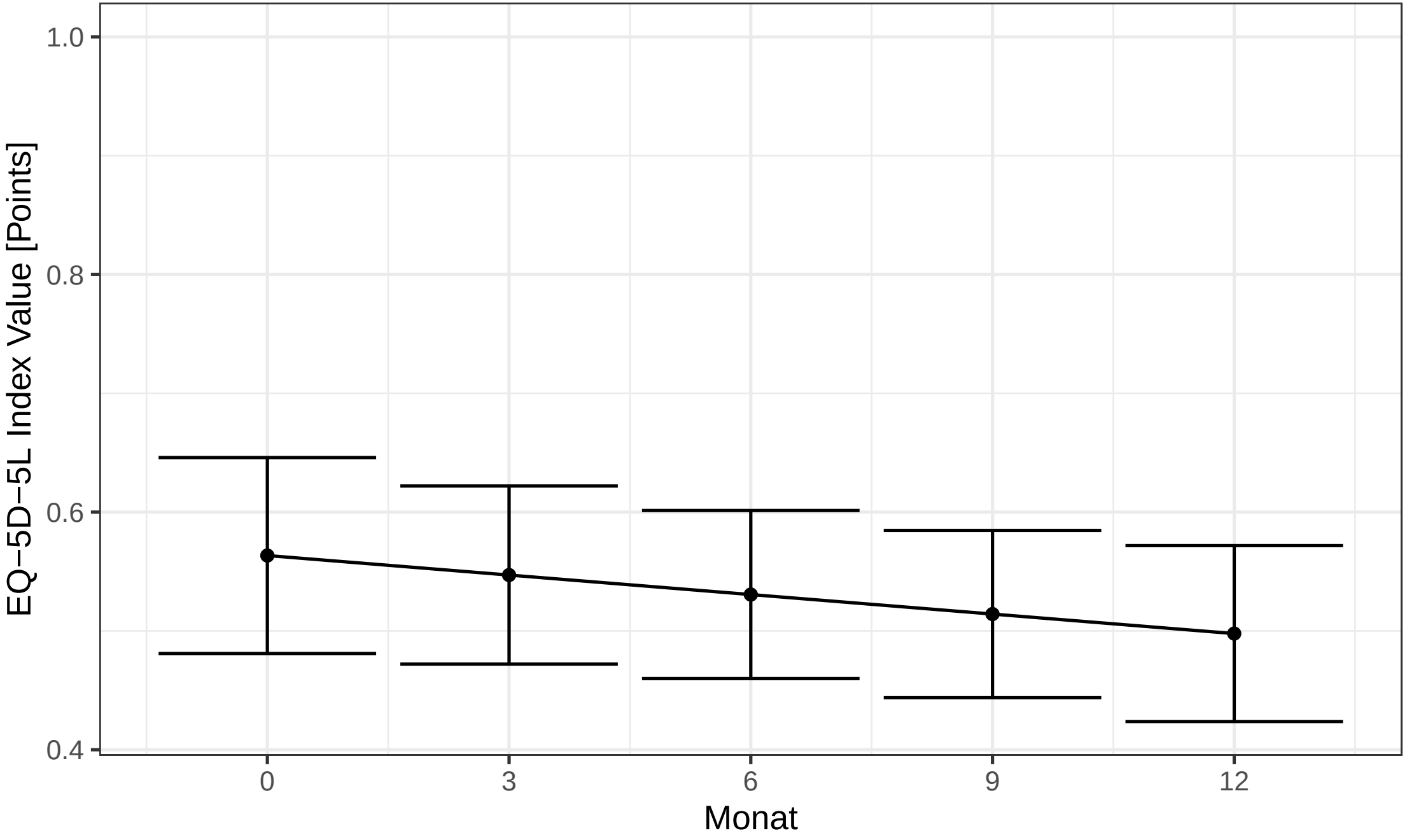
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

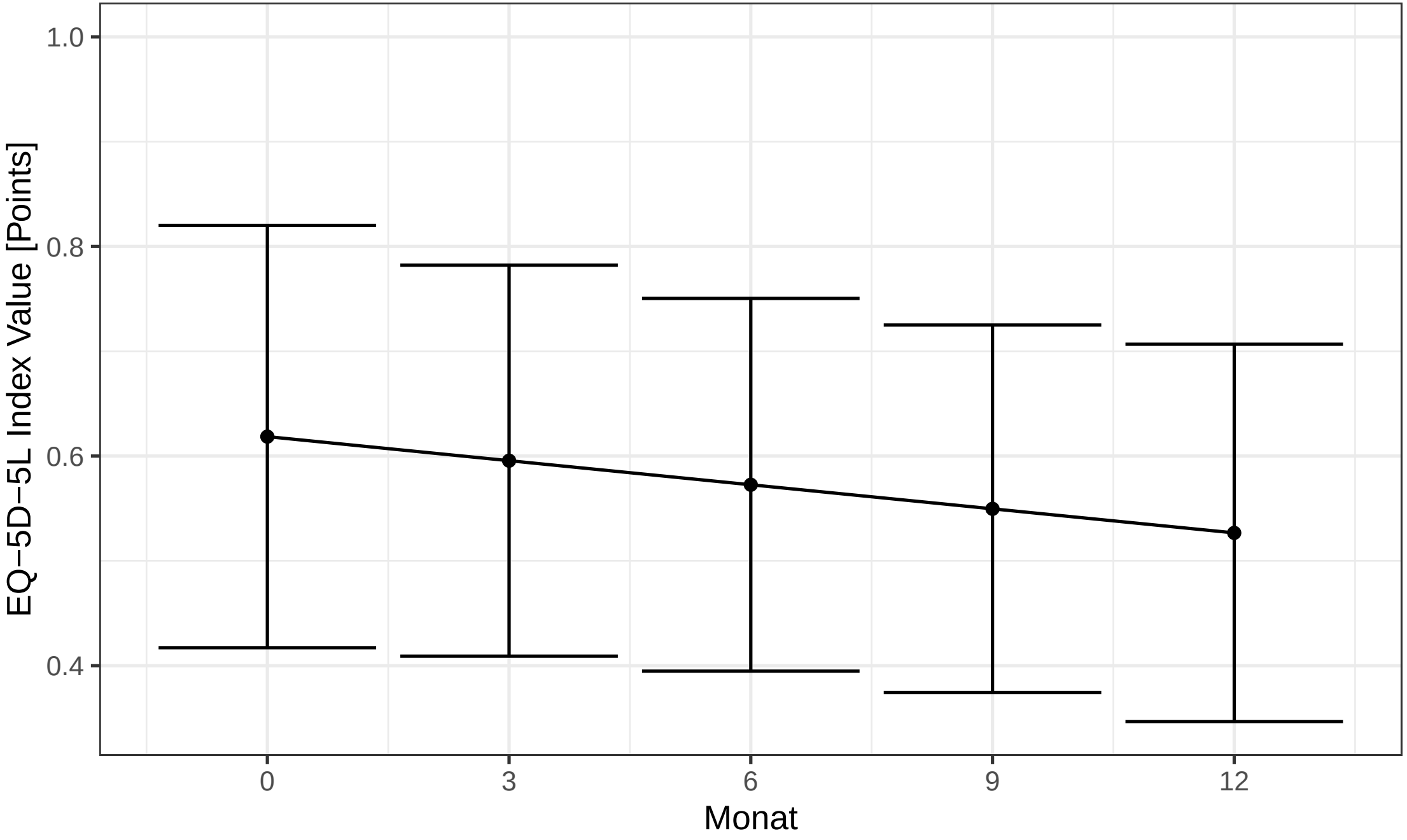
Adjustierter Mittelwert



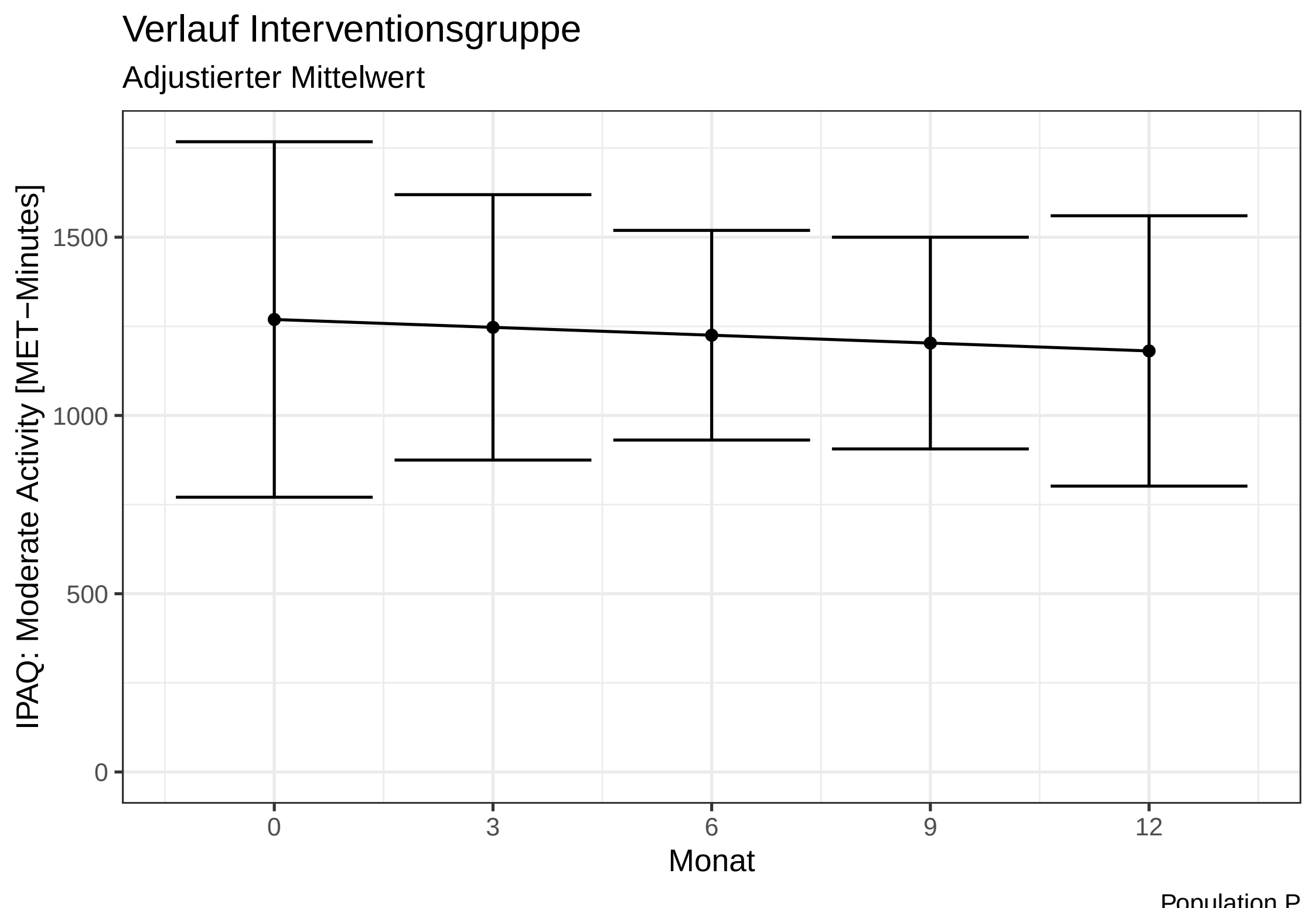
Population P

Verlauf Interventionsgruppe

Adjustierter Mittelwert

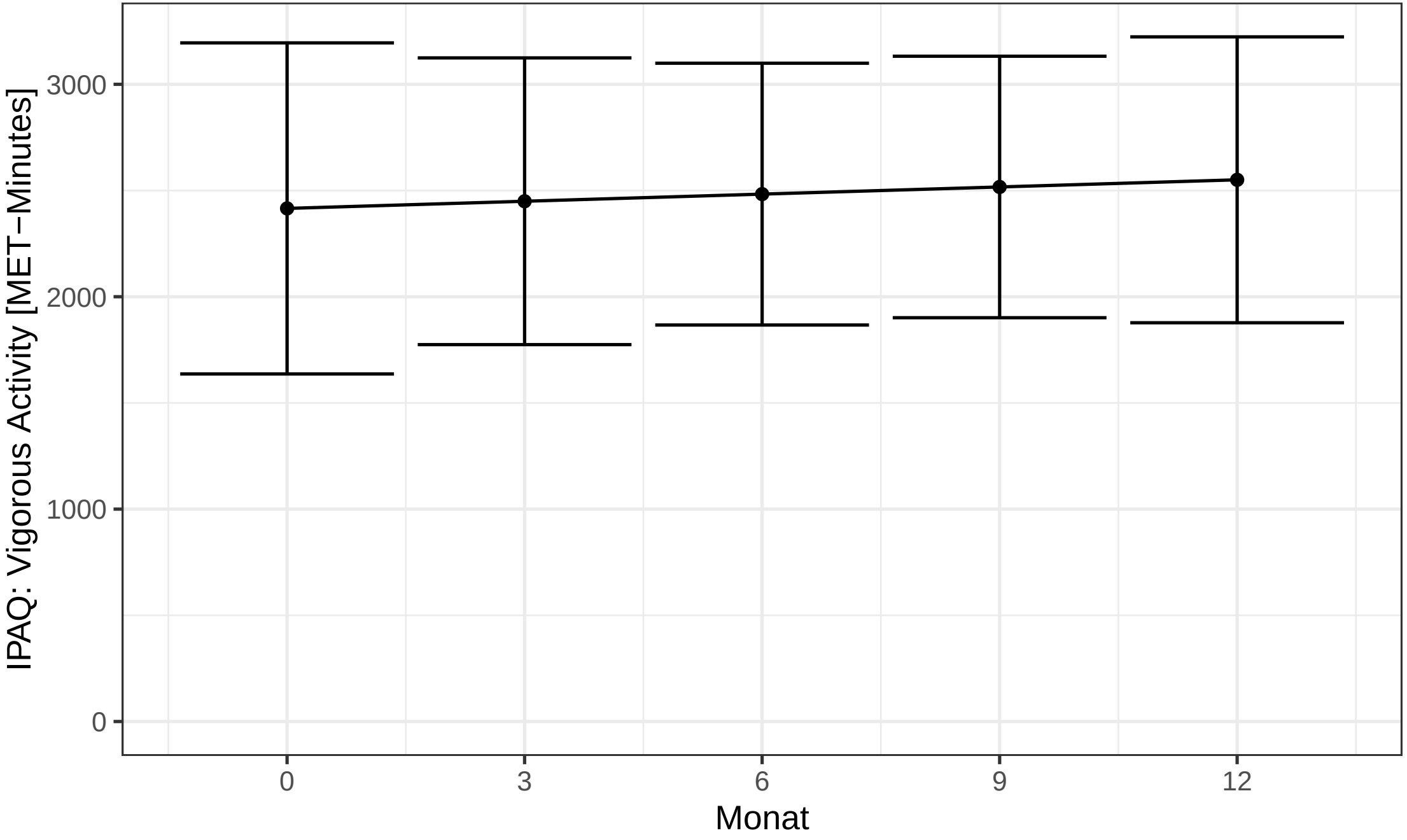


Population P


Verlauf Interventionsgruppe
Adjustierter Mittelwert
IPAQ: Moderate Activity [MET−Minutes]
1500
1000
500
0
0
3
6
9
12
Monat
Population P

Verlauf Interventionsgruppe

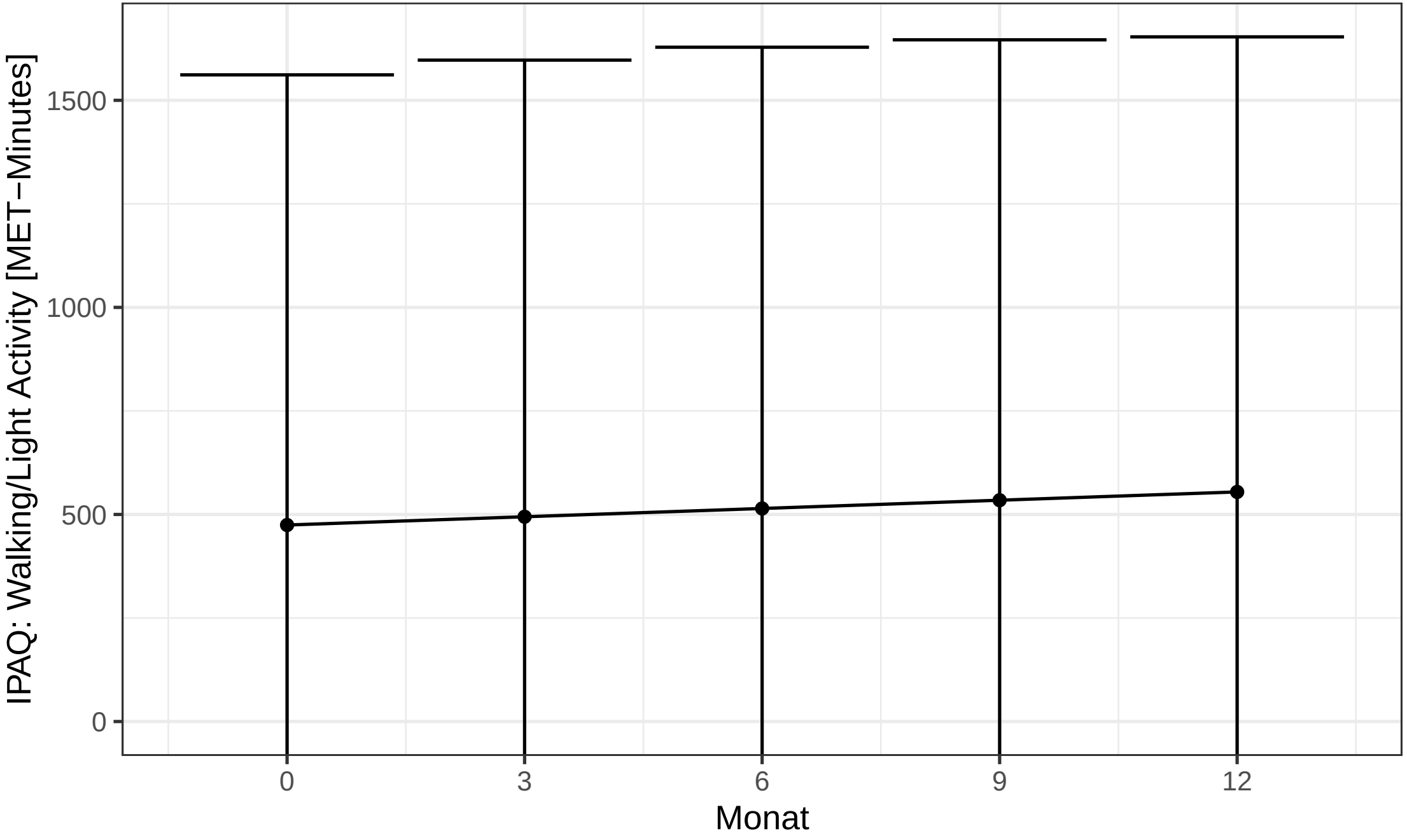
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

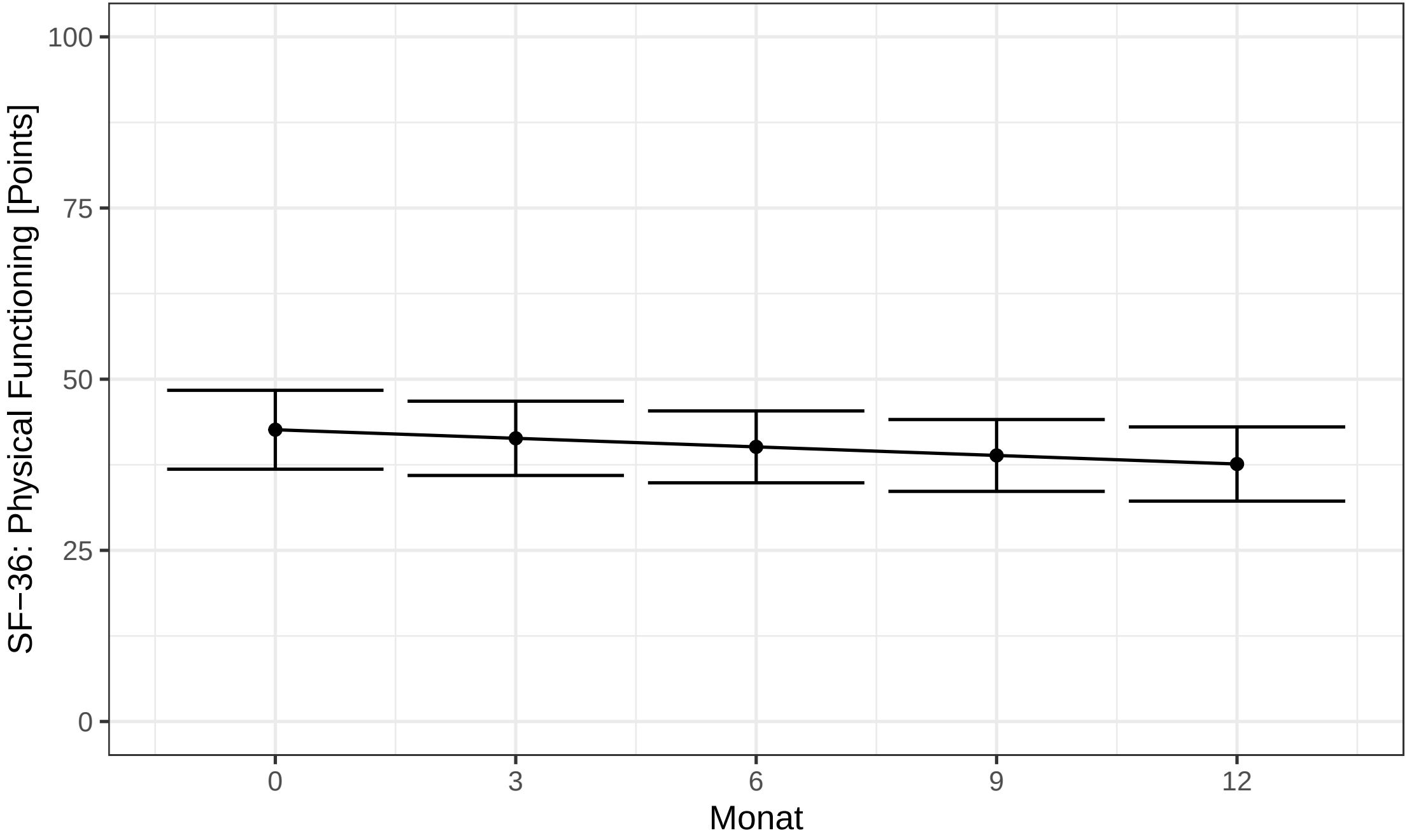
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

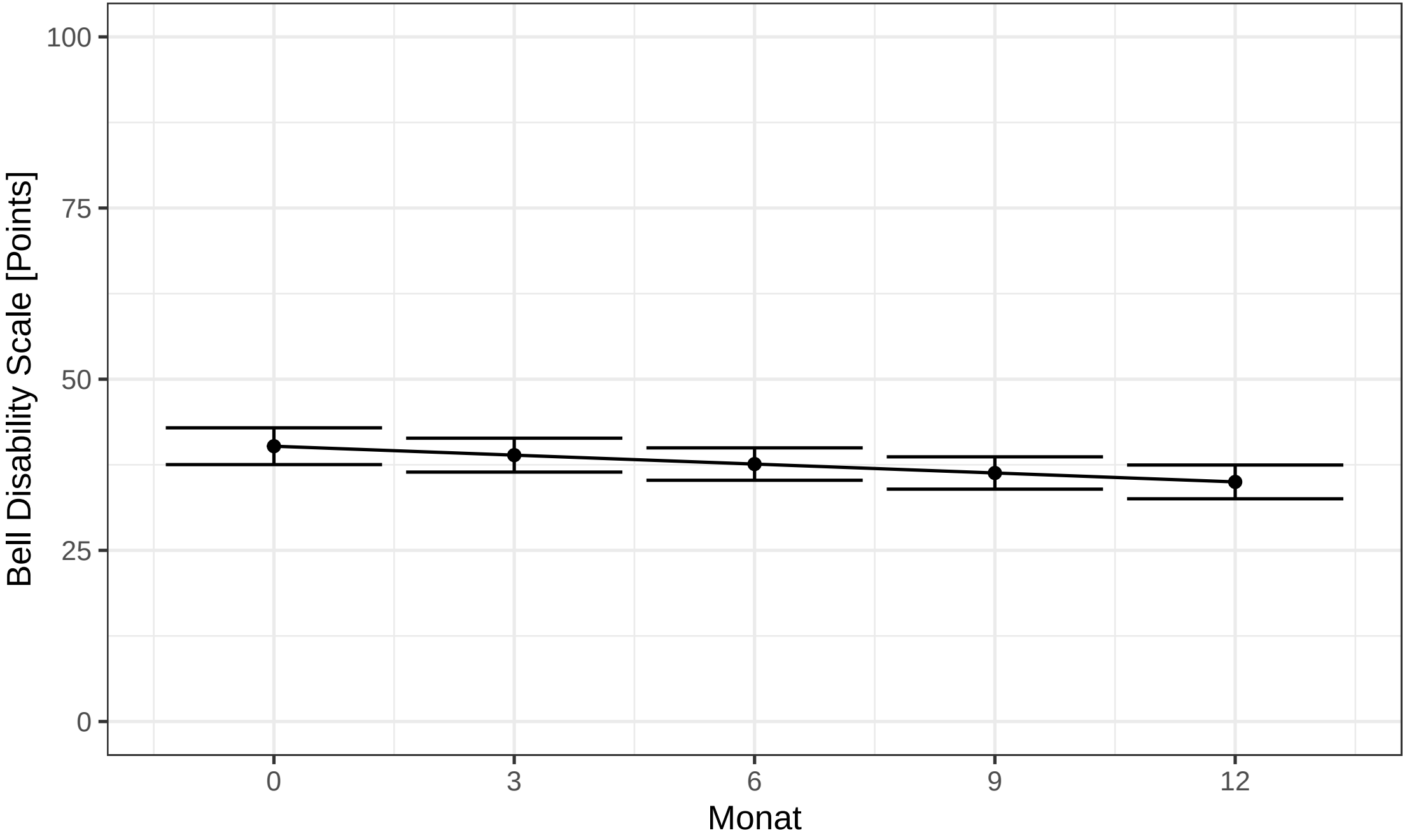
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

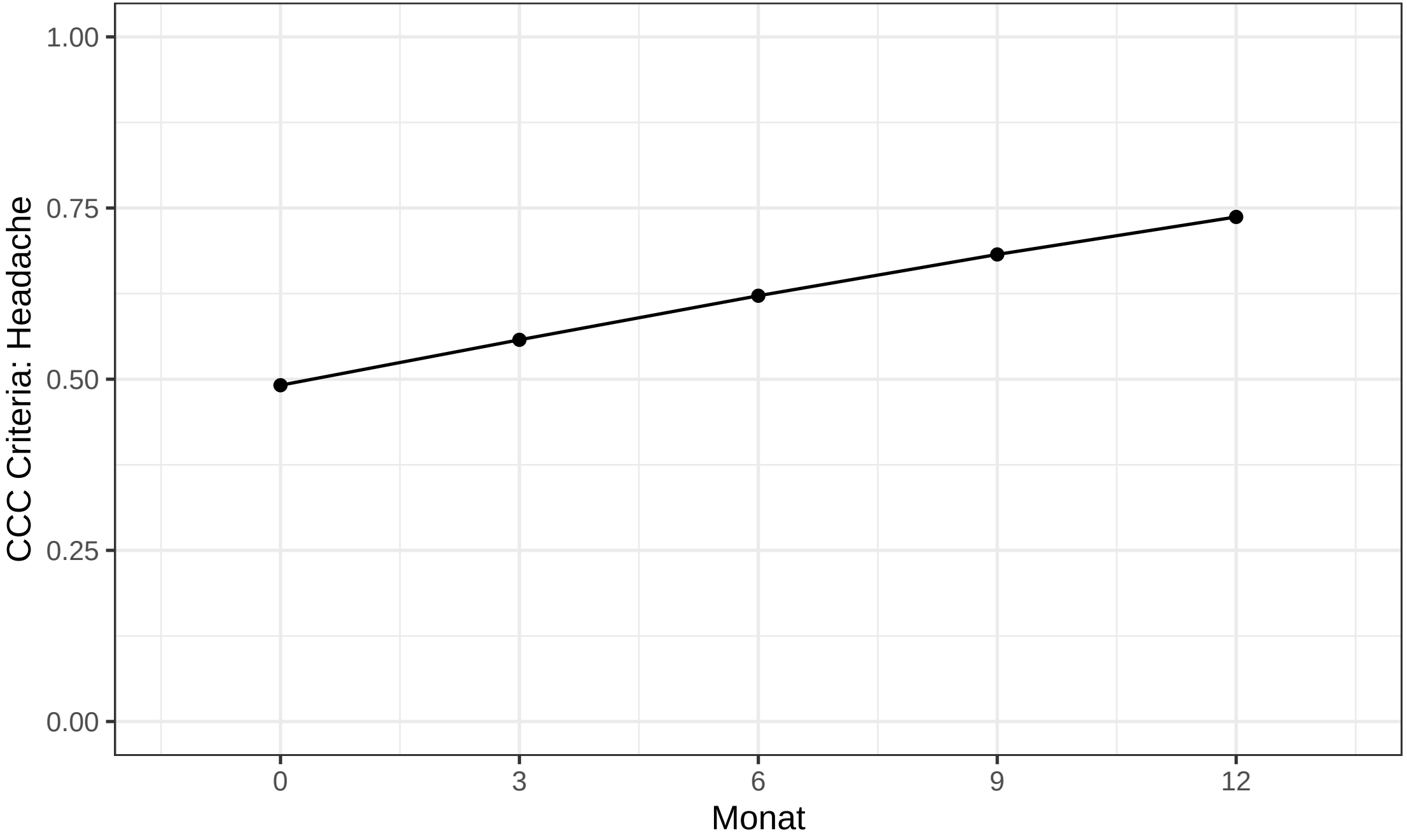
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

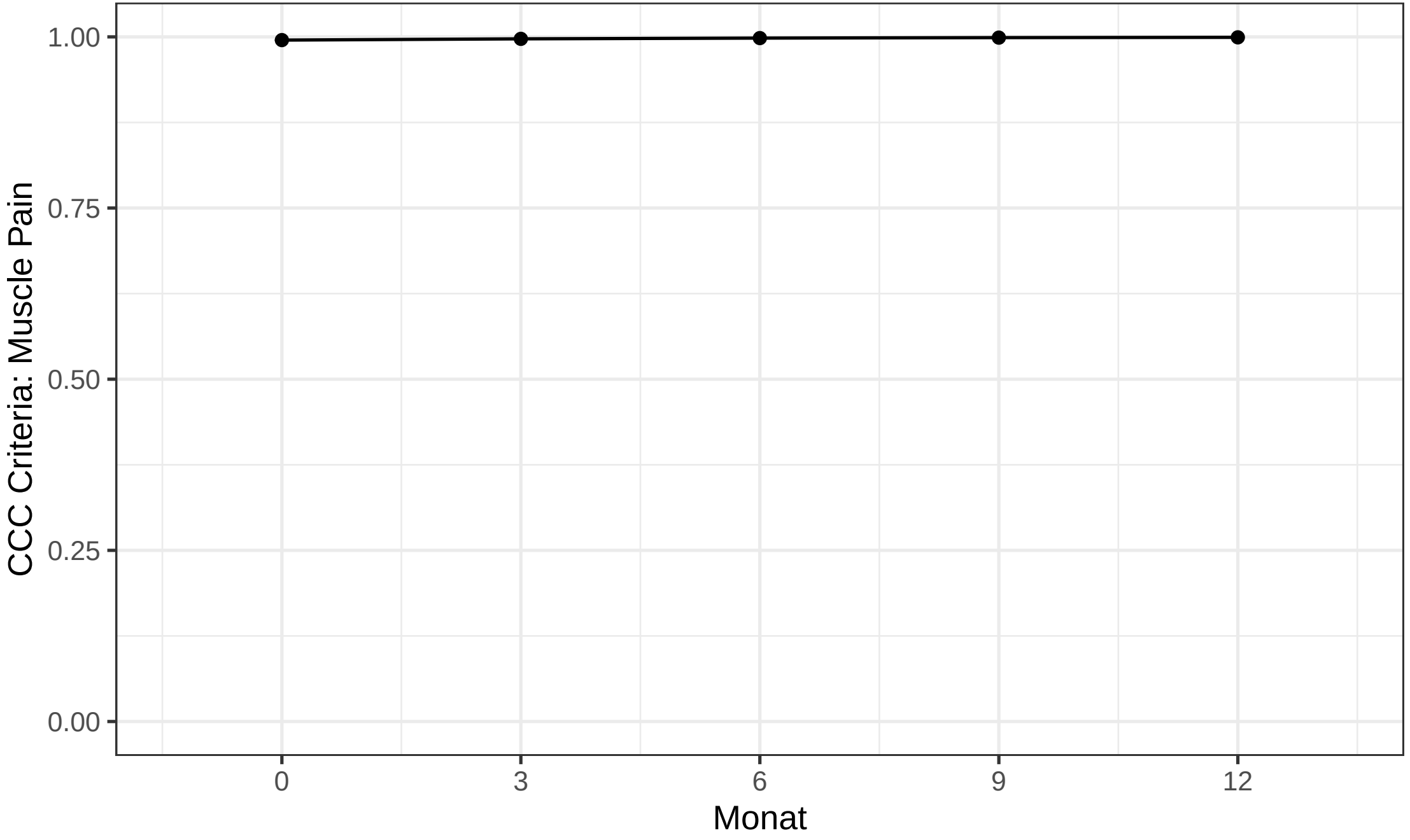
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

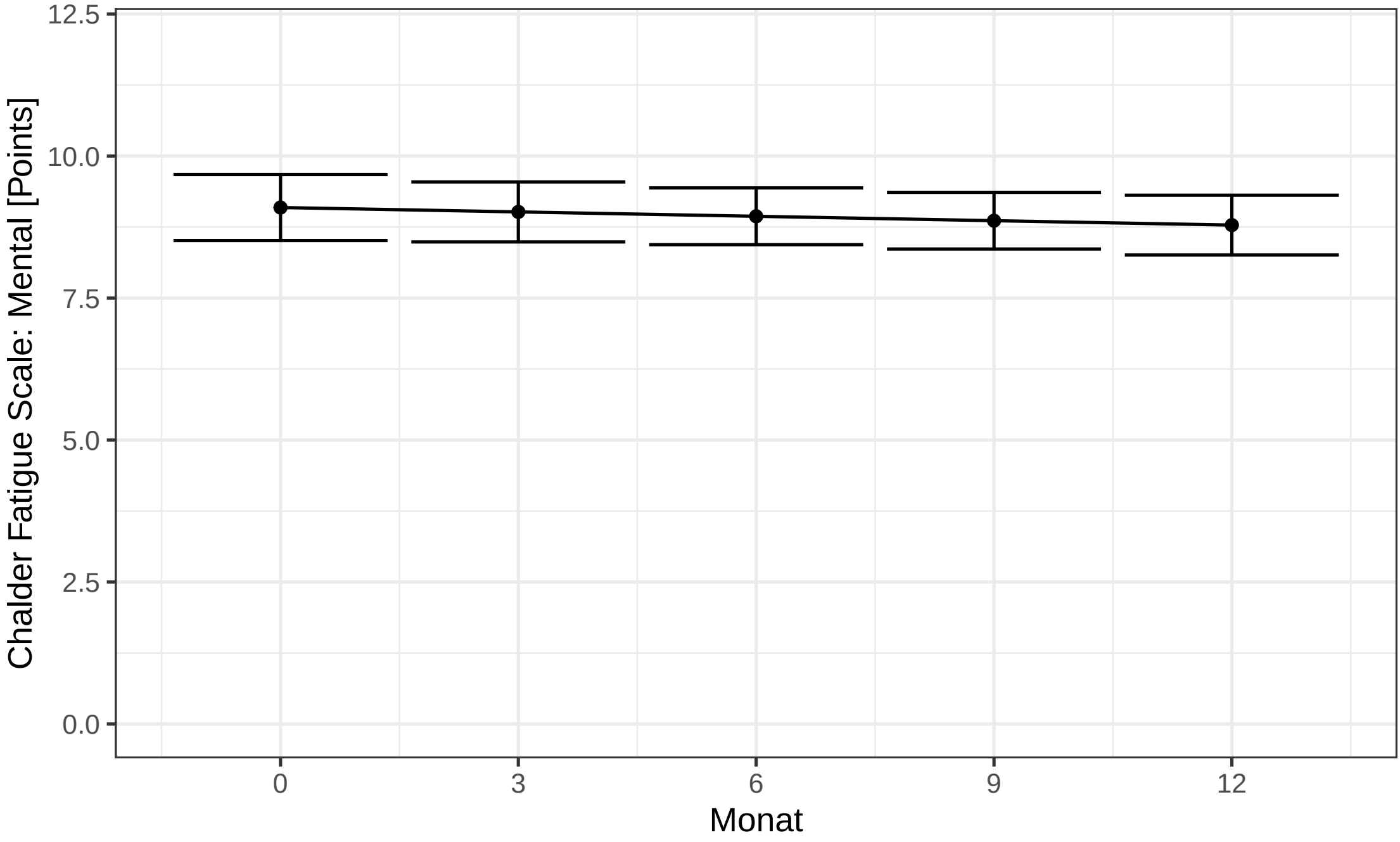
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

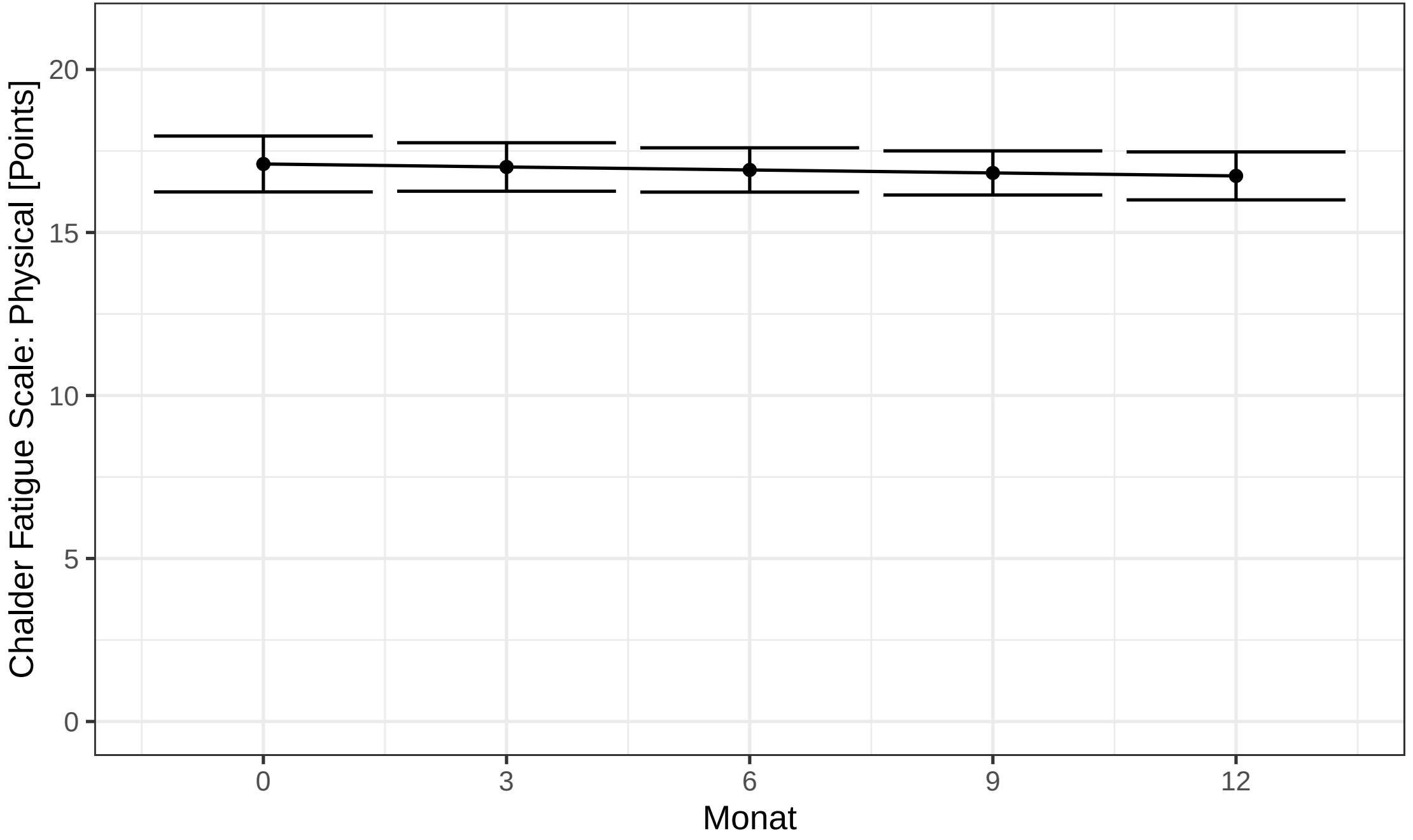
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

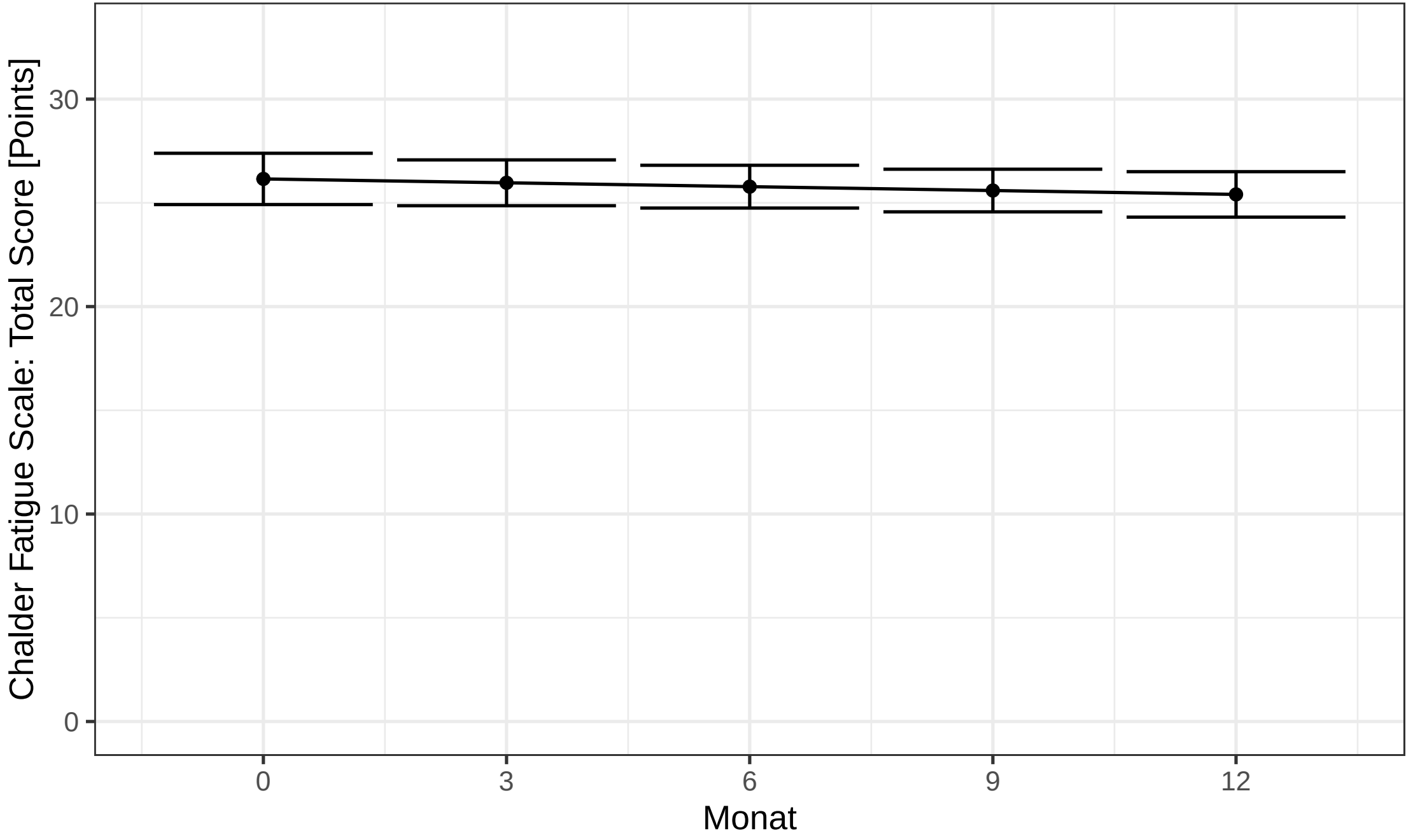
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

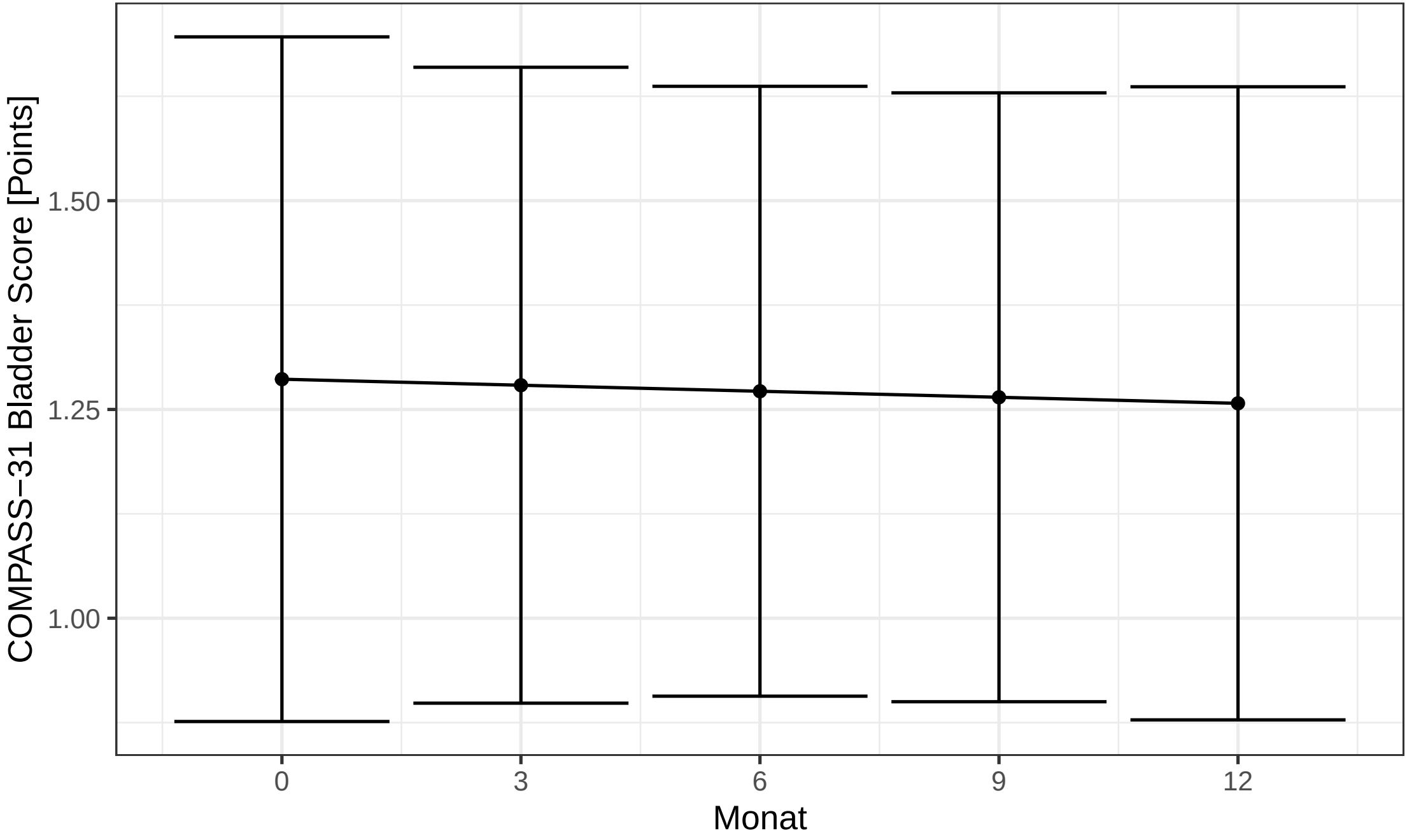
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

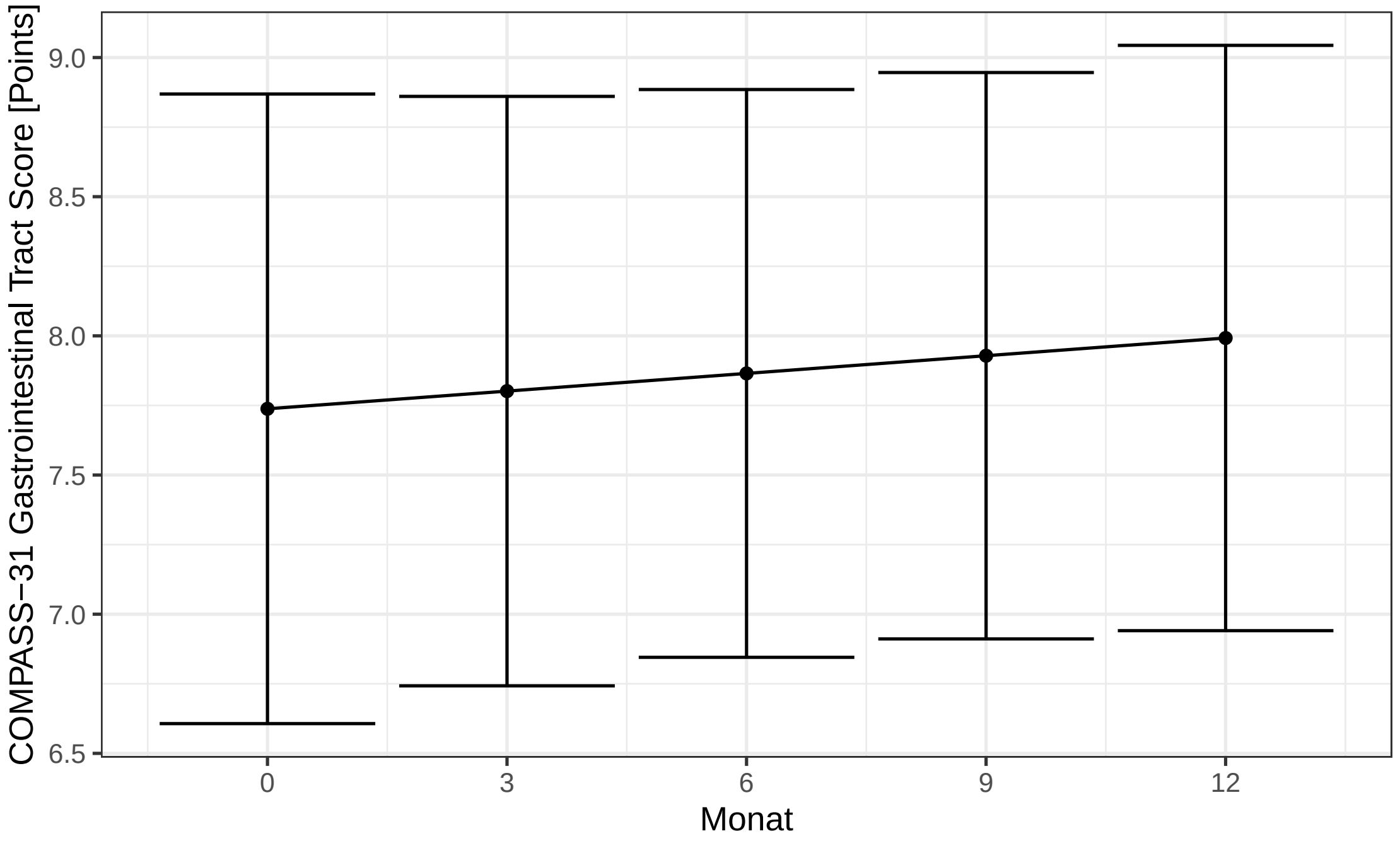
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

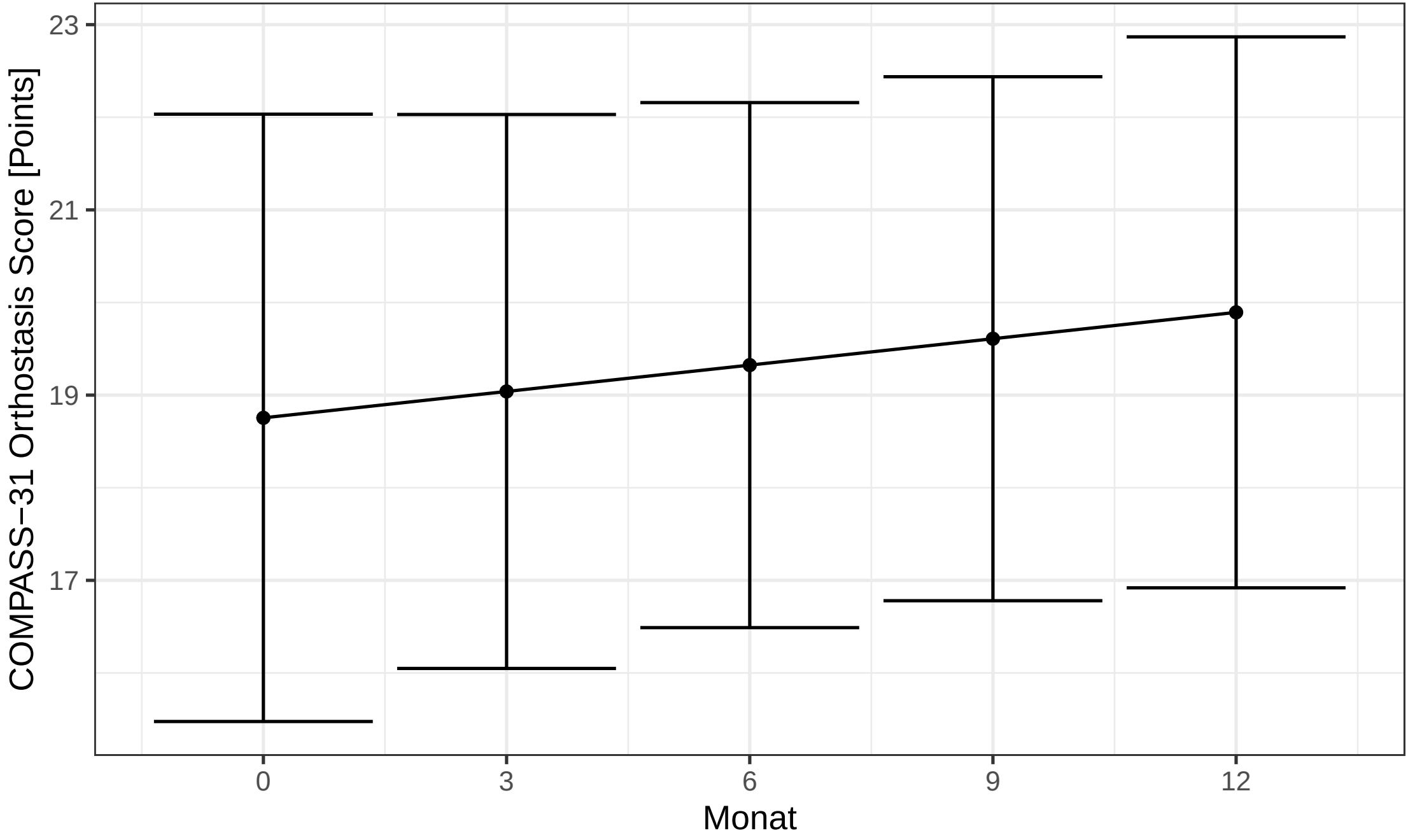
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

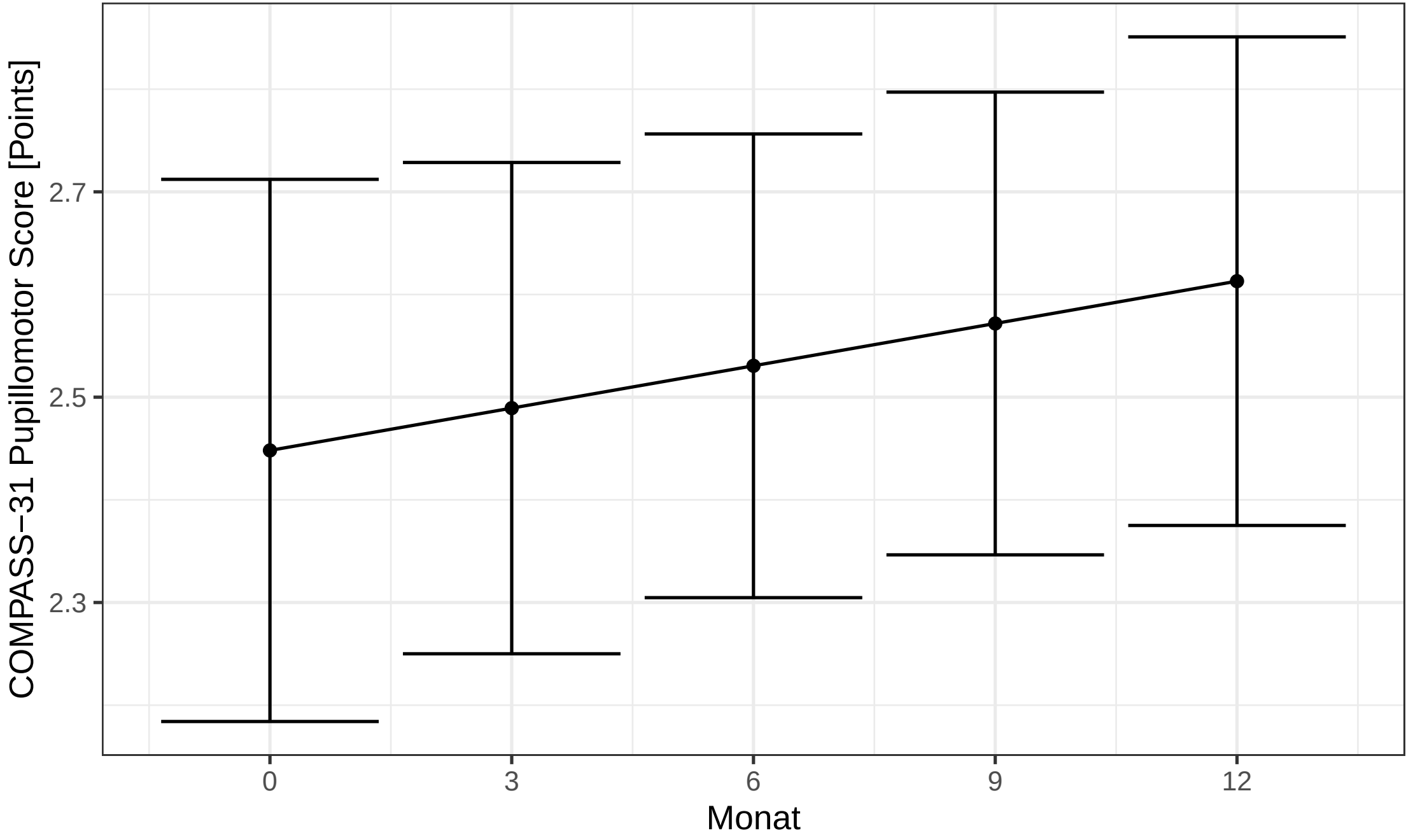
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

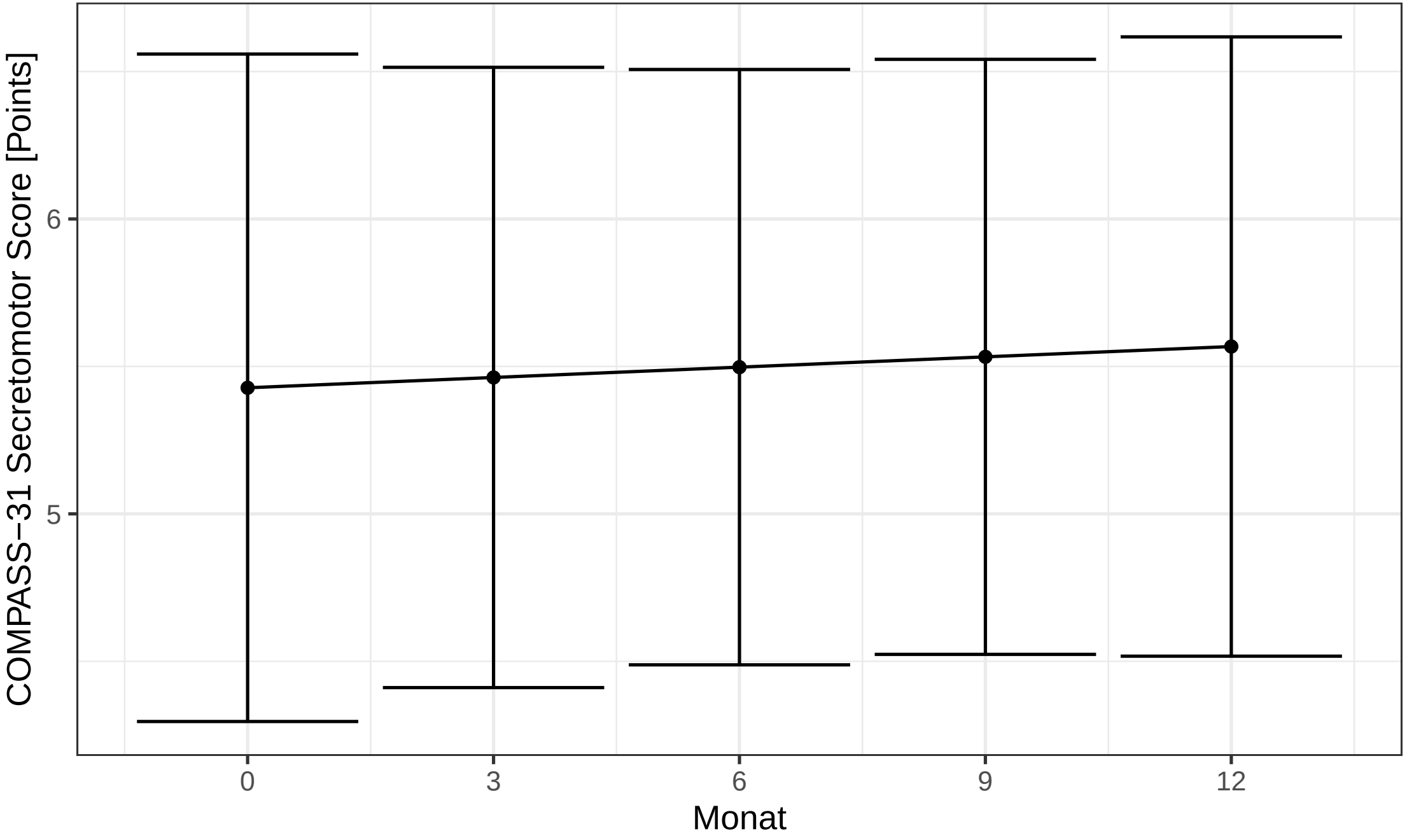
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

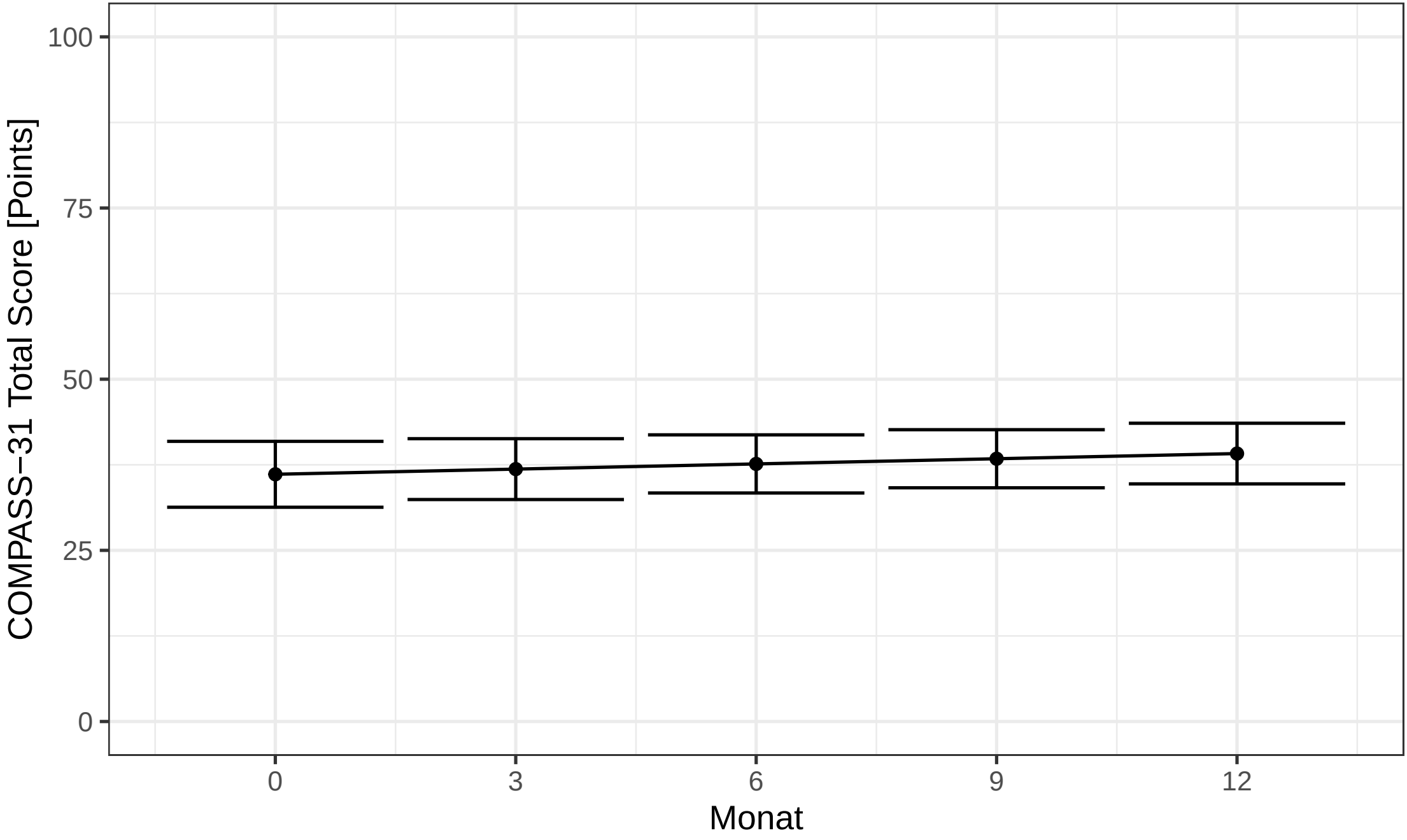
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

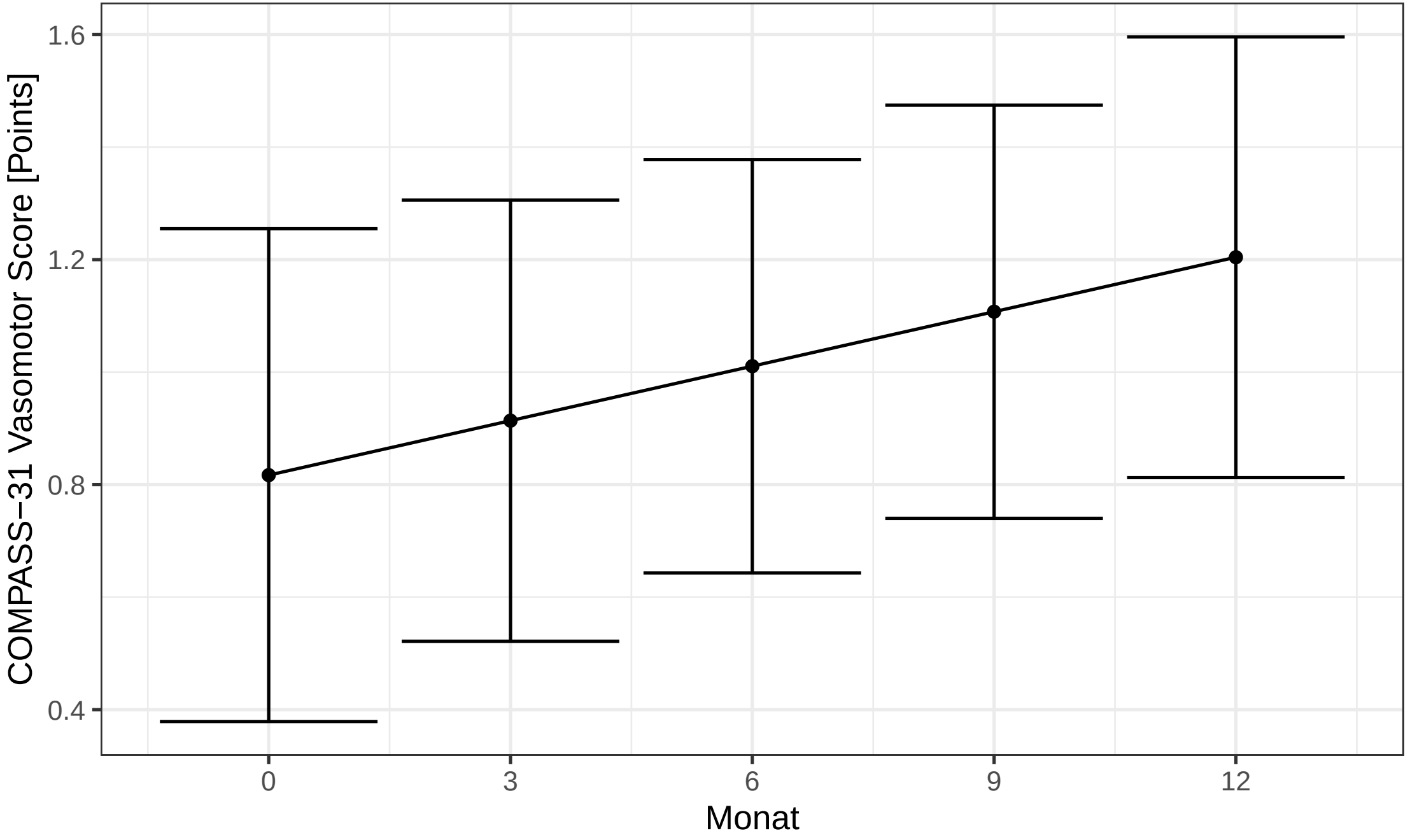
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

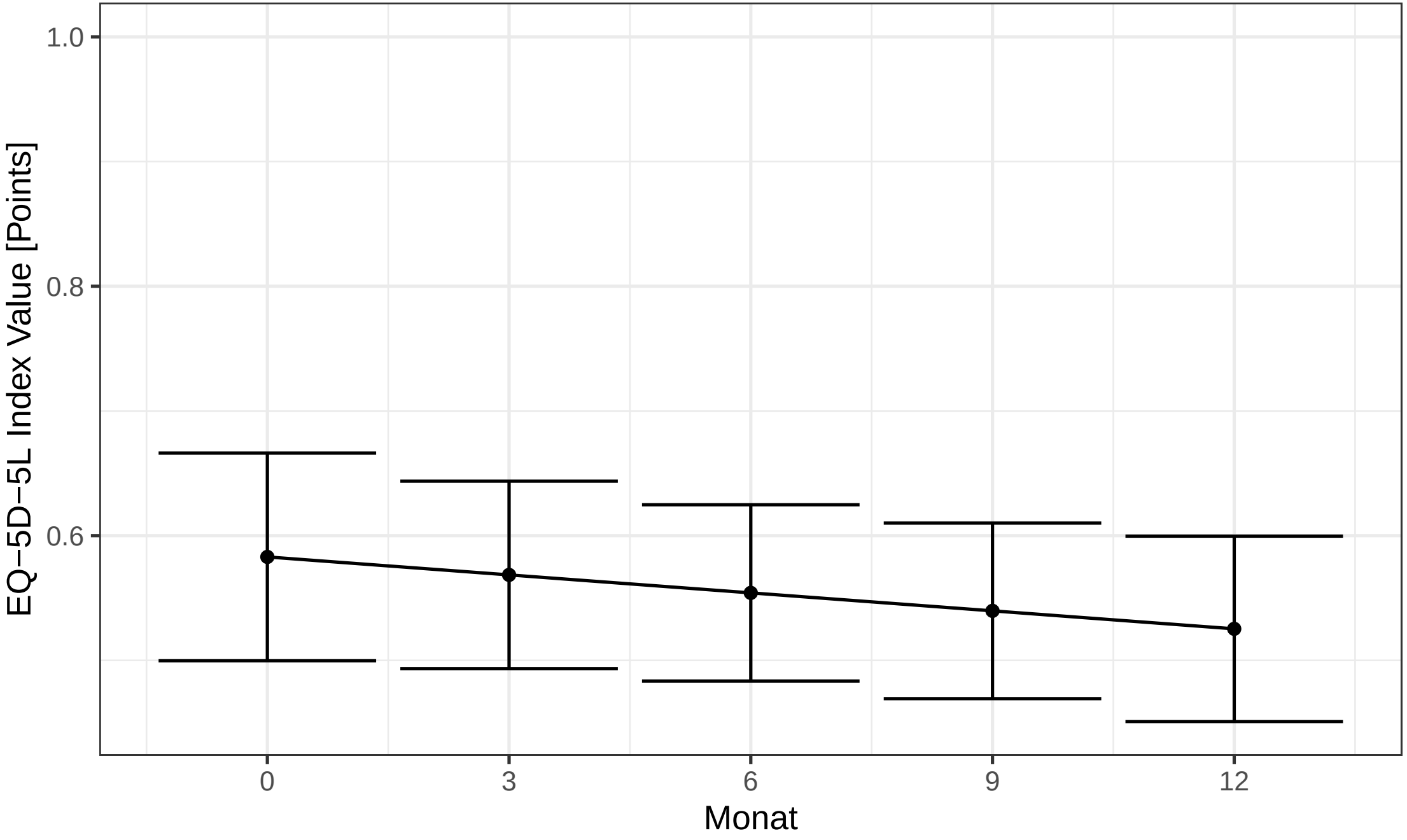
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

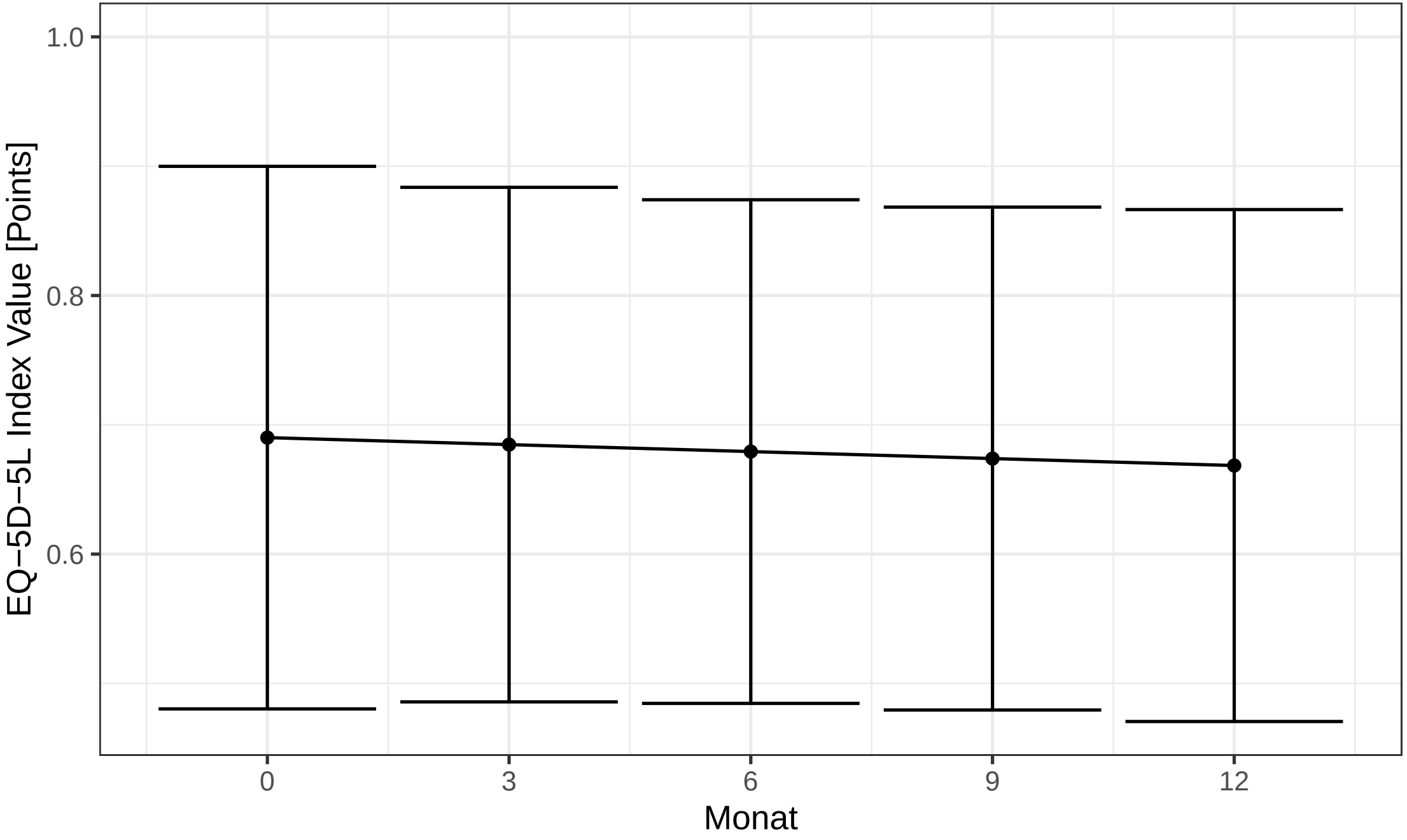
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

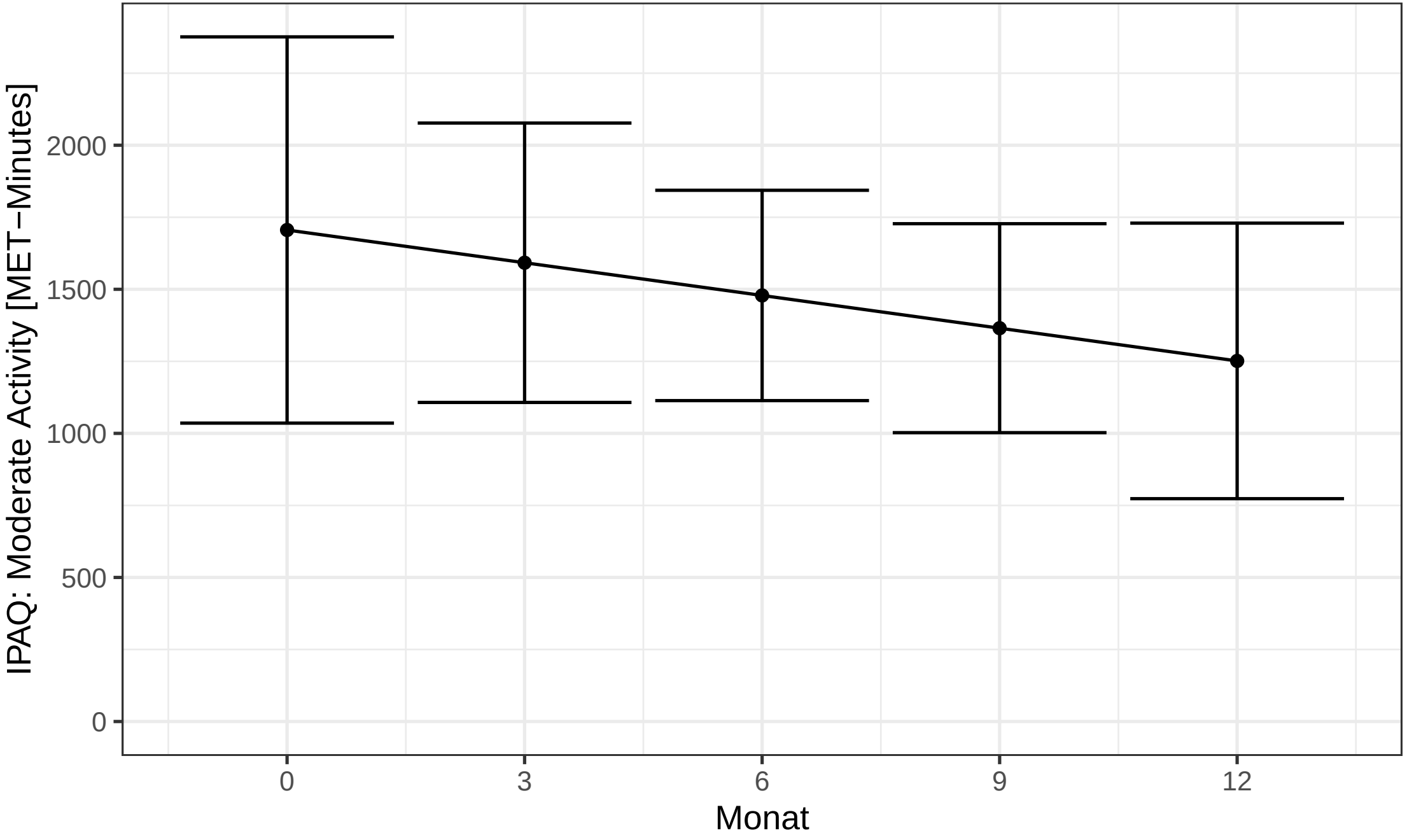
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

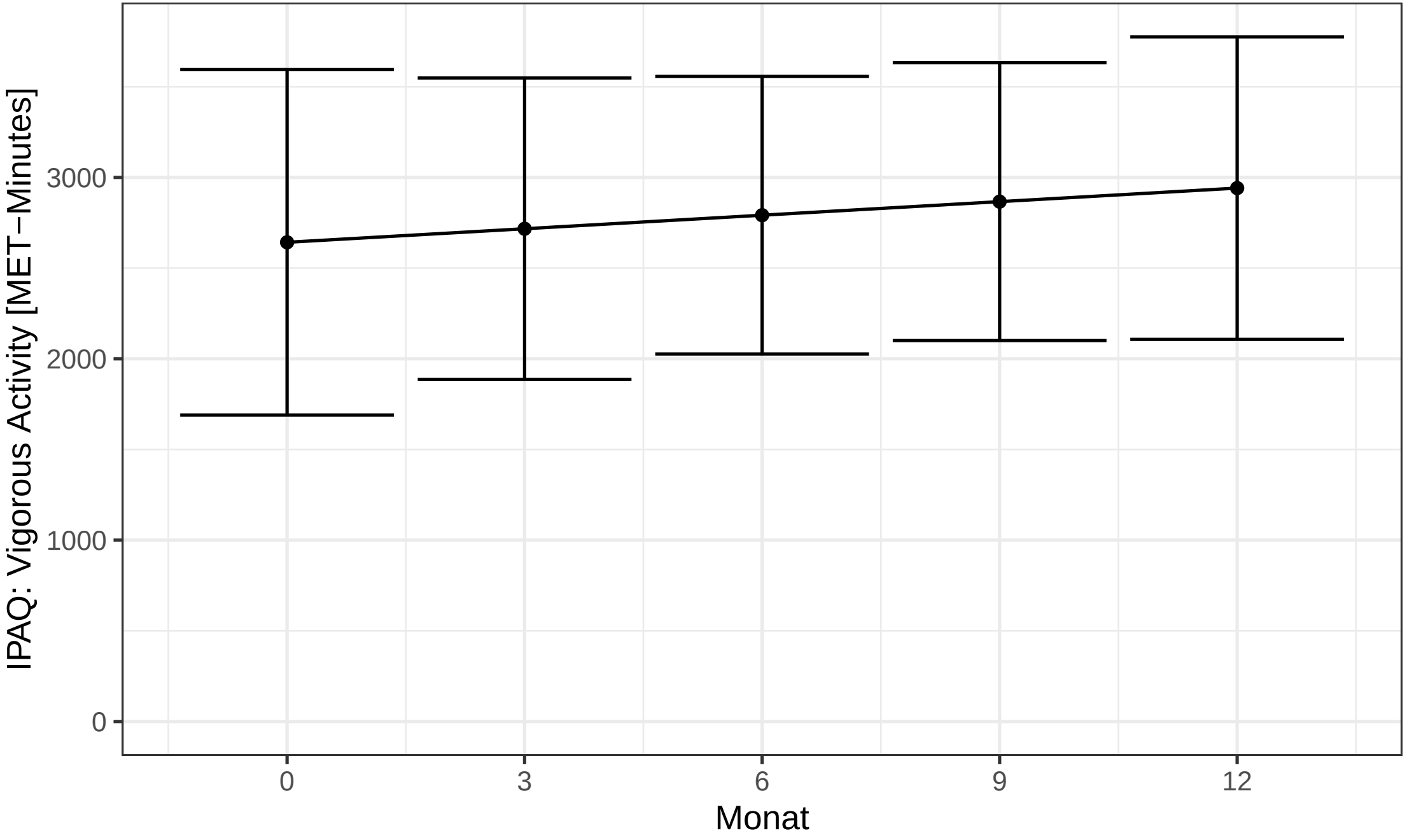
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

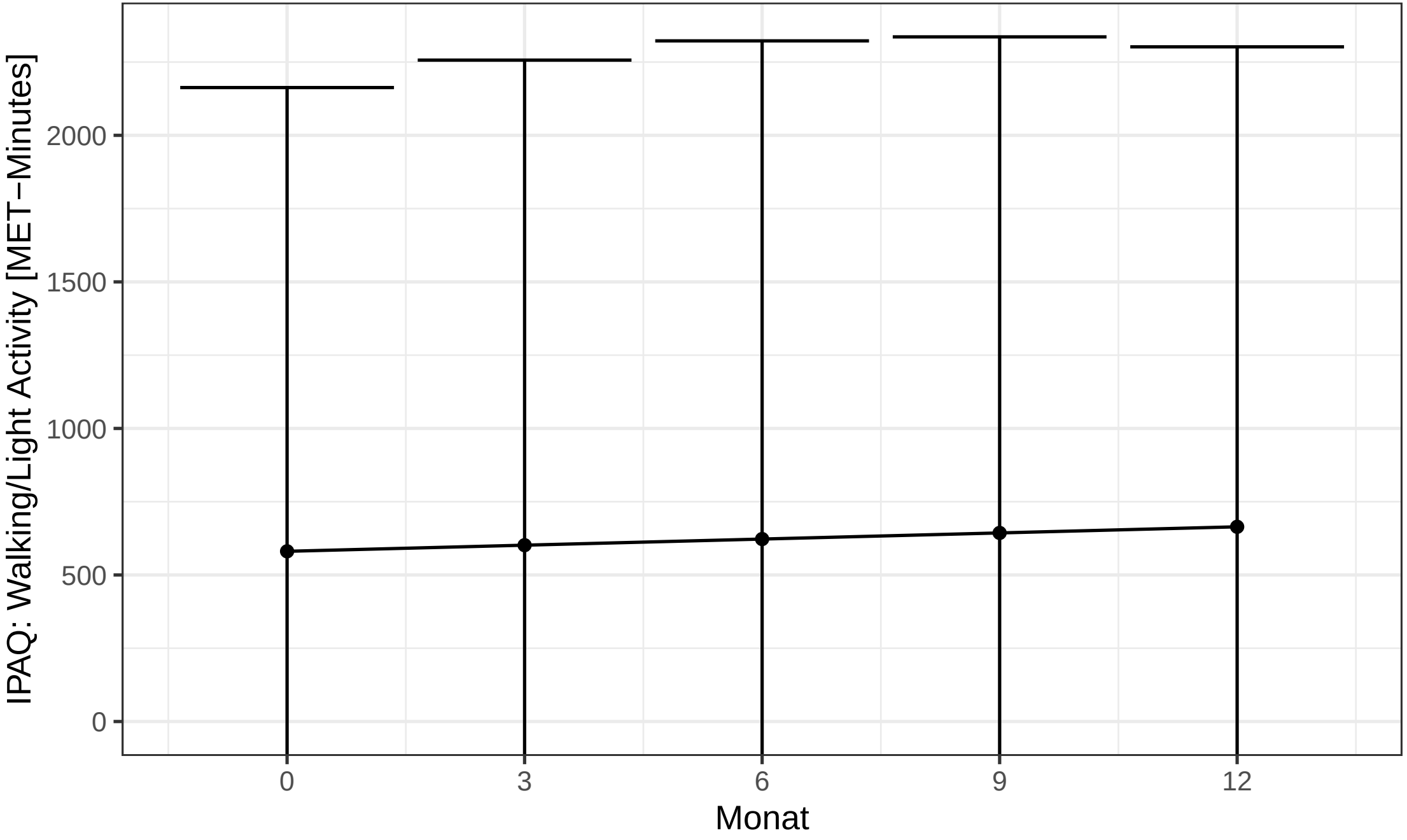
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

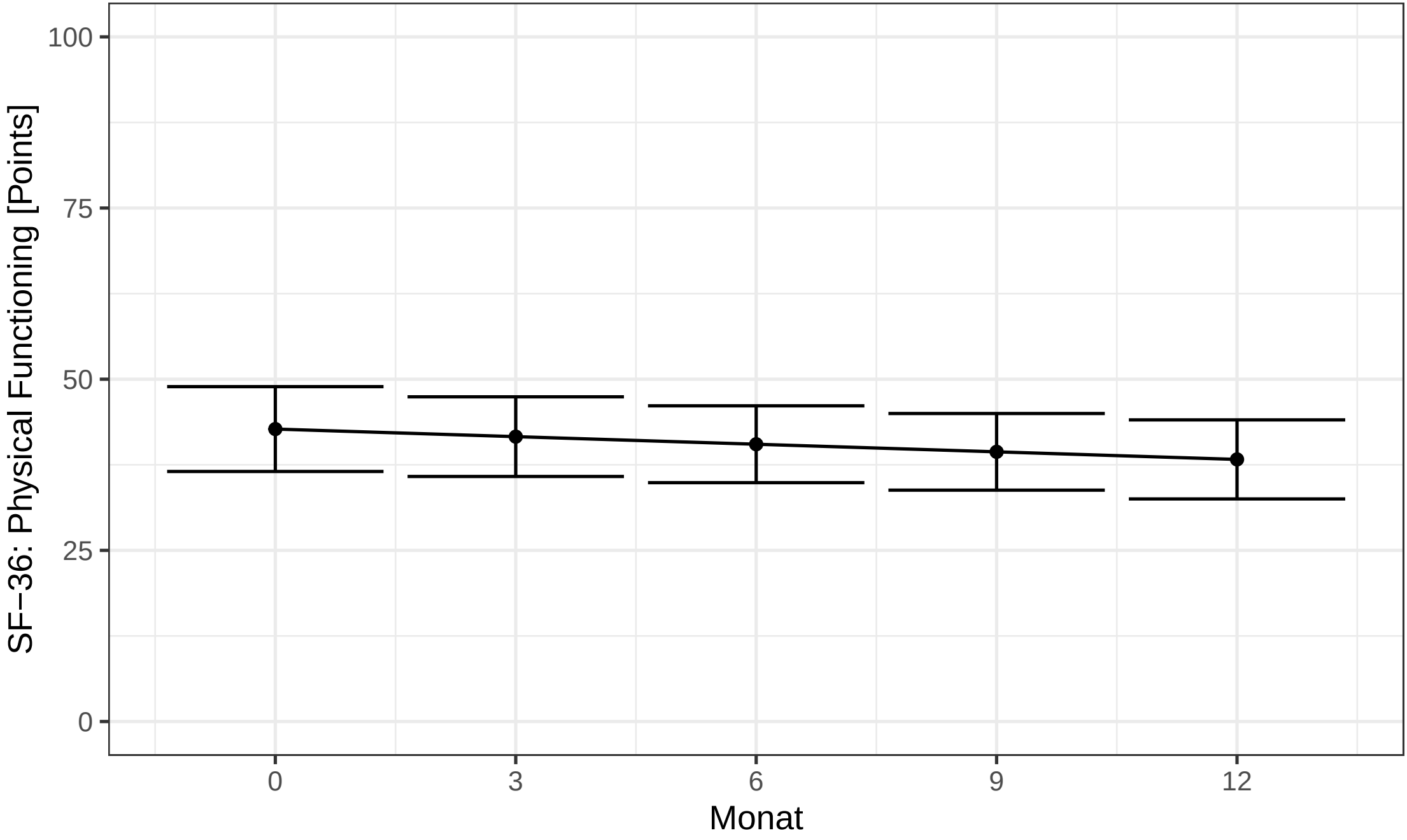
Adjustierter Mittelwert



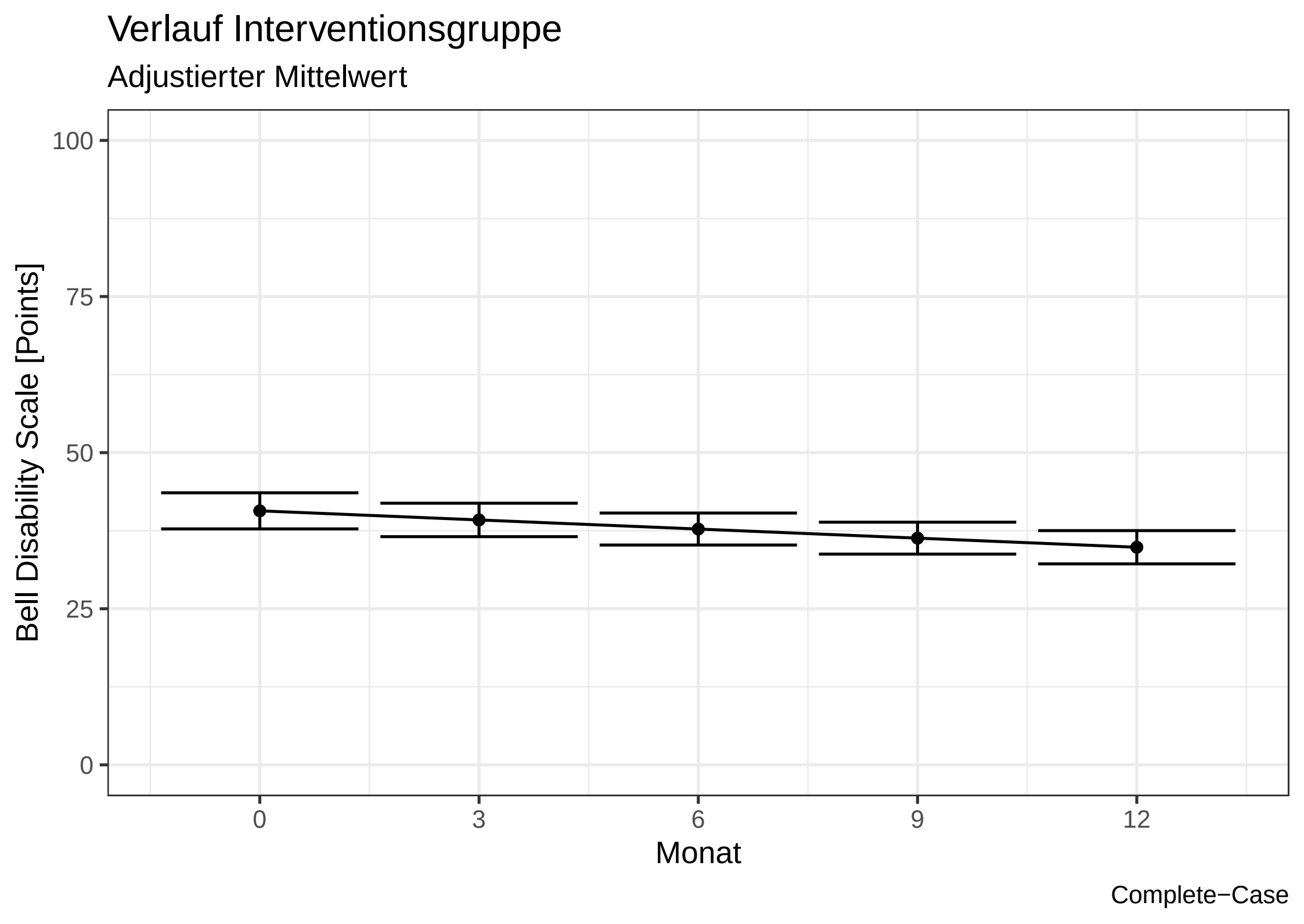
Population C

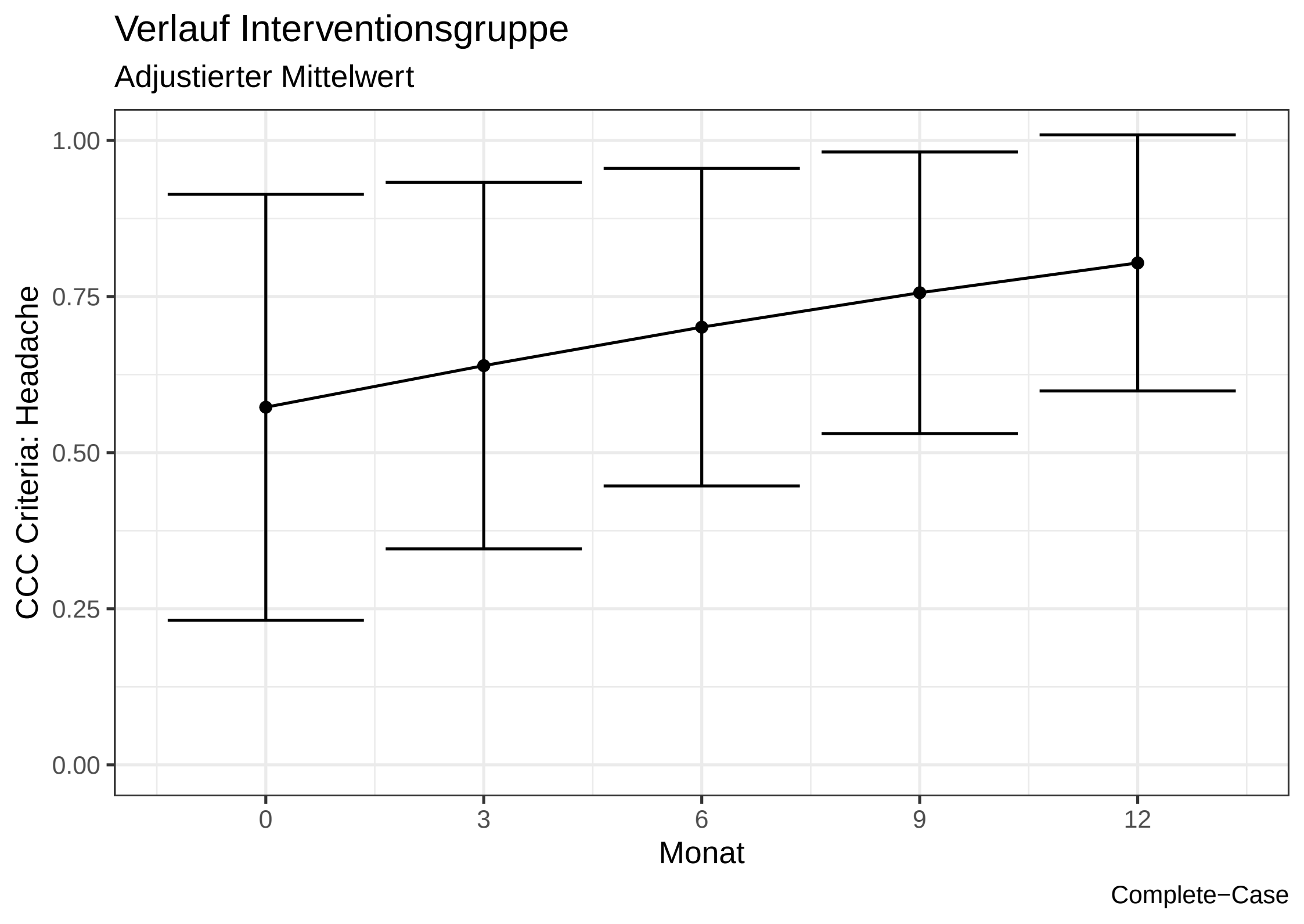
Verlauf Interventionsgruppe

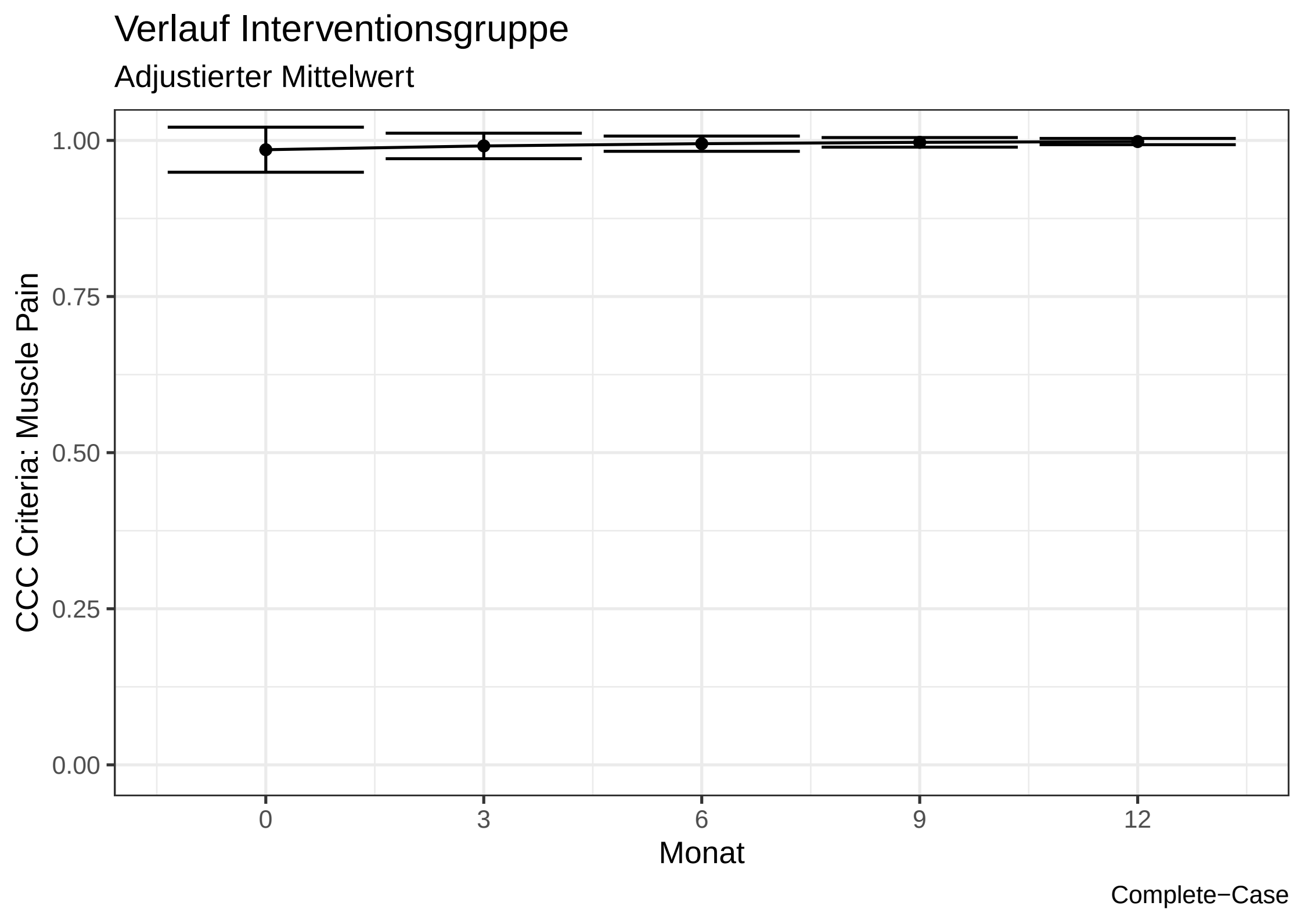
Adjustierter Mittelwert

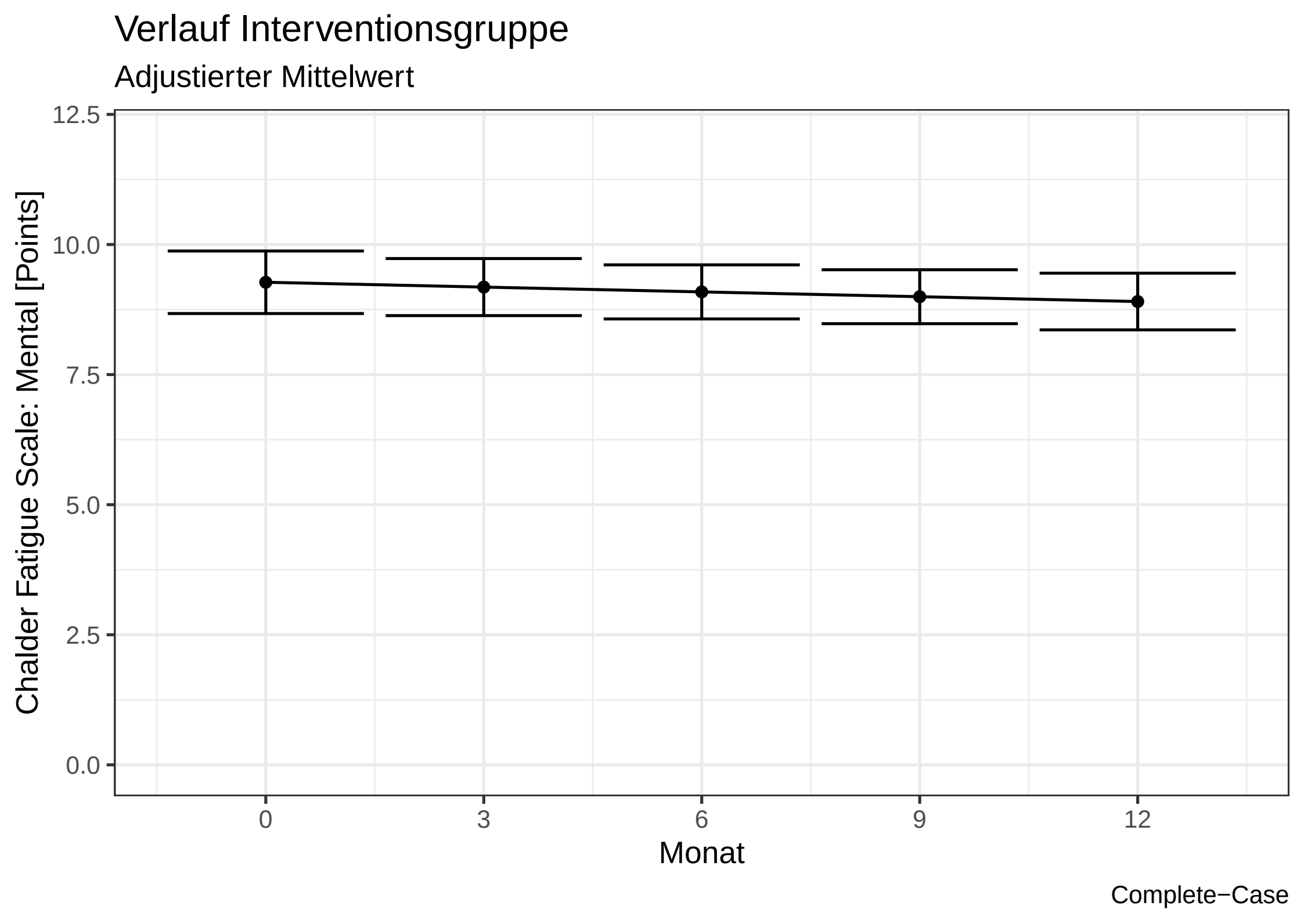


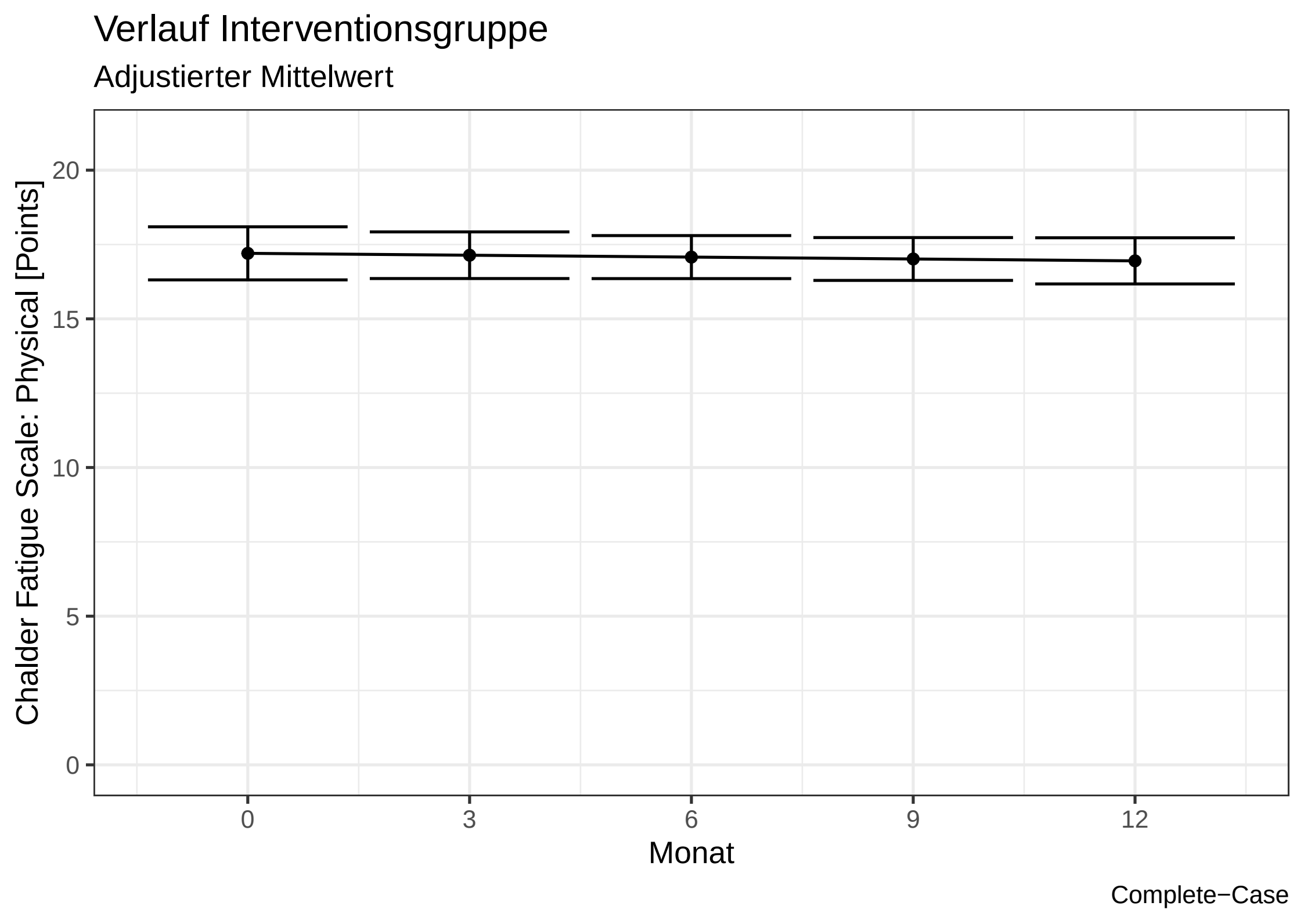
Population C

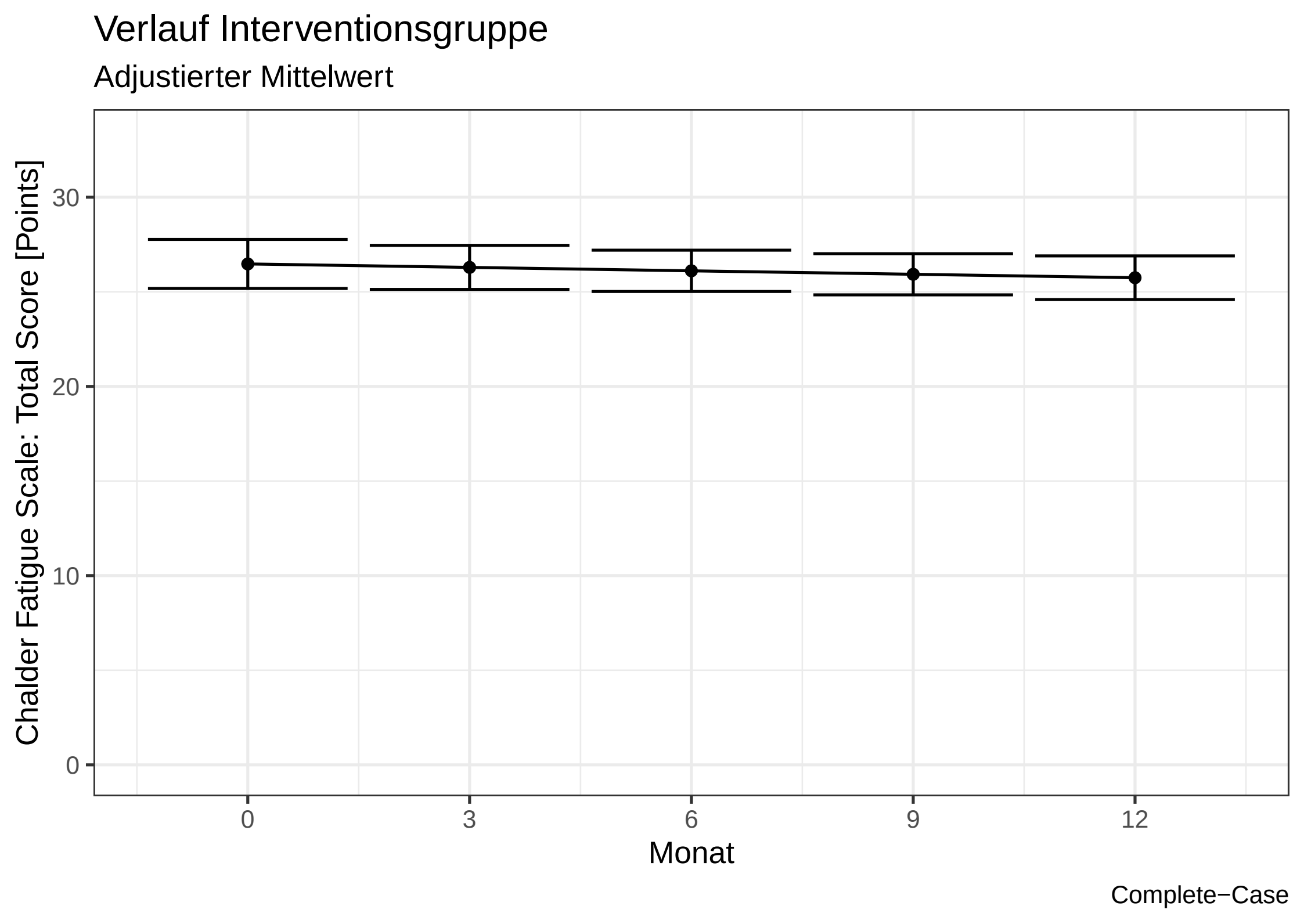

Verlauf Interventionsgruppe
Adjustierter Mittelwert
Bell Disability Scale [Points]
100
75
50
25
0
0
3
6
9
12
Monat
Complete−Case

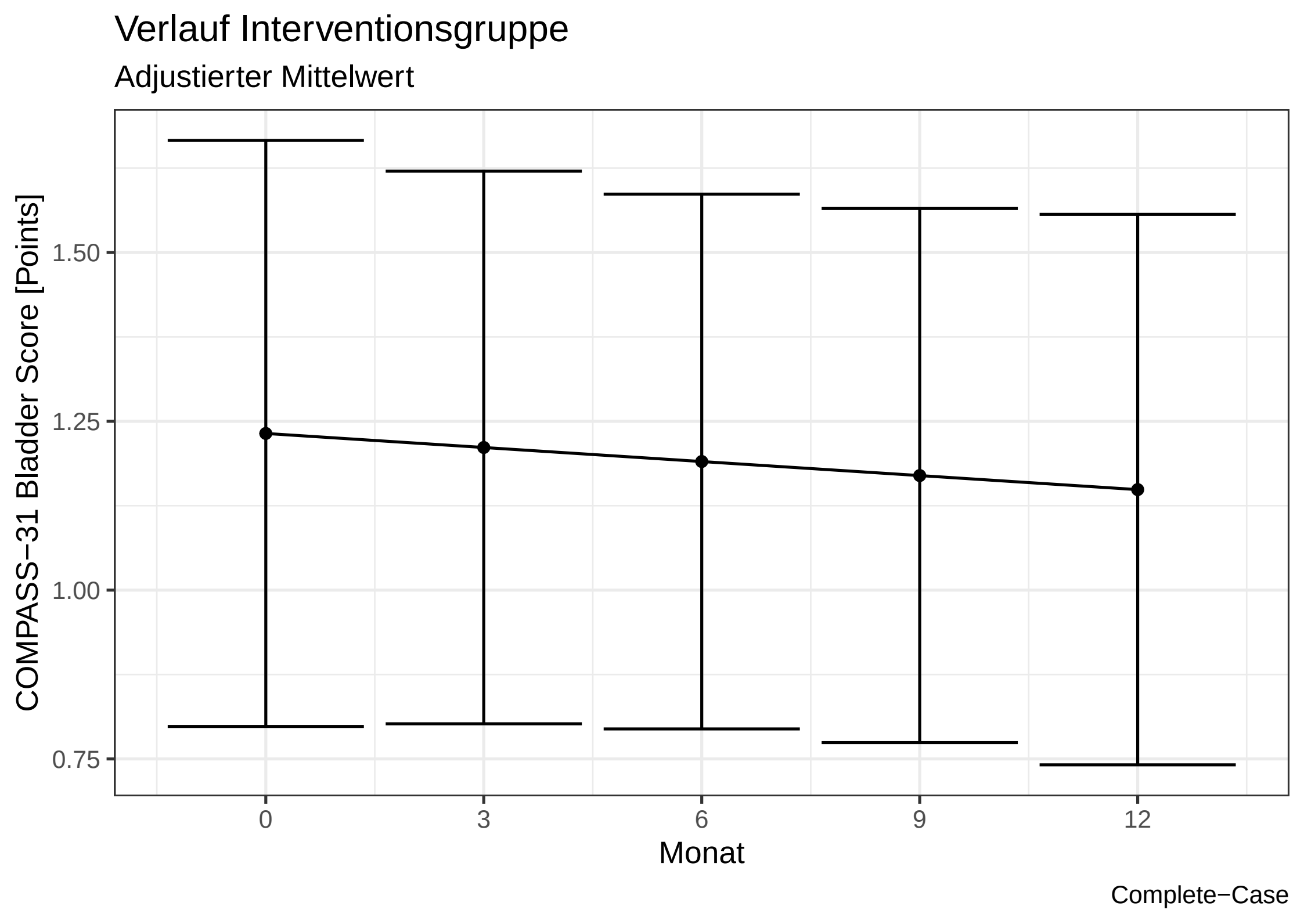

Verlauf Interventionsgruppe
Adjustierter Mittelwert
CCC Criteria: Headache
1.00
0.75
0.50
0.25
0.00
0
3
6
9
12
Monat
Complete−Case

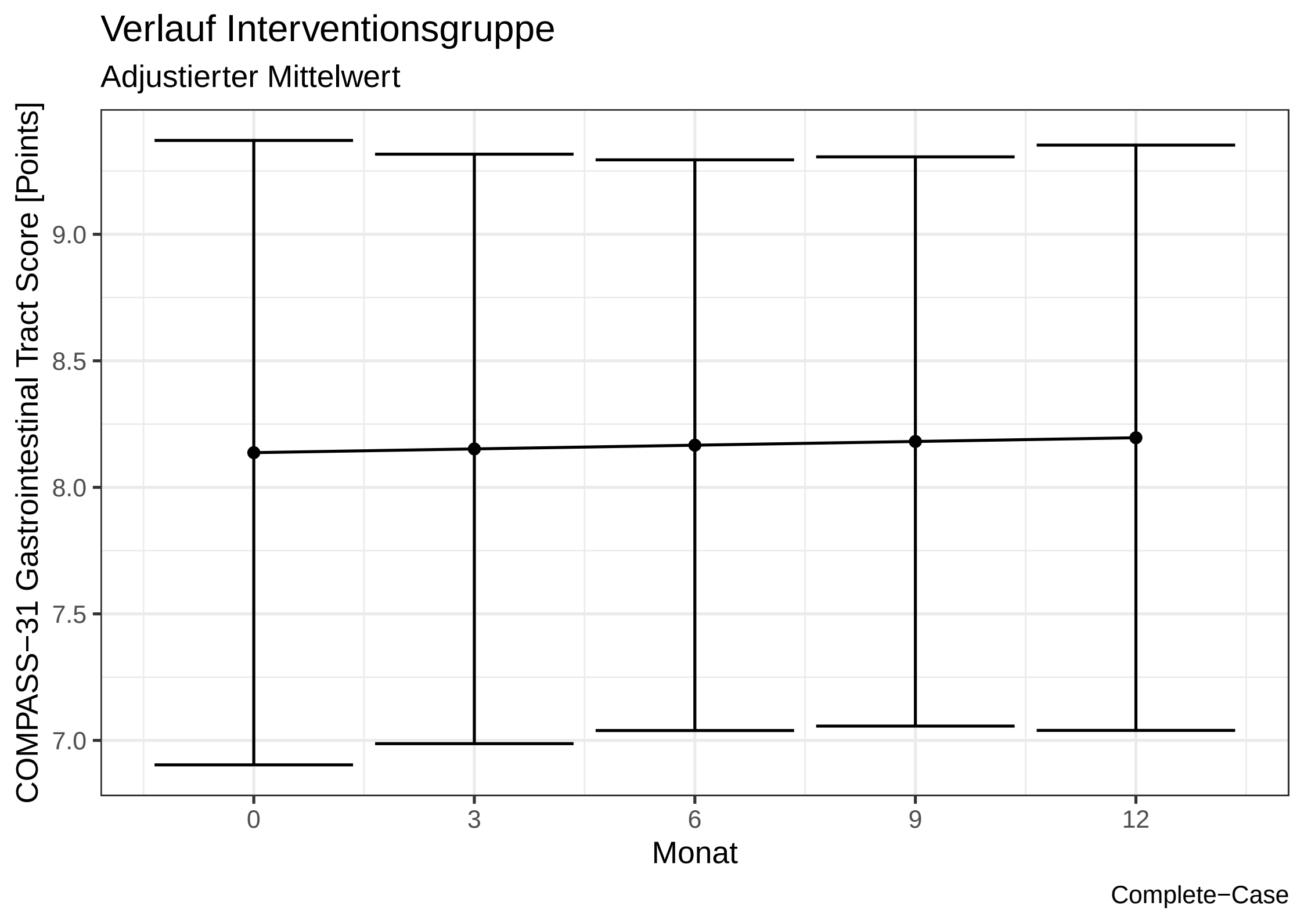

Verlauf Interventionsgruppe
Adjustierter Mittelwert
CCC Criteria: Muscle Pain
1.00
0.75
0.50
0.25
0.00
0
3
6
9
12
Monat
Complete−Case

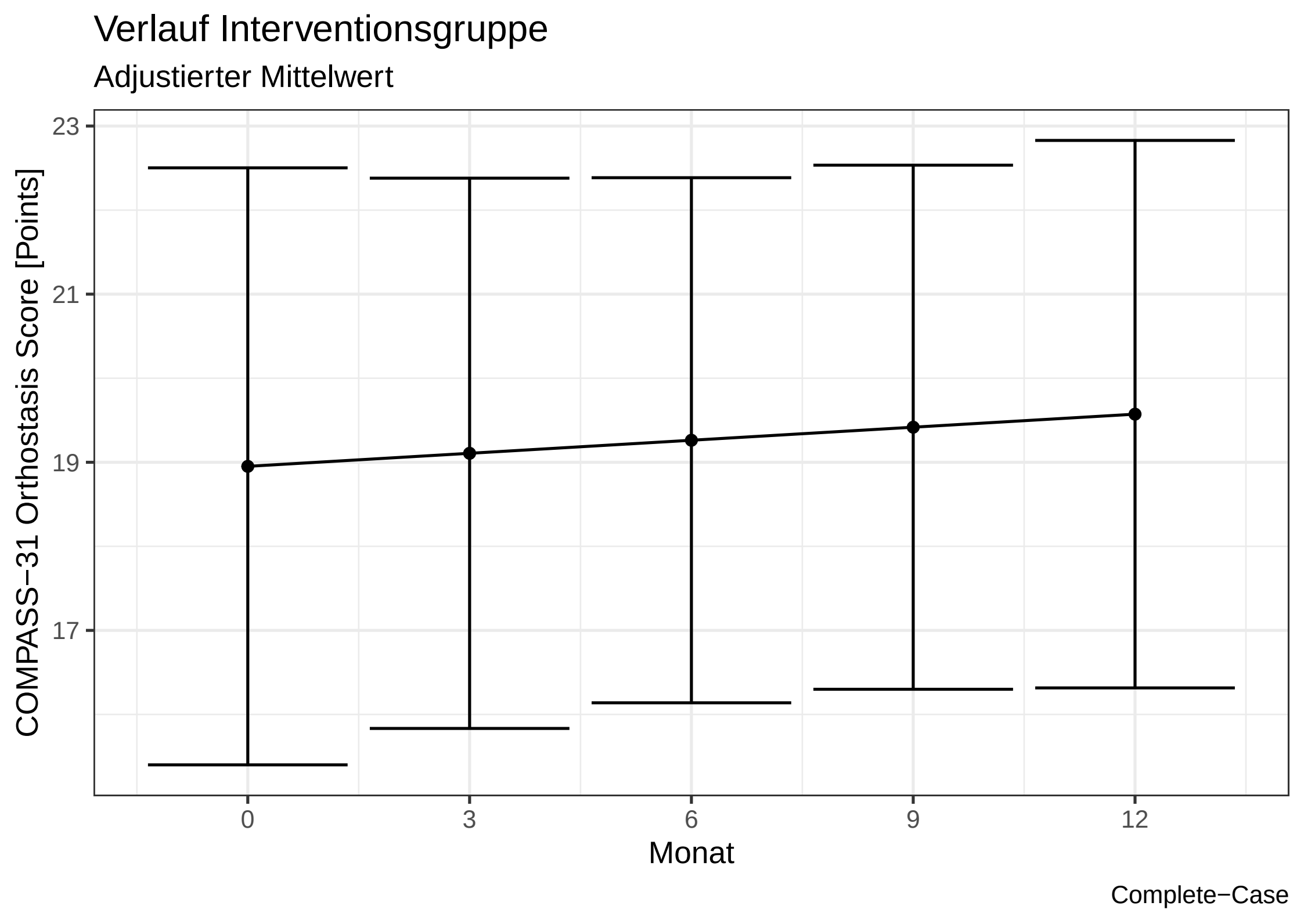

Verlauf Interventionsgruppe
Adjustierter Mittelwert
Chalder Fatigue Scale: Mental [Points]
12.5
10.0
7.5
5.0
2.5
0.0
0
3
6
9
12
Monat
Complete−Case

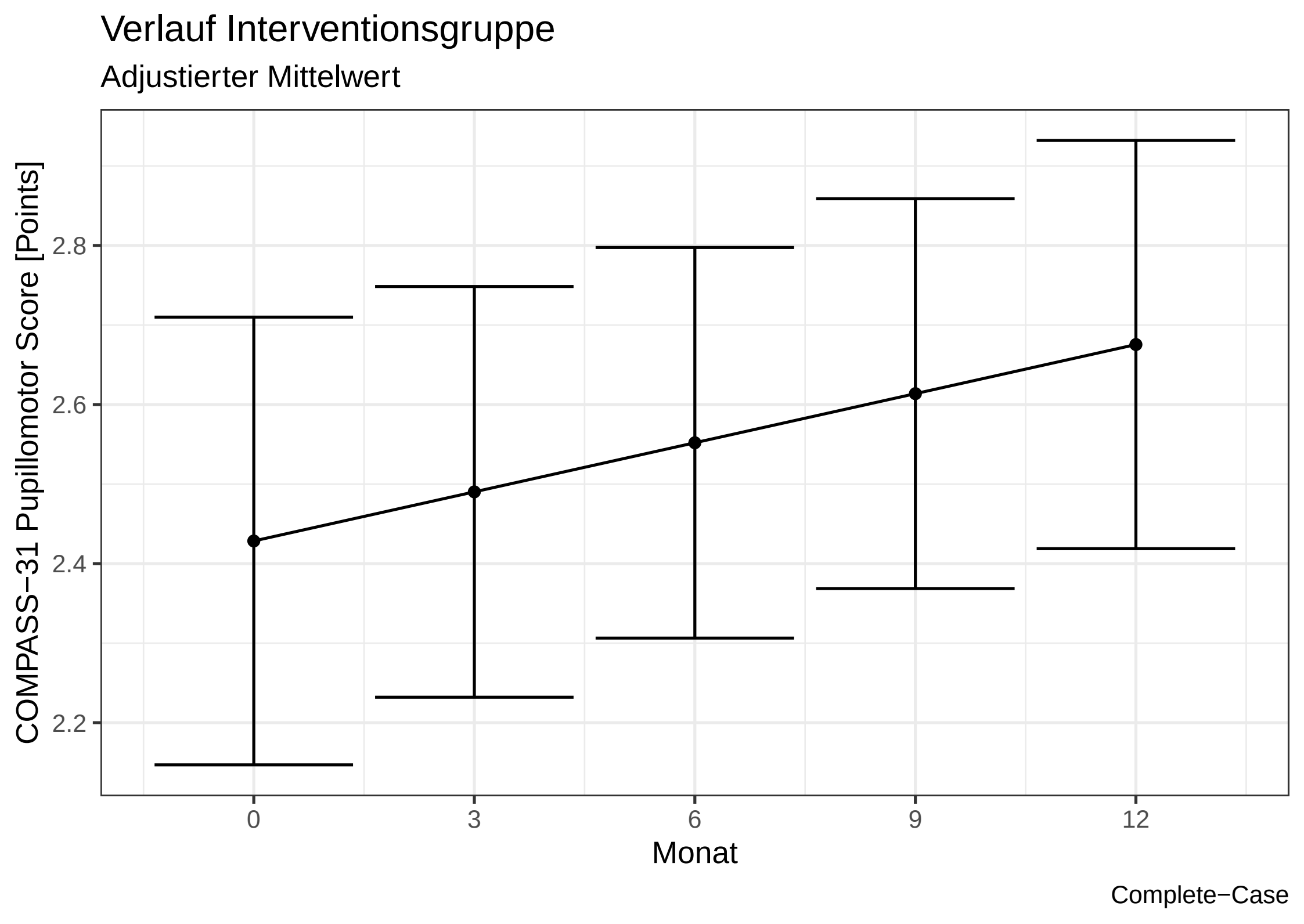

Verlauf Interventionsgruppe
Adjustierter Mittelwert
Chalder Fatigue Scale: Physical [Points]
20
15
10
5
0
0
3
6
9
12
Monat
Complete−Case

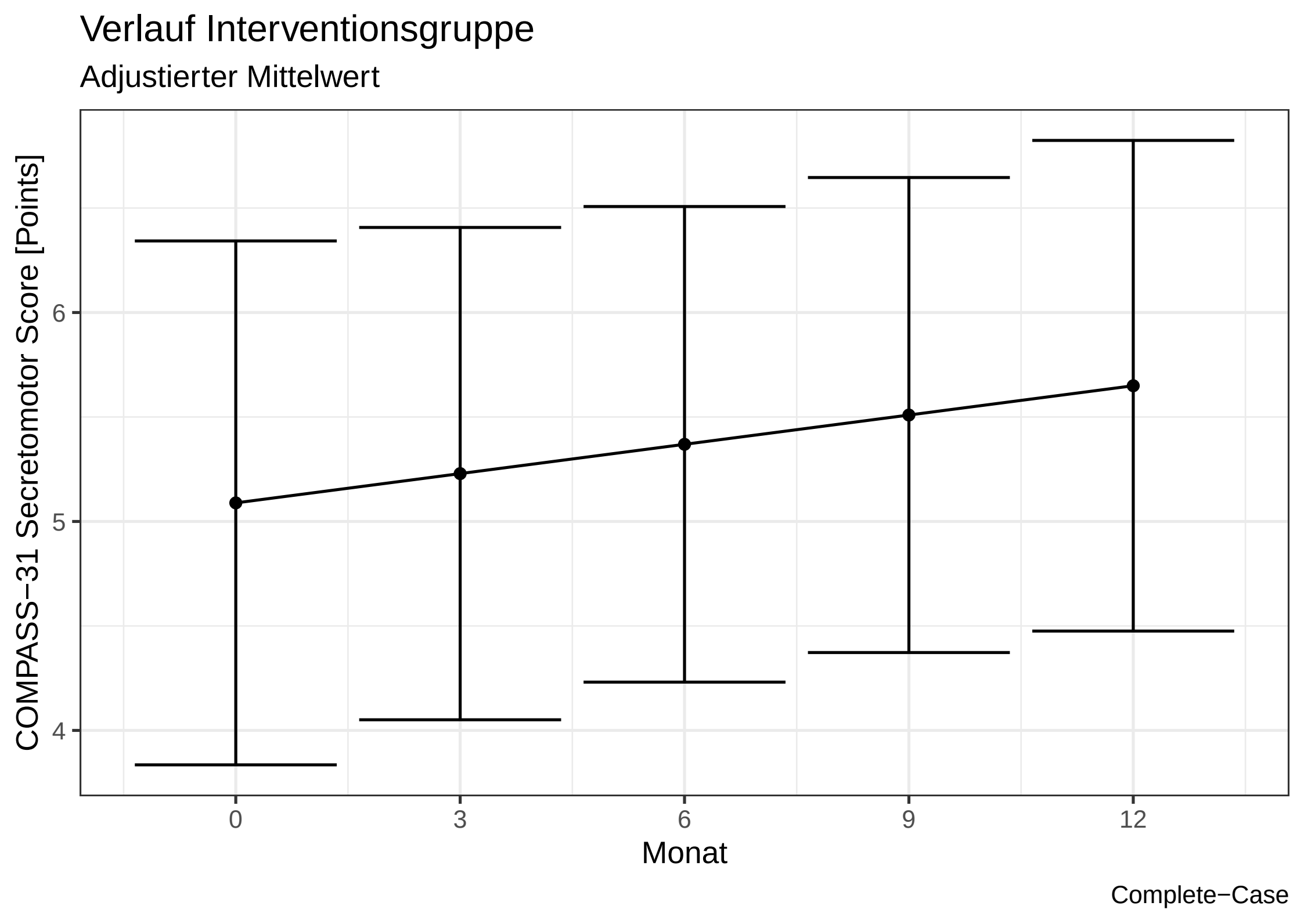

Verlauf Interventionsgruppe
Adjustierter Mittelwert
Chalder Fatigue Scale: Total Score [Points]
30
20
10
0
0
3
6
9
12
Monat
Complete−Case

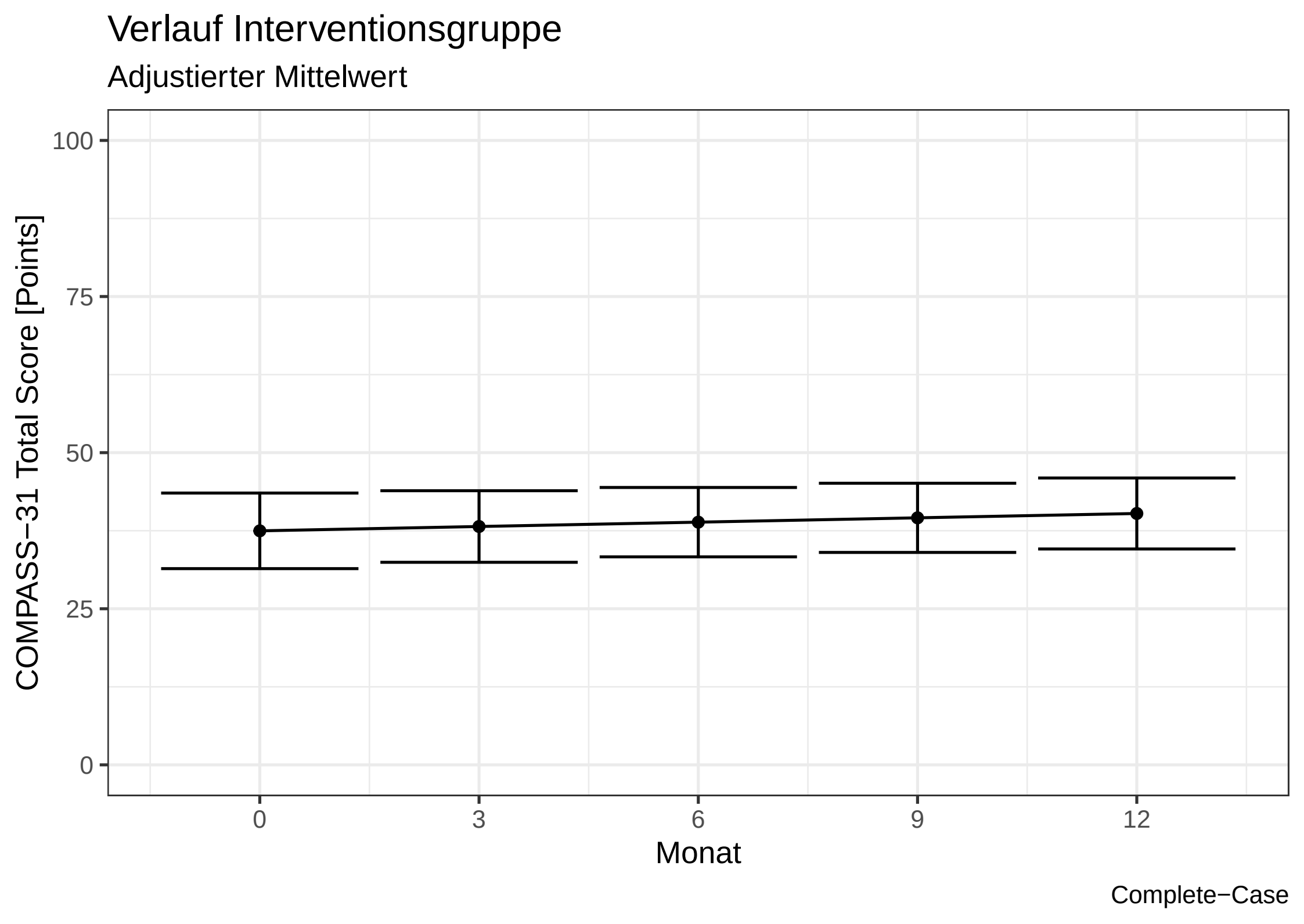

Verlauf Interventionsgruppe
Adjustierter Mittelwert
COMPASS−31 Bladder Score [Points]
1.50
1.25
1.00
0.75
0
3
6
9
12
Monat
Complete−Case

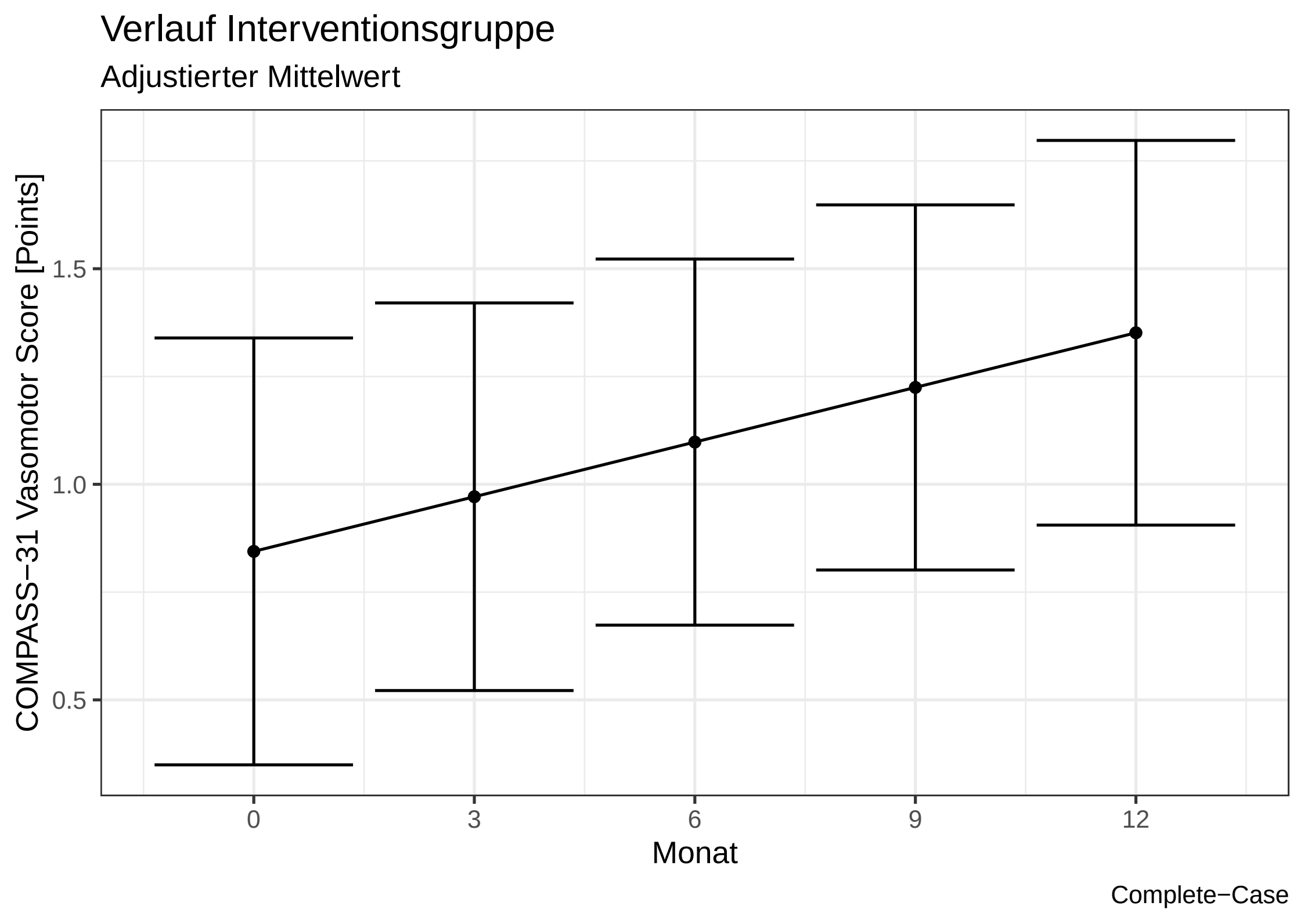

Verlauf Interventionsgruppe
Adjustierter Mittelwert
COMPASS−31 Gastrointestinal Tract Score [Points]
9.0
8.5
8.0
7.5
7.0
0
3
6
9
12
Monat
Complete−Case

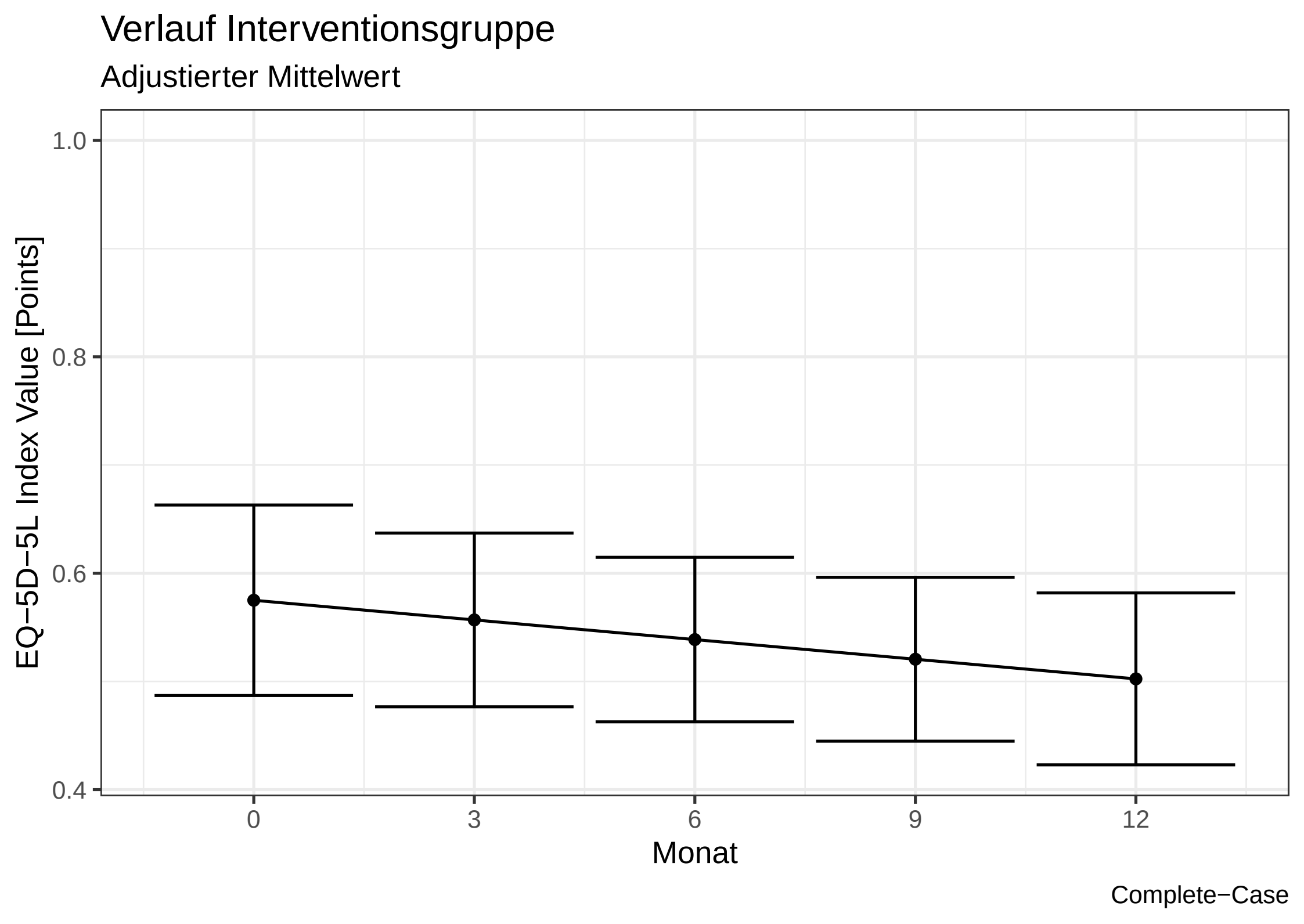

Verlauf Interventionsgruppe
Adjustierter Mittelwert
COMPASS−31 Orthostasis Score [Points]
23
21
19
17
0
3
6
9
12
Monat
Complete−Case

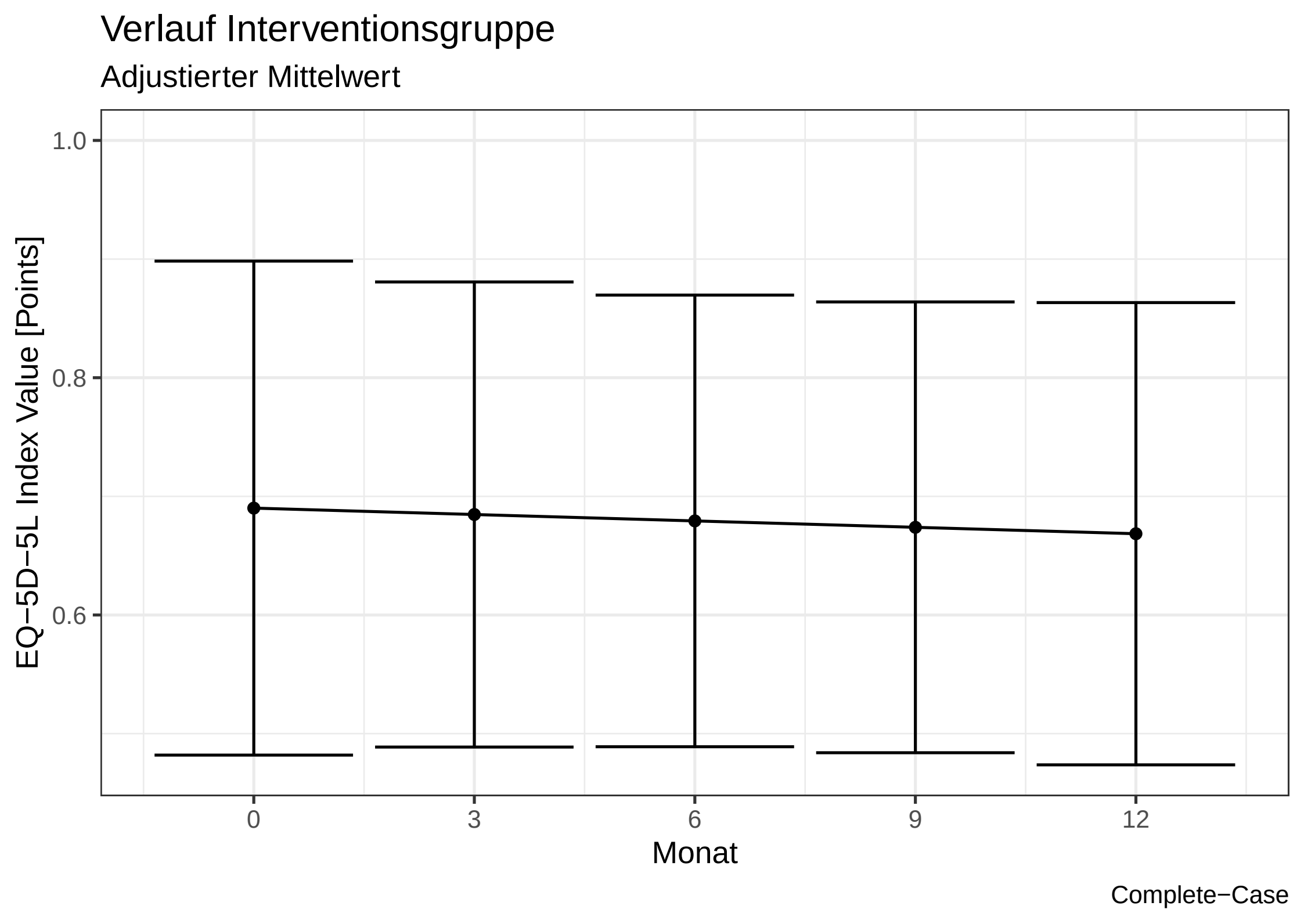

Verlauf Interventionsgruppe
Adjustierter Mittelwert
COMPASS−31 Pupillomotor Score [Points]
2.8
2.6
2.4
2.2
0
3
6
9
12
Monat
Complete−Case

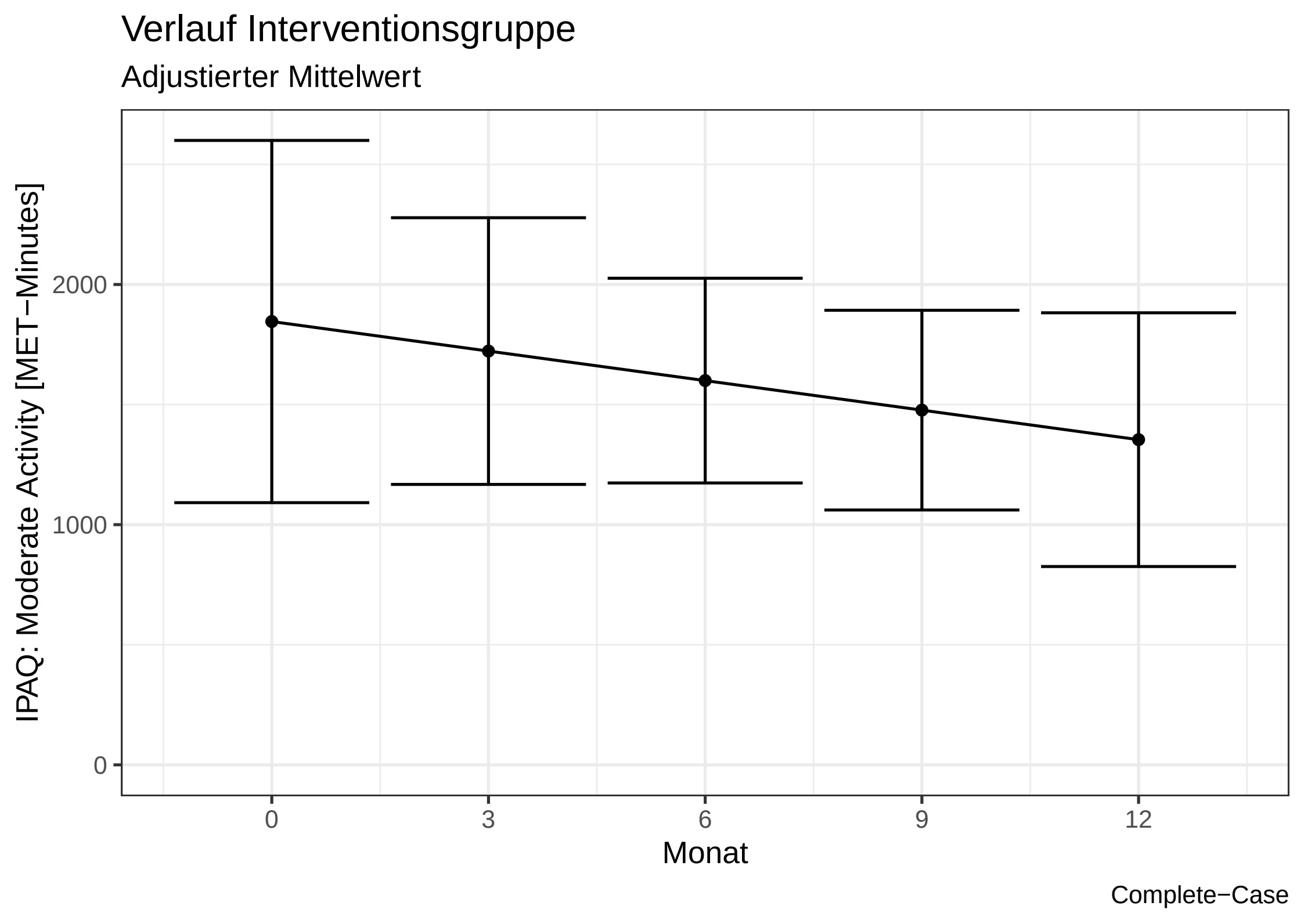

Verlauf Interventionsgruppe
Adjustierter Mittelwert
COMPASS−31 Secretomotor Score [Points]
6
5
4
0
3
6
9
12
Monat
Complete−Case

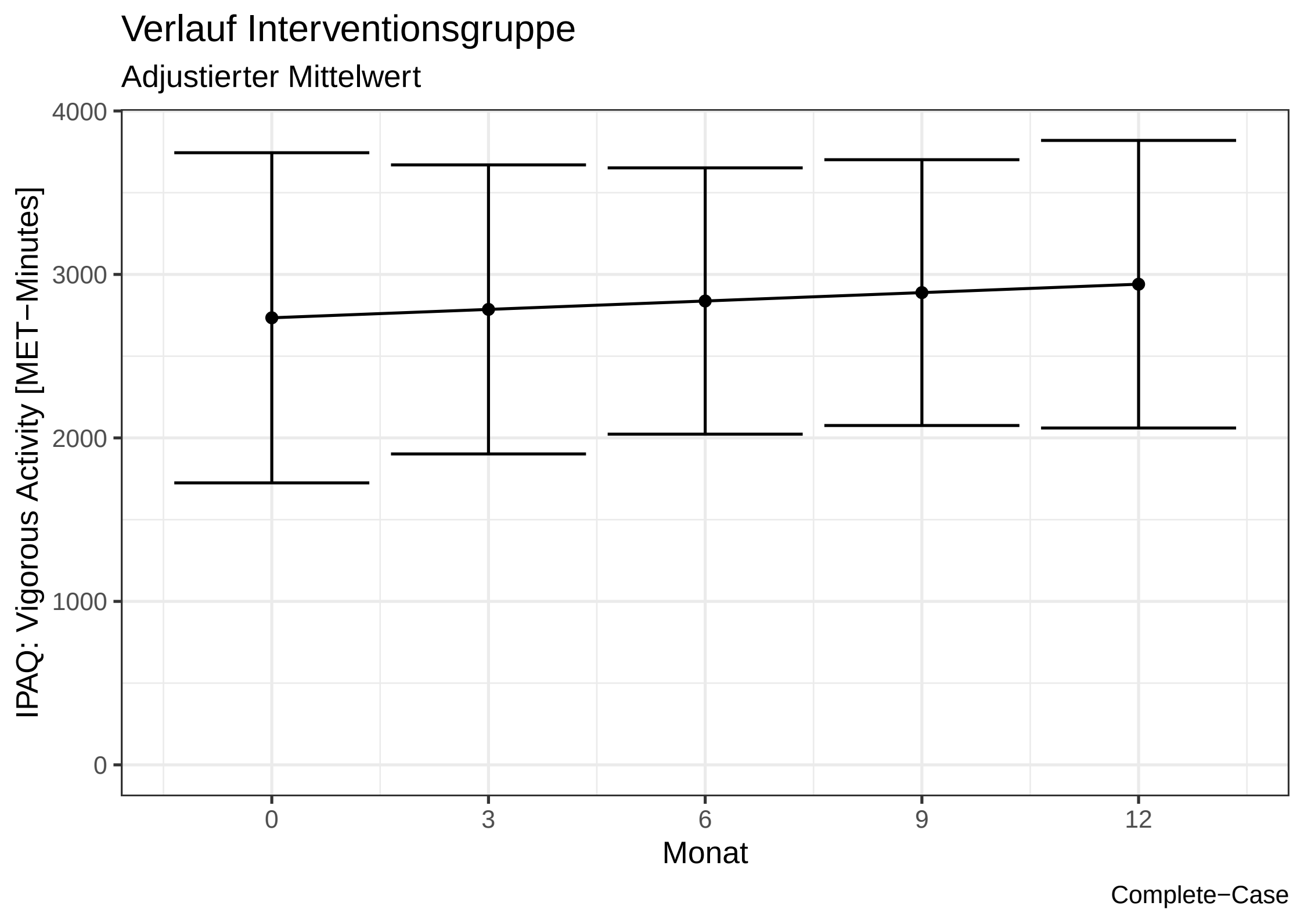

Verlauf Interventionsgruppe
Adjustierter Mittelwert
COMPASS−31 Total Score [Points]
100
75
50
25
0
0
3
6
9
12
Monat
Complete−Case

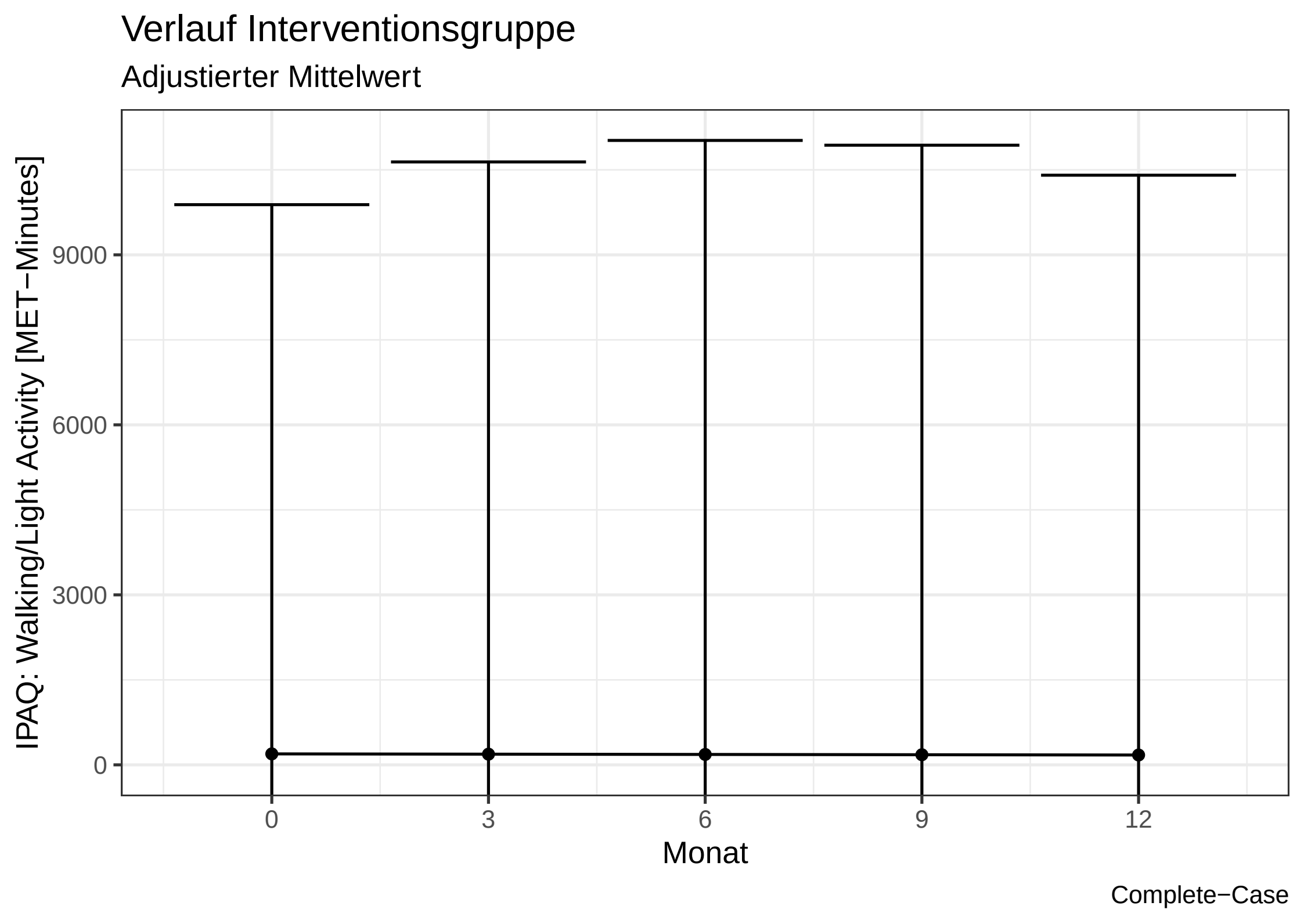

Verlauf Interventionsgruppe
Adjustierter Mittelwert
COMPASS−31 Vasomotor Score [Points]
1.5
1.0
0.5
0
3
6
9
12
Monat
Complete−Case

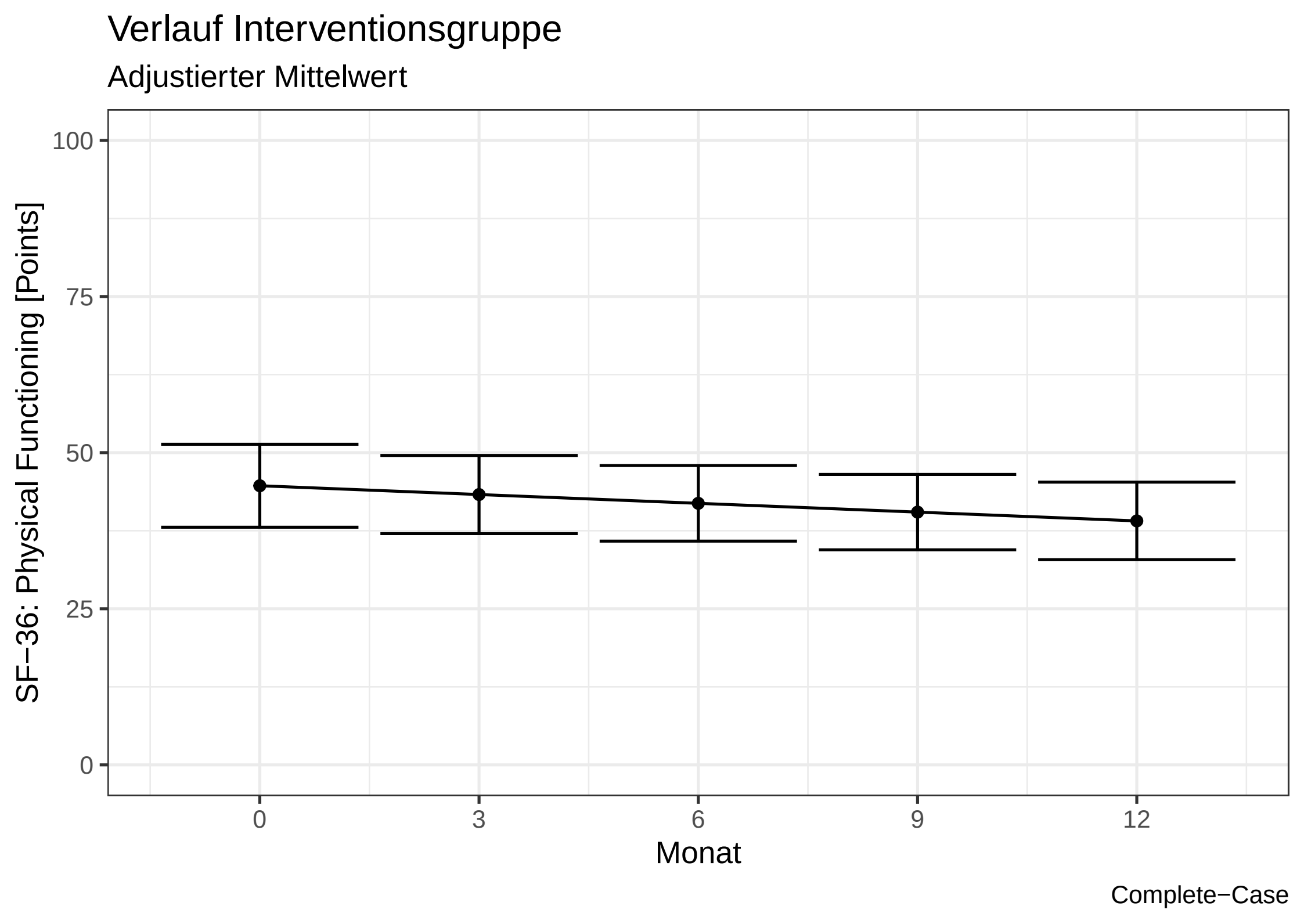

Verlauf Interventionsgruppe
Adjustierter Mittelwert
EQ−5D−5L Index Value [Points]
1.0
0.8
0.6
0.4
0
3
6
9
12
Monat
Complete−Case


Verlauf Interventionsgruppe
Adjustierter Mittelwert
1.0
0.8
0.6
EQ−5D−5L Index Value [Points]
0
3
6
9
12
Monat
Complete−Case


Verlauf Interventionsgruppe
Adjustierter Mittelwert
IPAQ: Moderate Activity [MET−Minutes]
2000
1000
0
0
3
6
9
12
Monat
Complete−Case

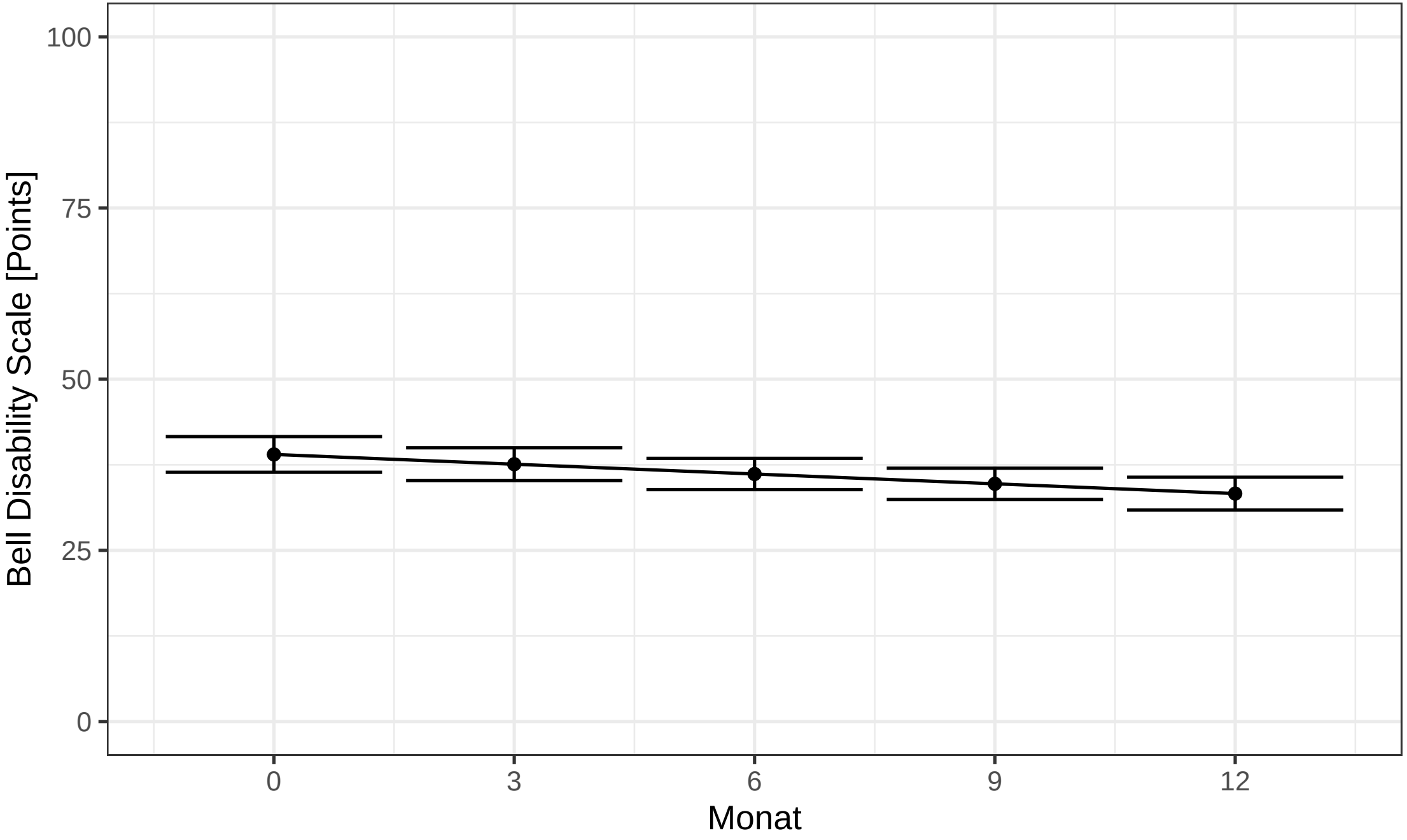

Verlauf Interventionsgruppe
Adjustierter Mittelwert
IPAQ: Vigorous Activity [MET−Minutes]
4000
3000
2000
1000
0
0
3
6
9
12
Monat
Complete−Case


Verlauf Interventionsgruppe
Adjustierter Mittelwert
IPAQ: Walking/Light Activity [MET−Minutes]
9000
6000
3000
0
0
3
6
9
12
Monat
Complete−Case


Verlauf Interventionsgruppe
Adjustierter Mittelwert
SF−36: Physical Functioning [Points]
100
75
50
25
0
0
3
6
9
12
Monat
Complete−Case

Verlauf Interventionsgruppe

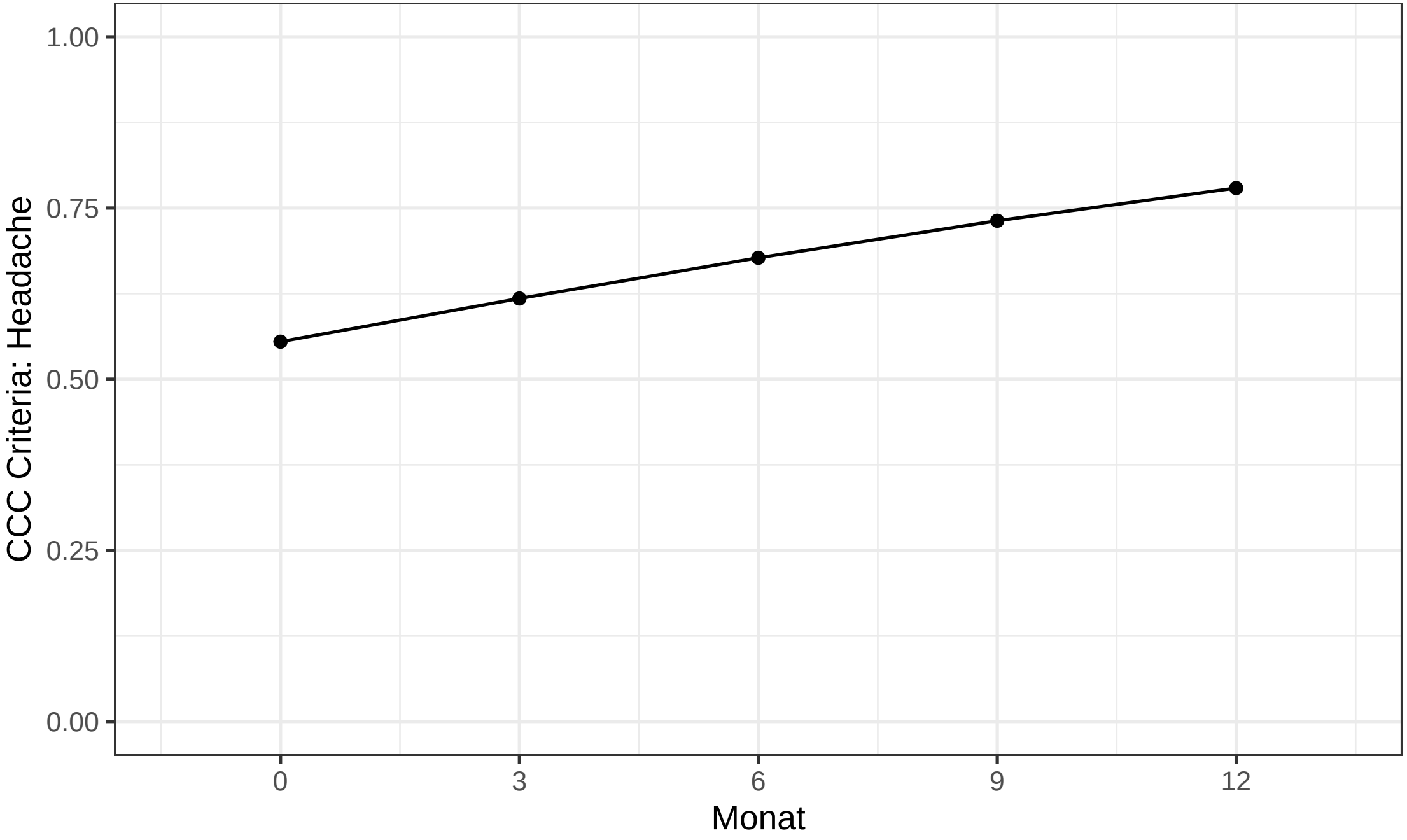
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

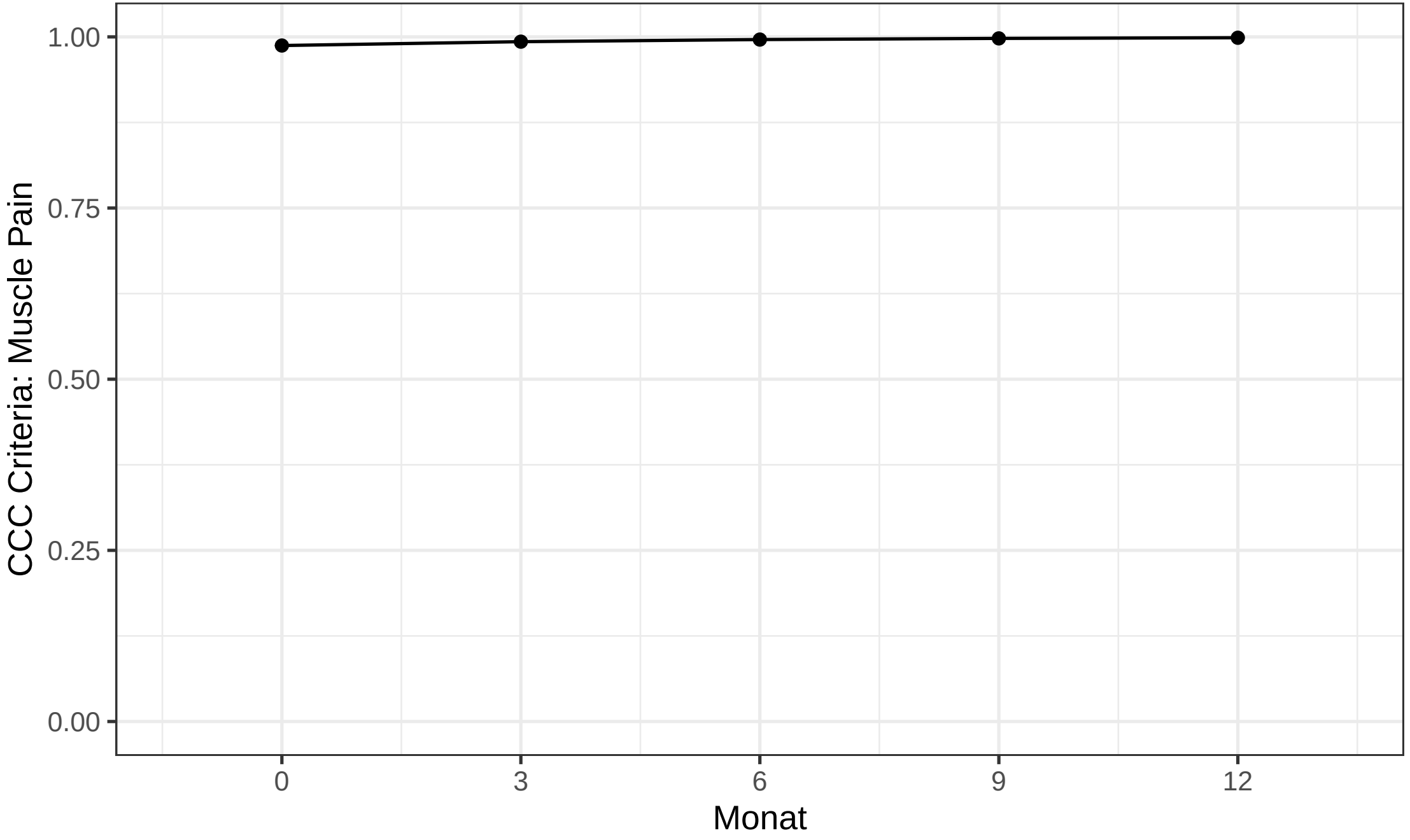
Adjustierter Mittelwert



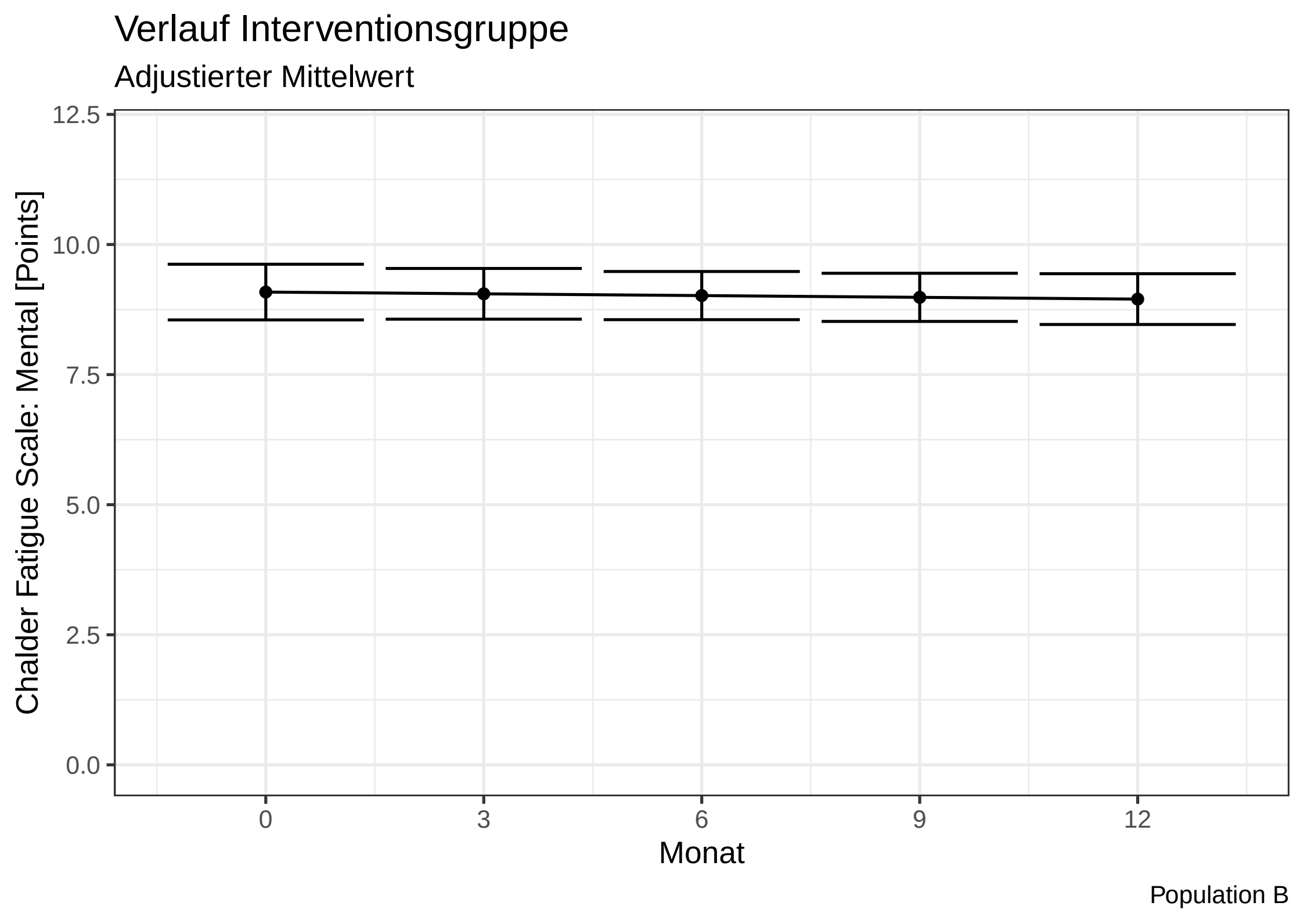
Population B

Verlauf Interventionsgruppe

Adjustierter Mittelwert

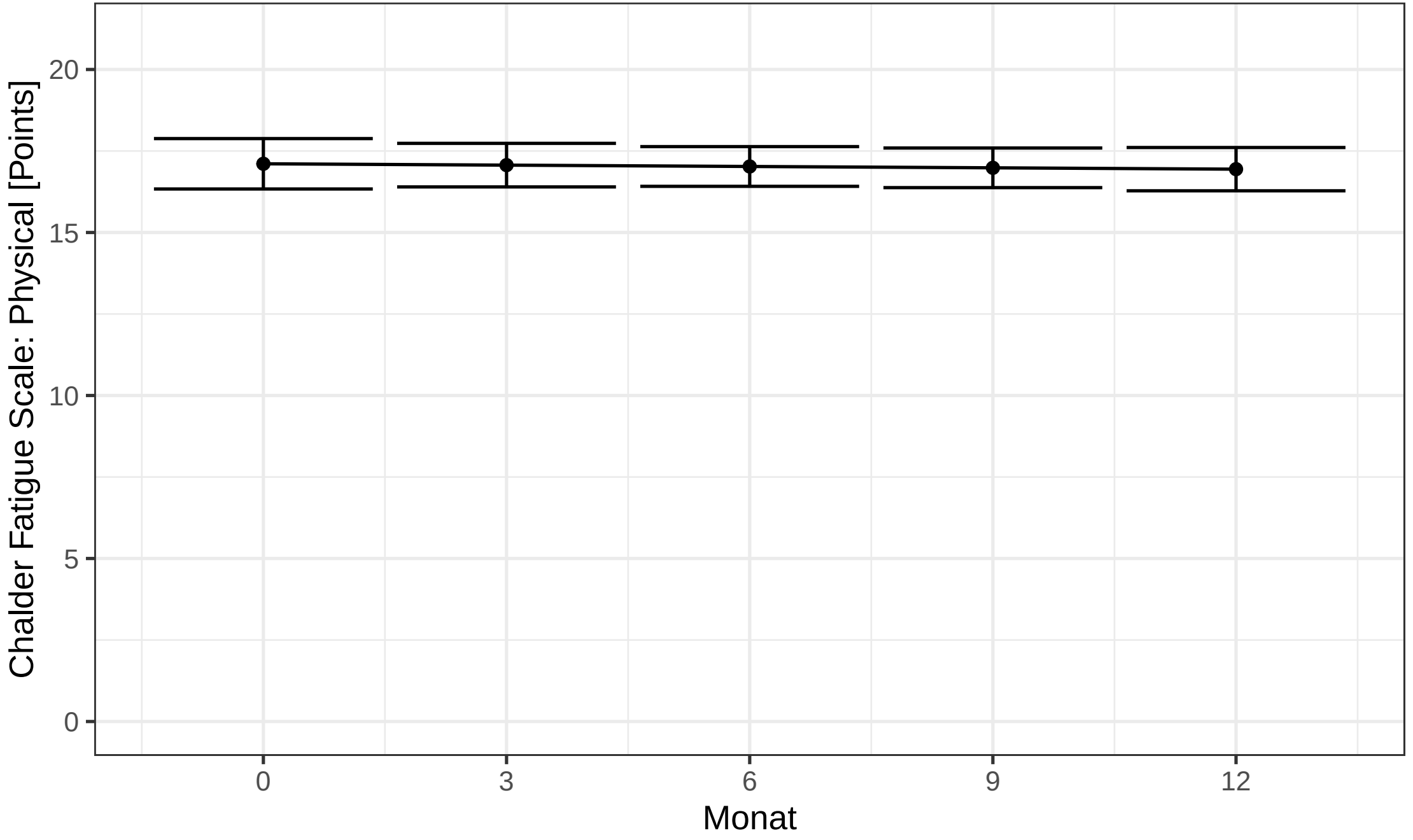


Population B


Verlauf Interventionsgruppe
Adjustierter Mittelwert
Chalder Fatigue Scale: Mental [Points]
12.5
10.0
7.5
5.0
2.5
0.0
0
3
6
9
12
Monat
Population B

Verlauf Interventionsgruppe

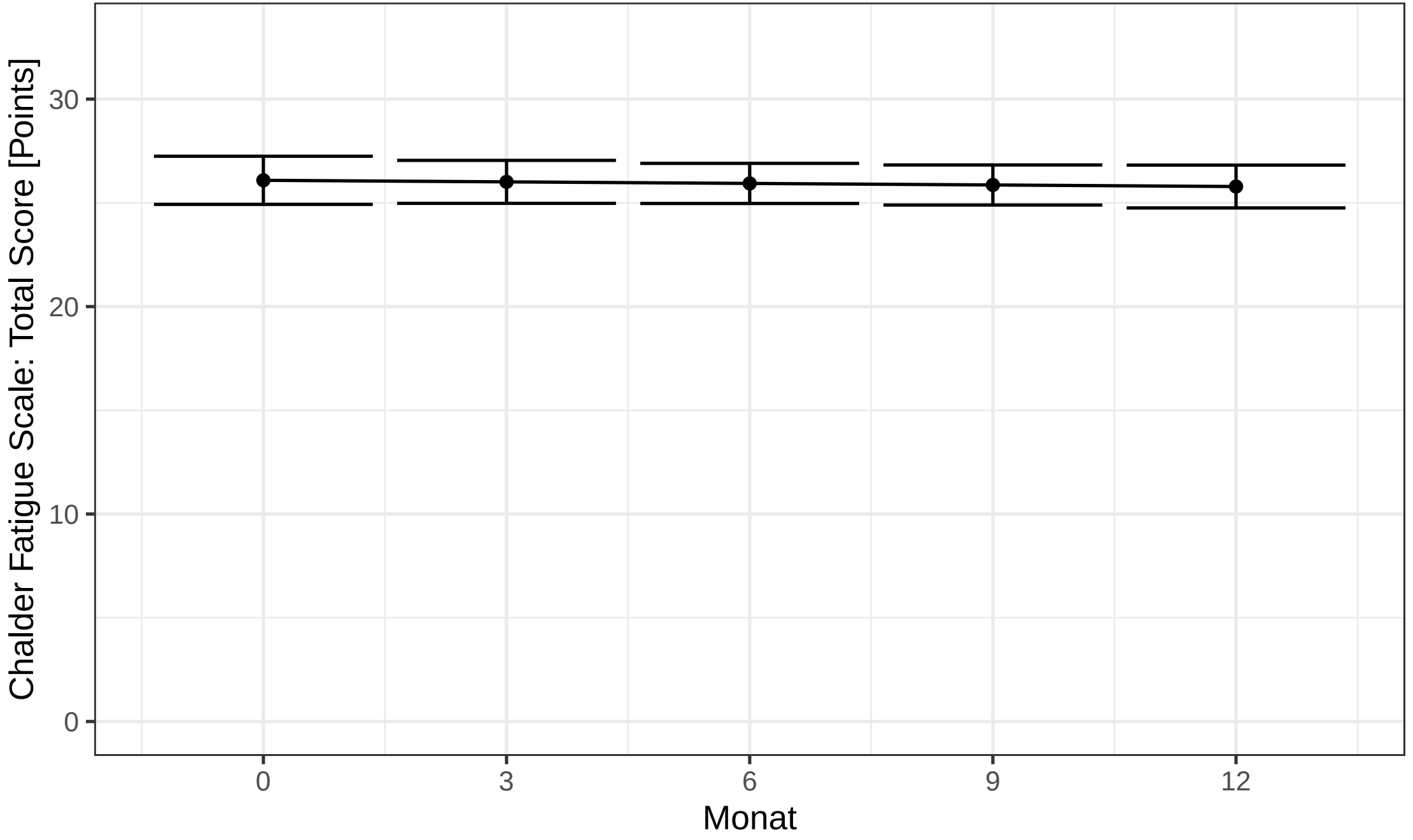
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

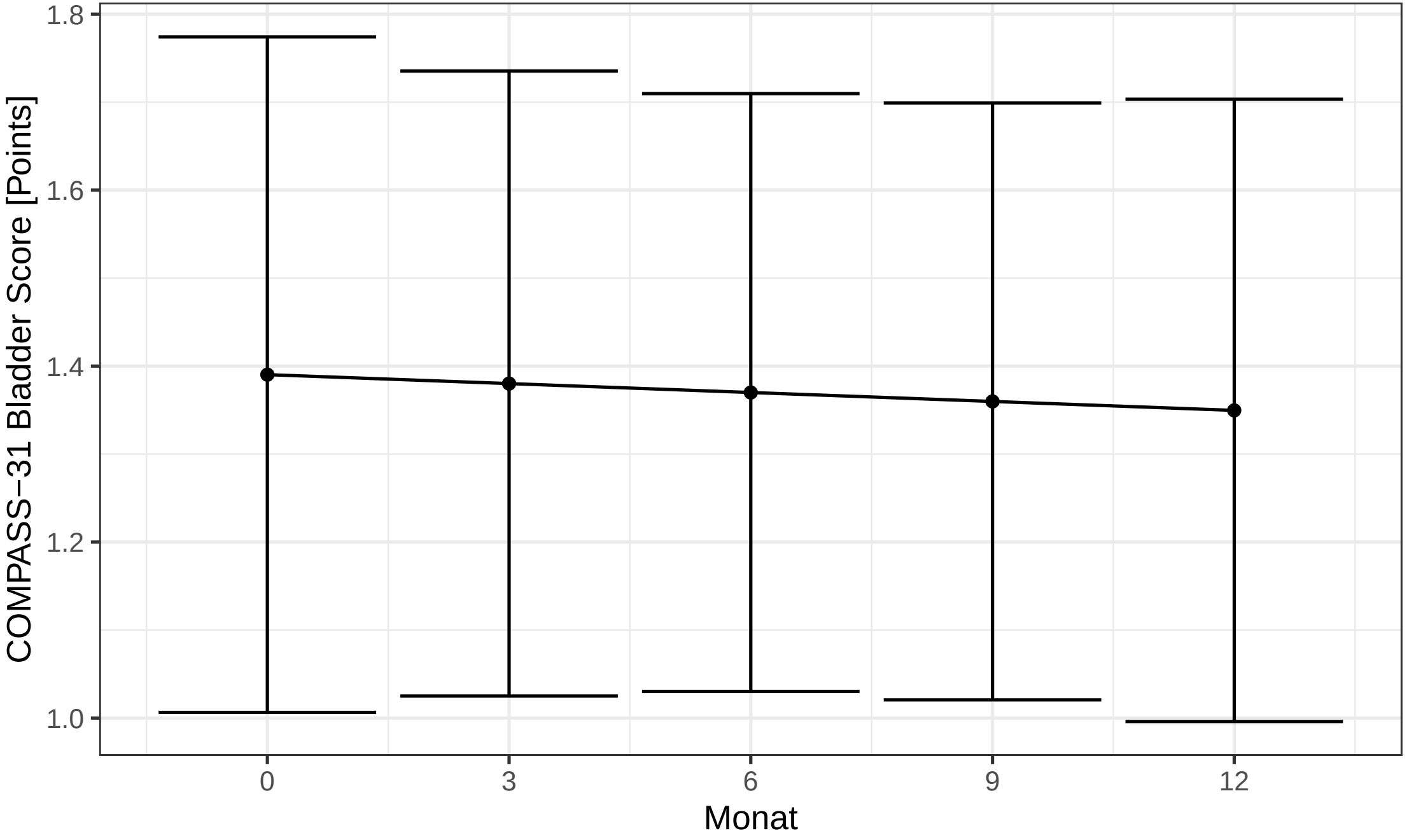
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

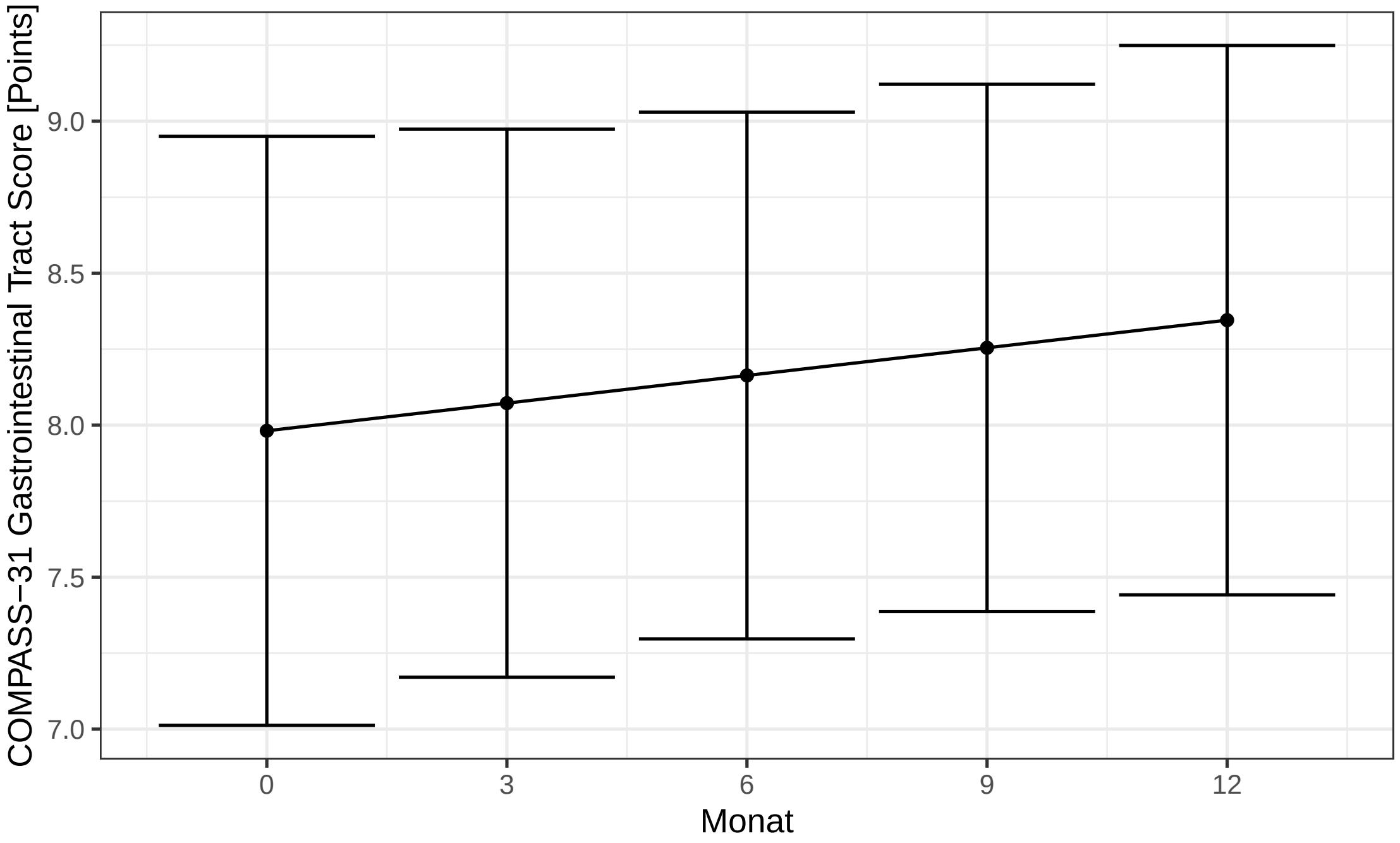
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

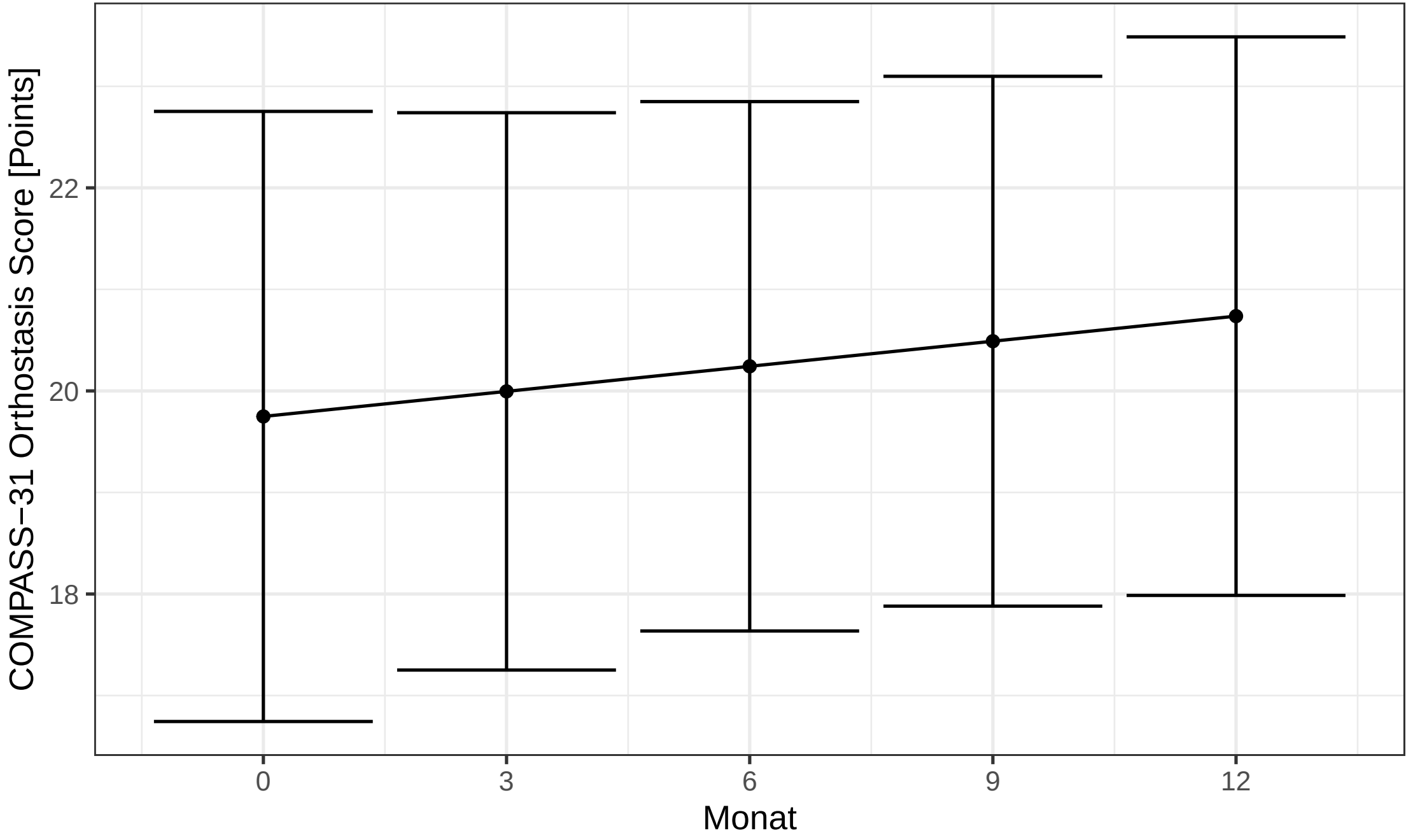
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

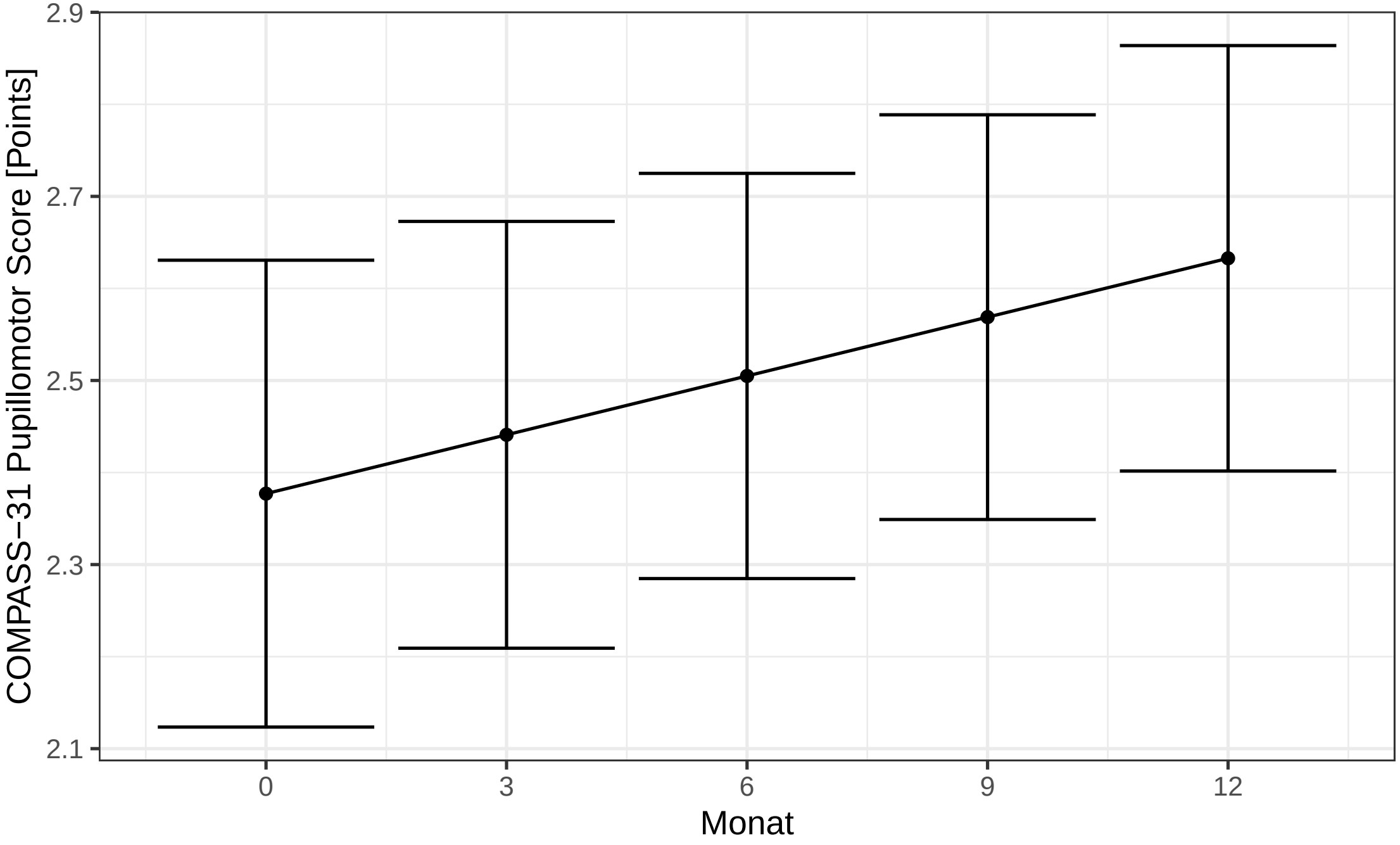
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

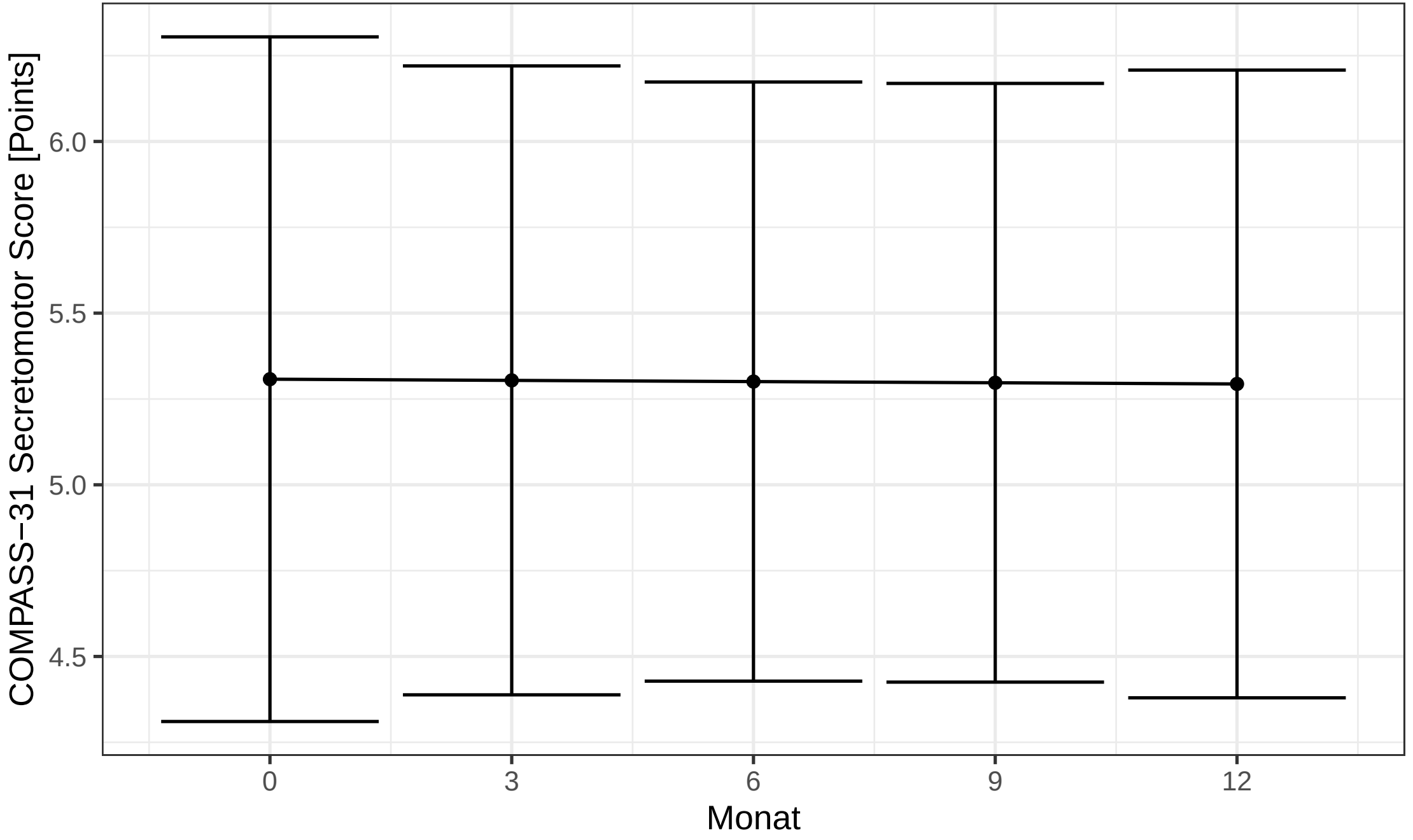
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

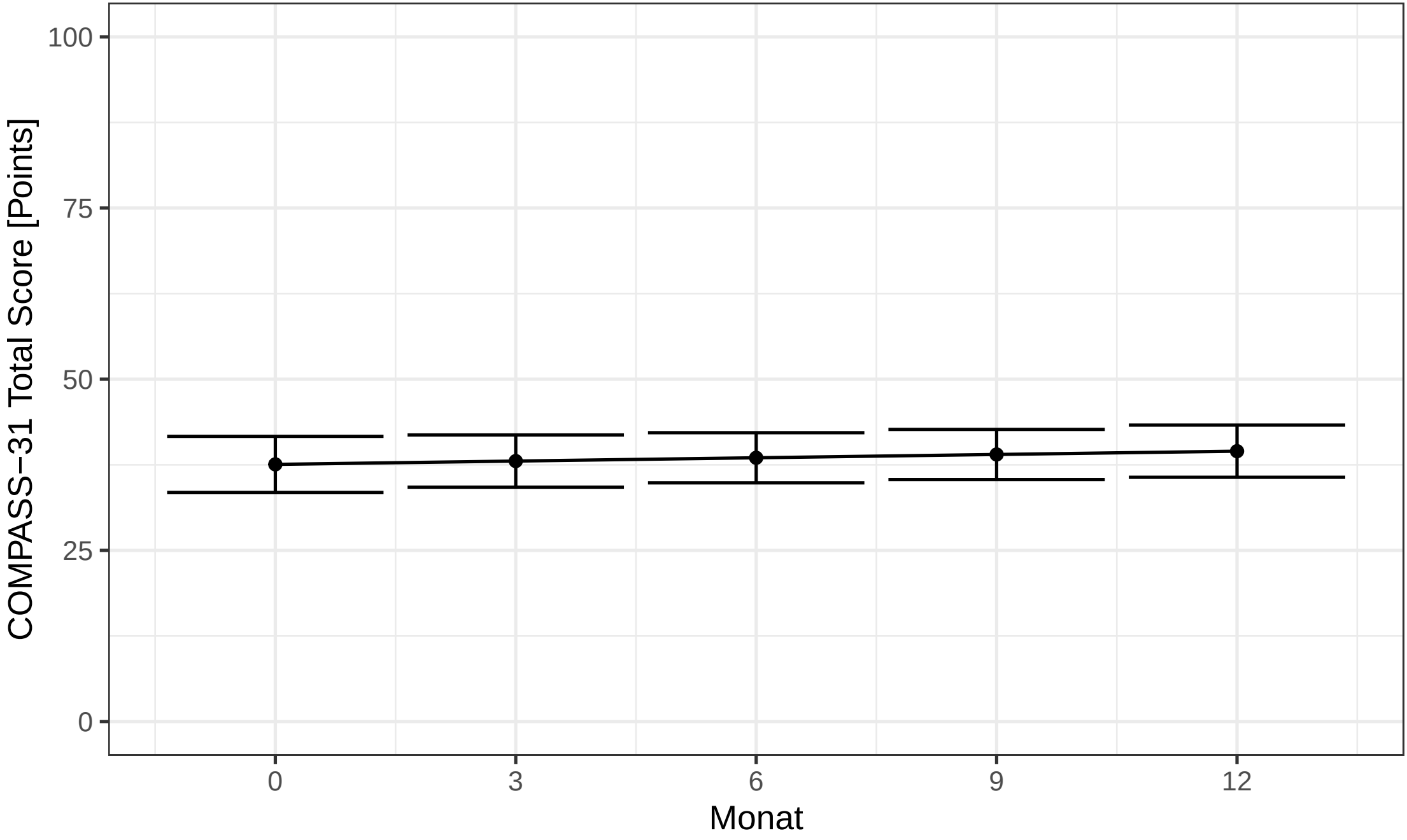
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

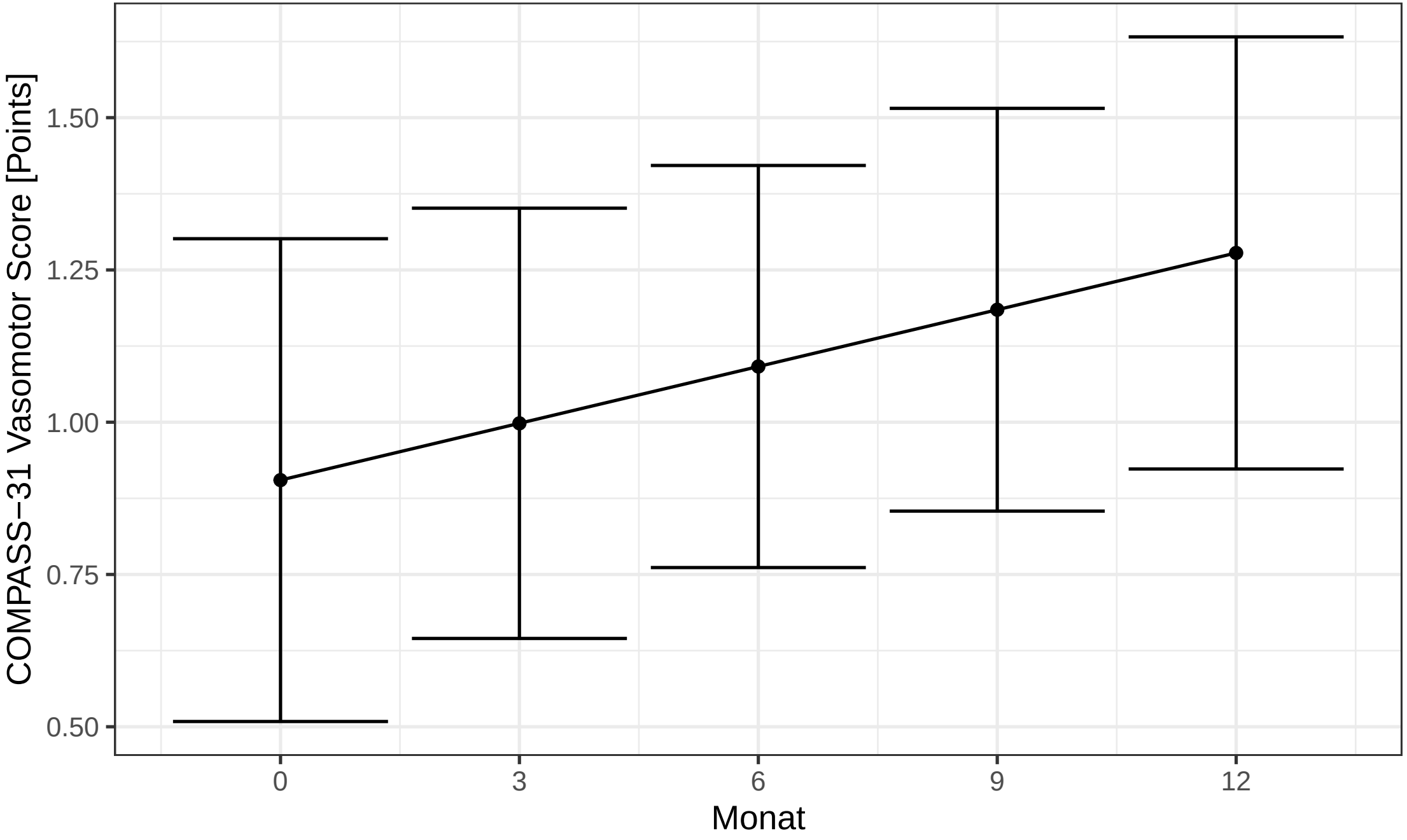
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

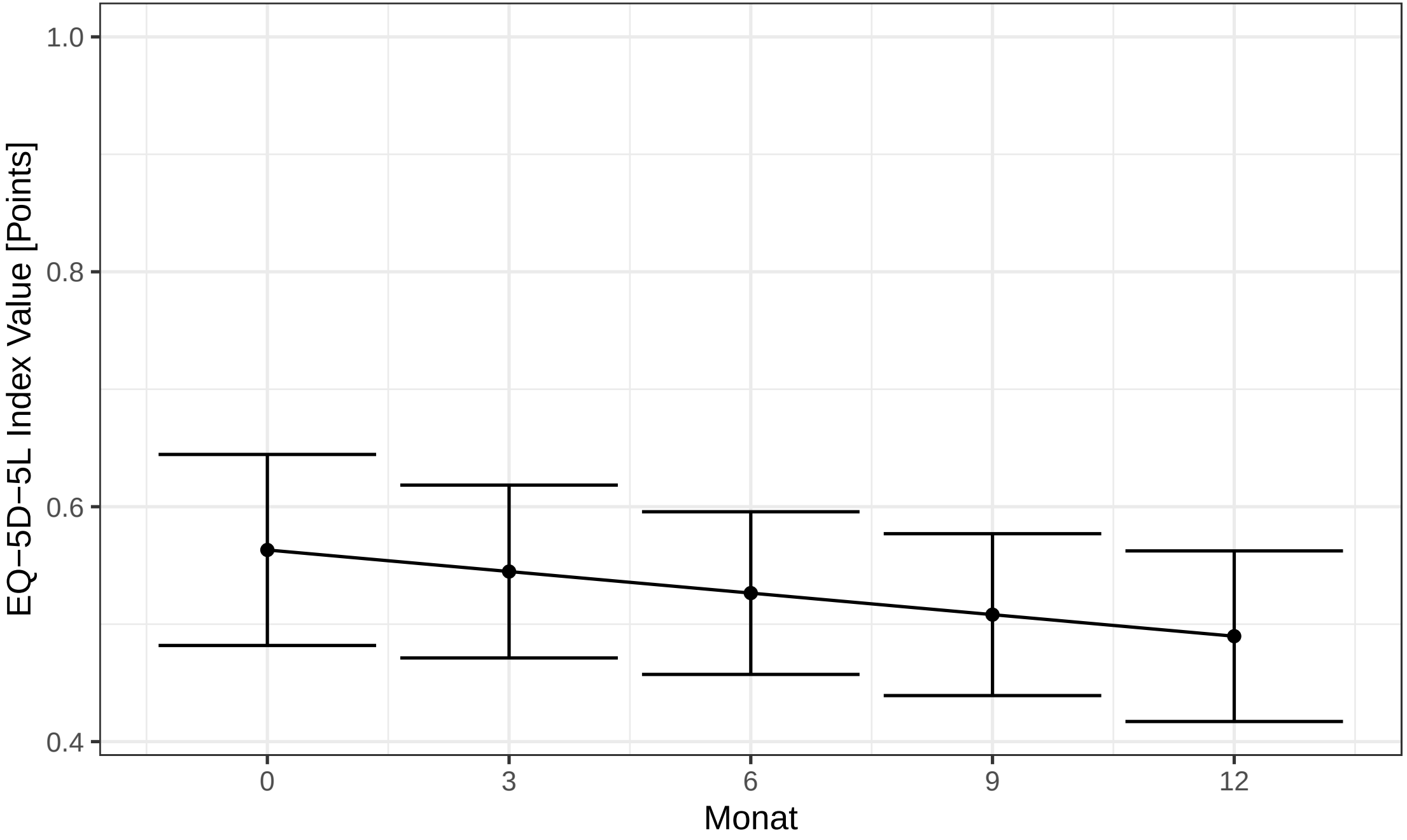
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

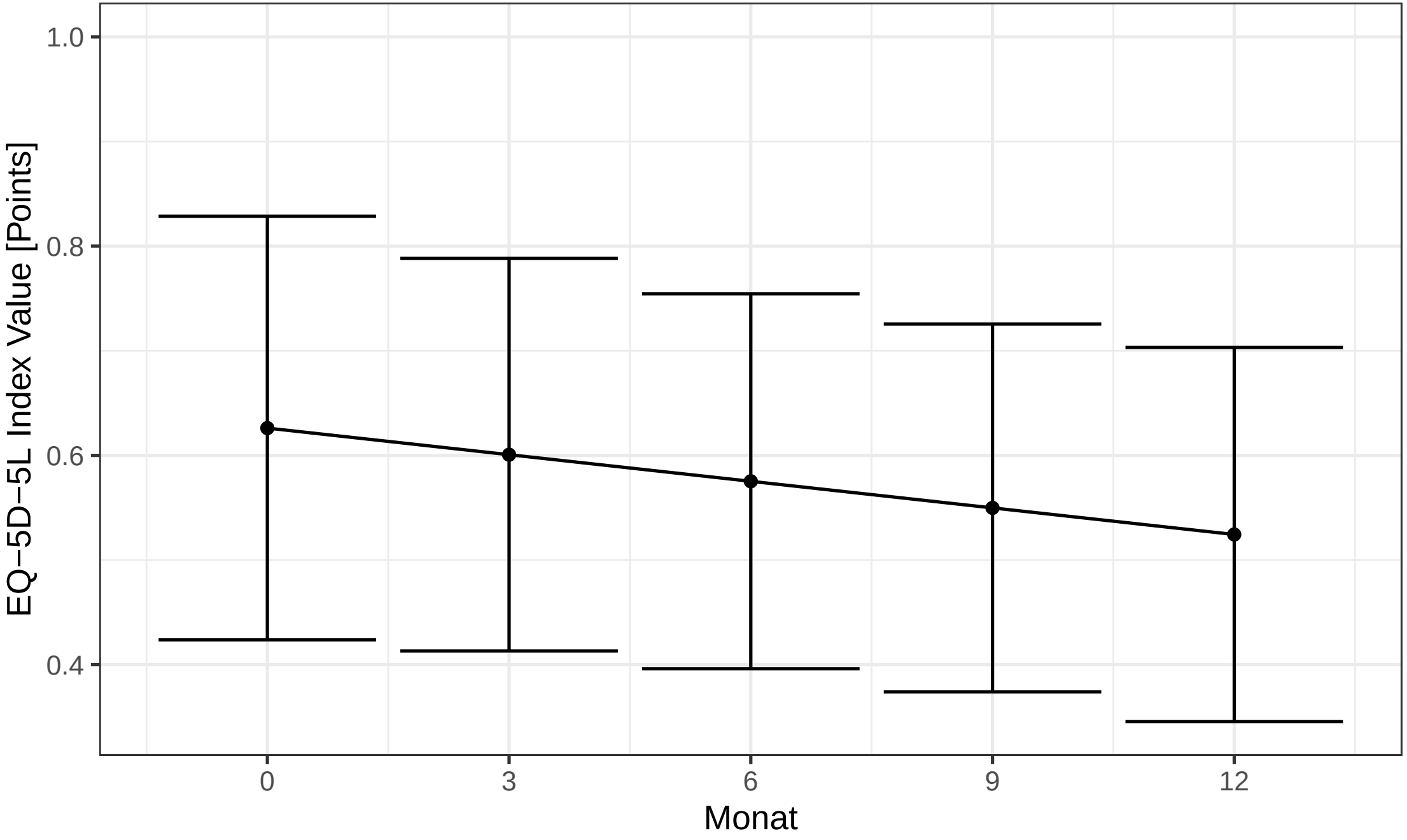
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

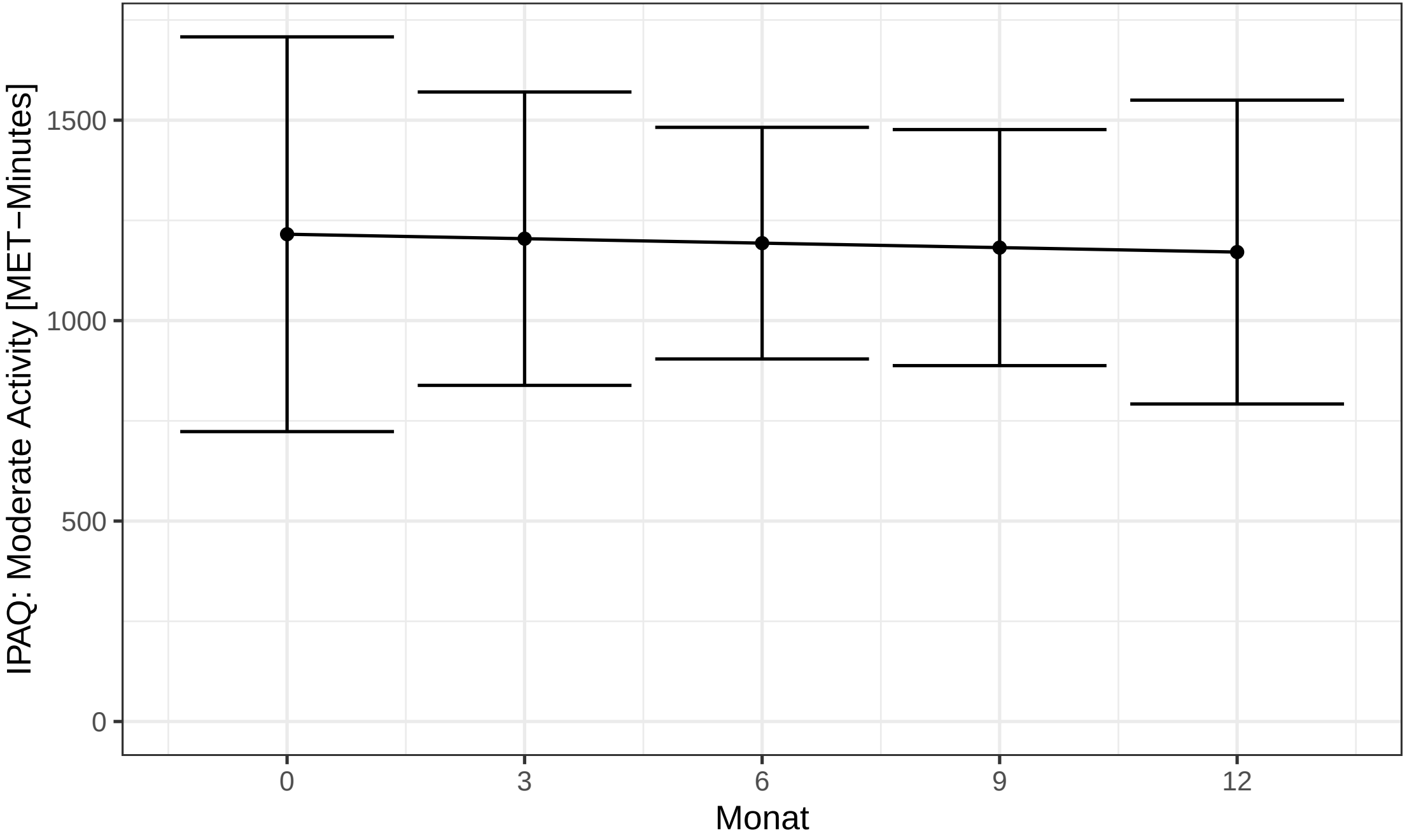
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

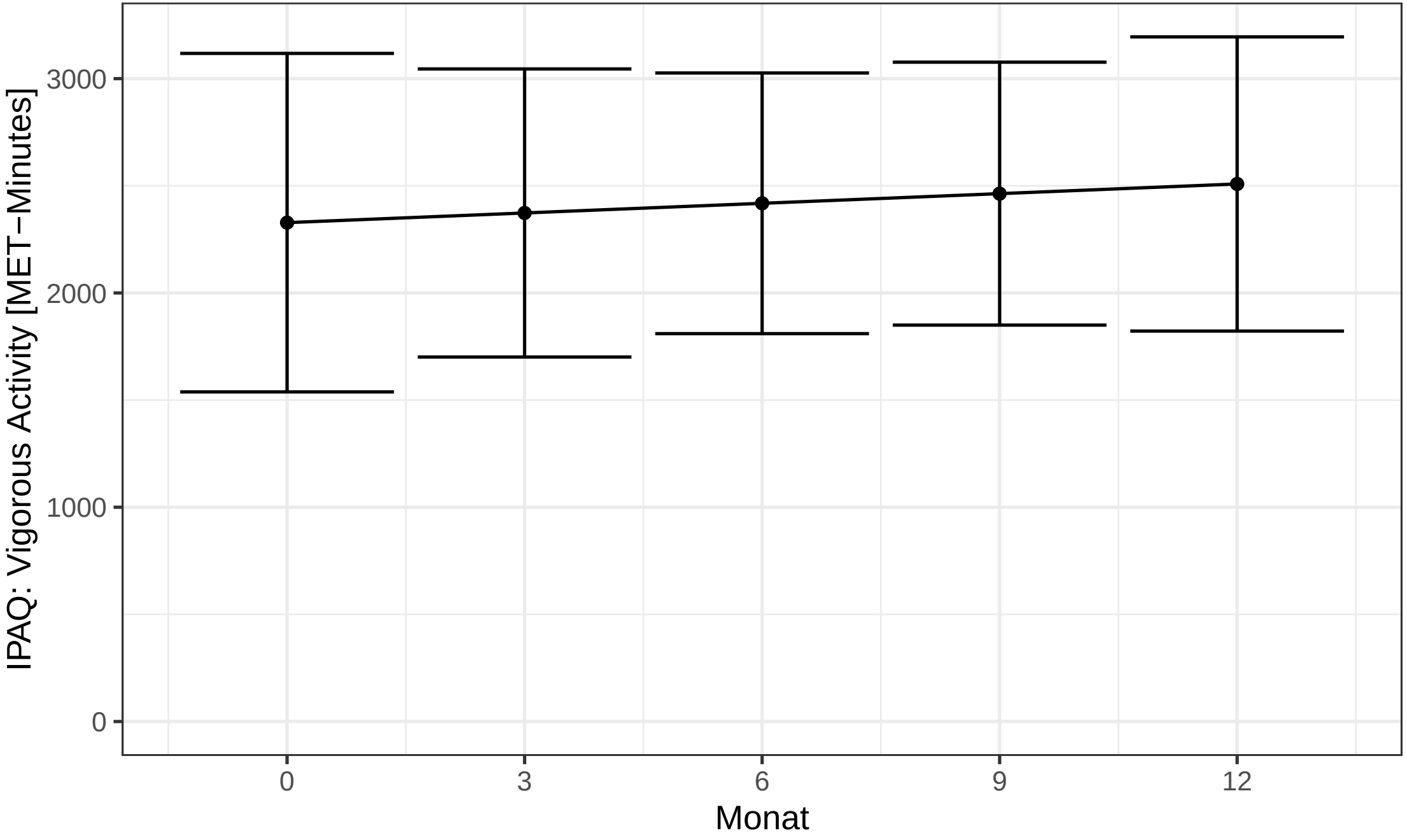
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

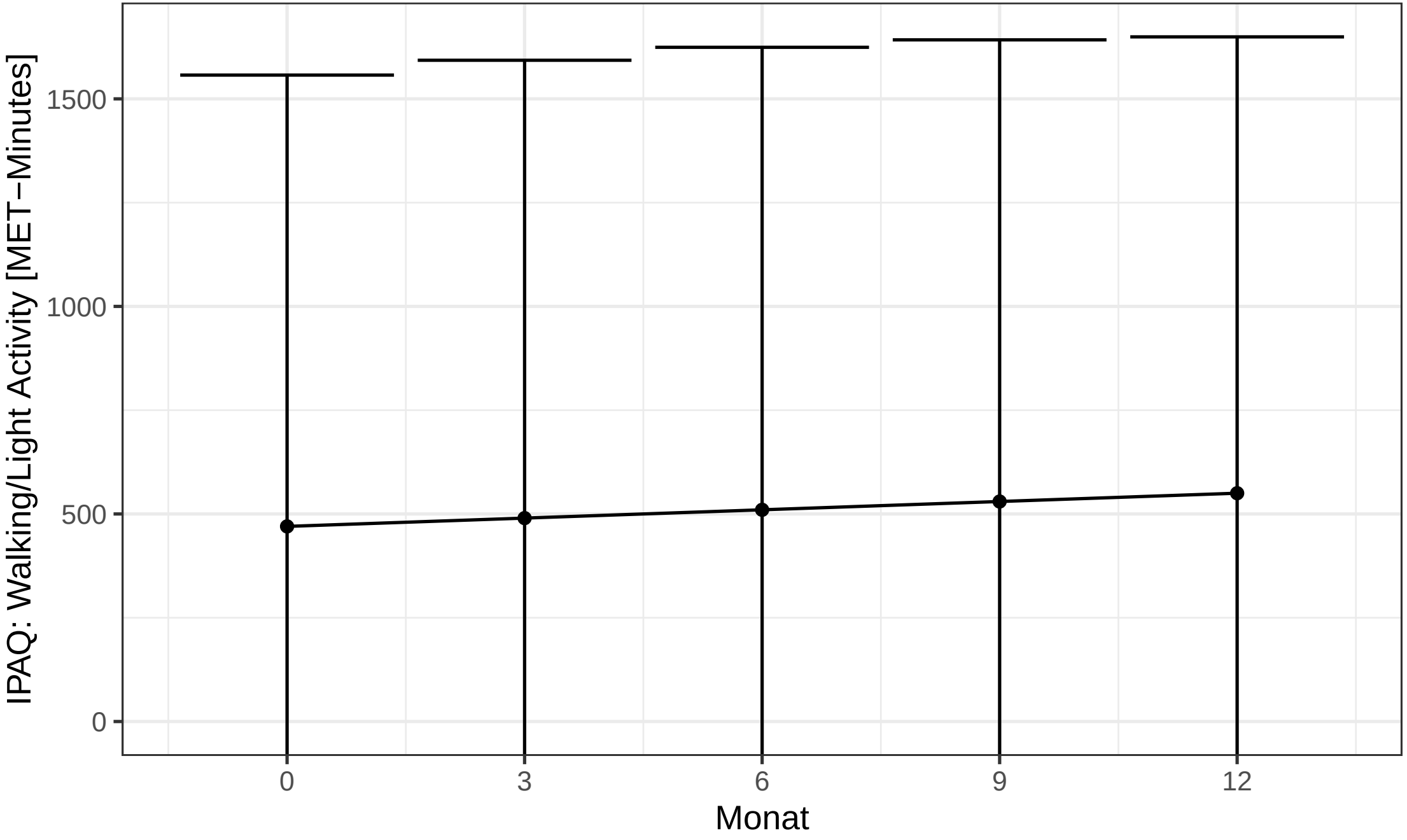
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

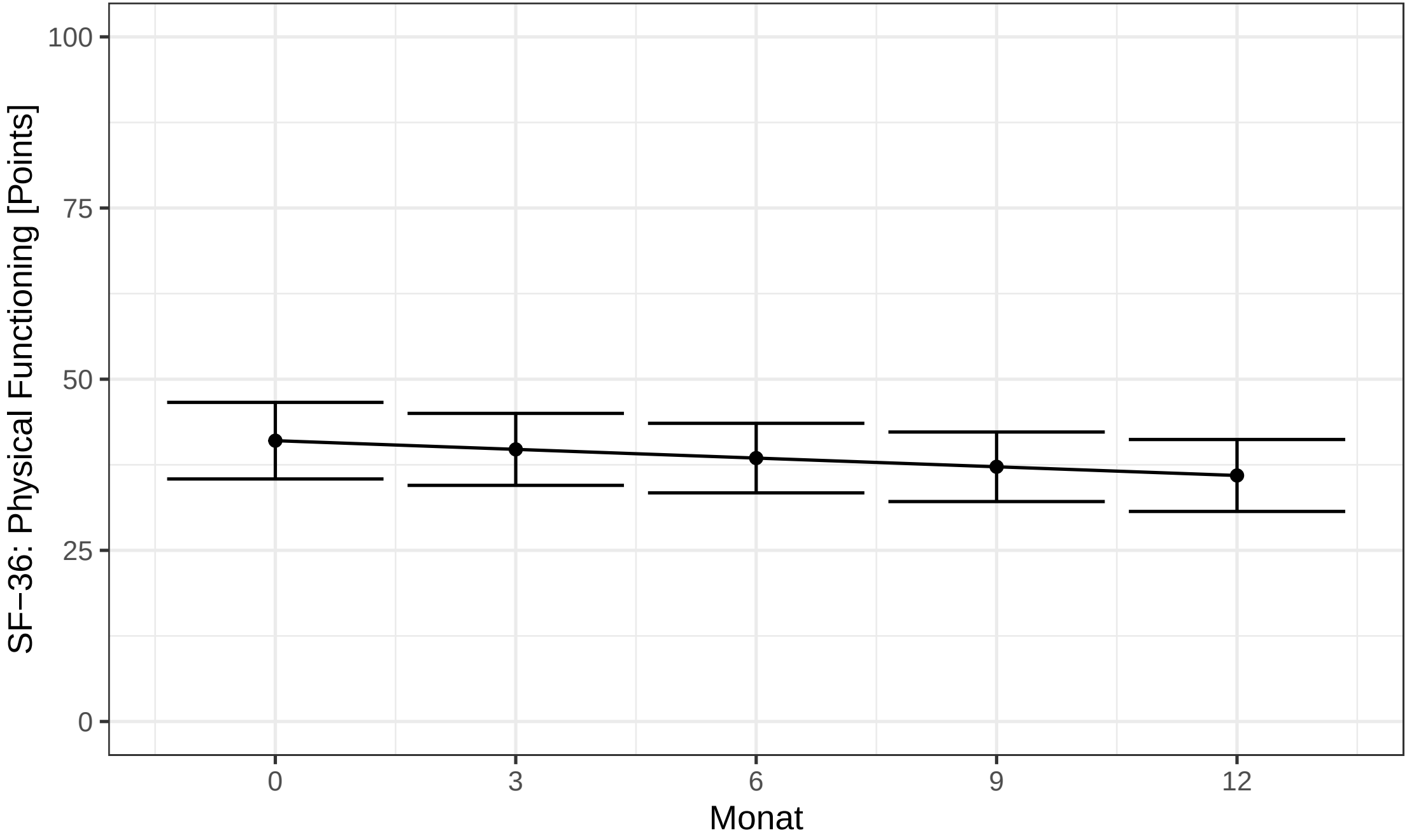
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

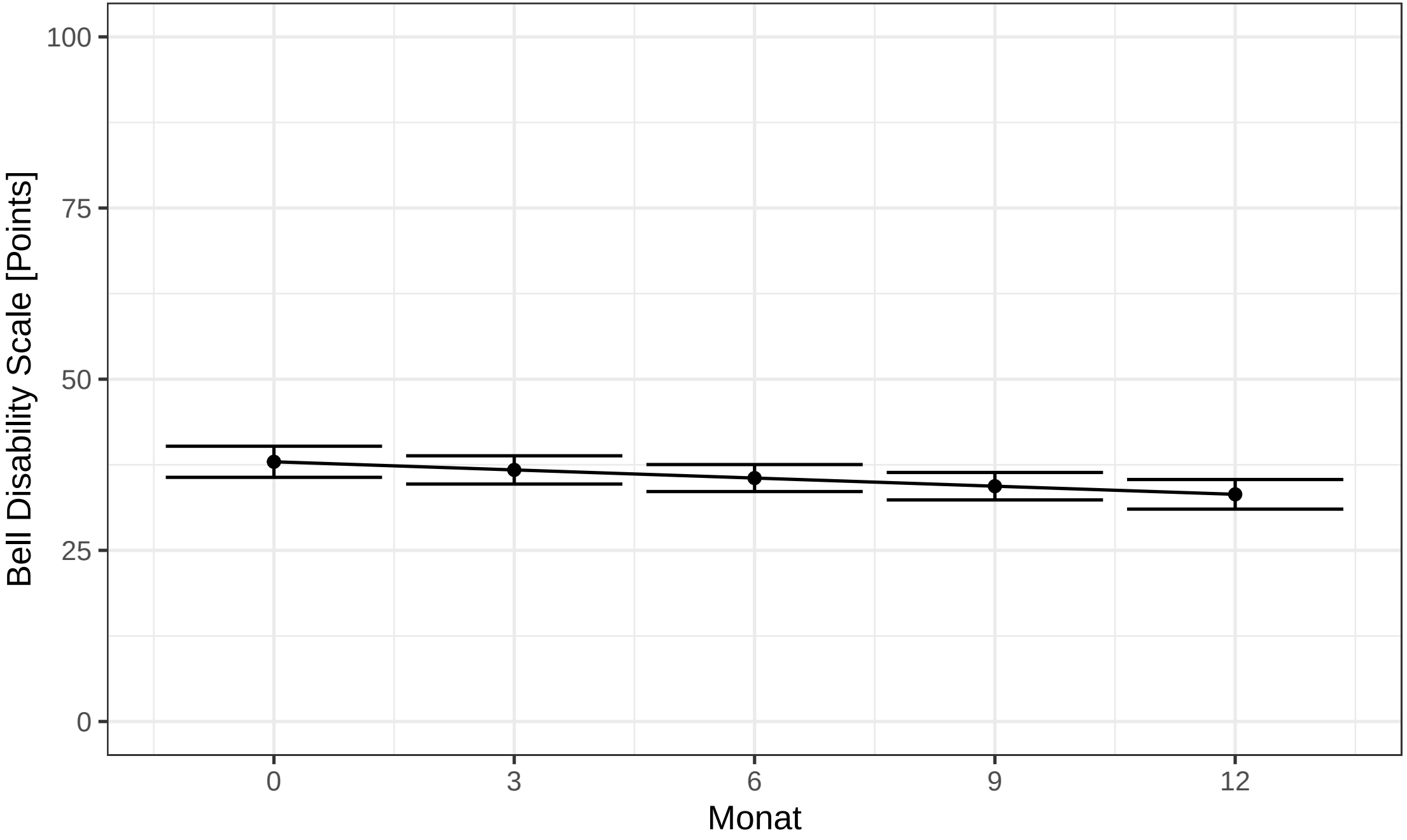
Adjustierter Mittelwert



Population B

Verlauf Interventionsgruppe

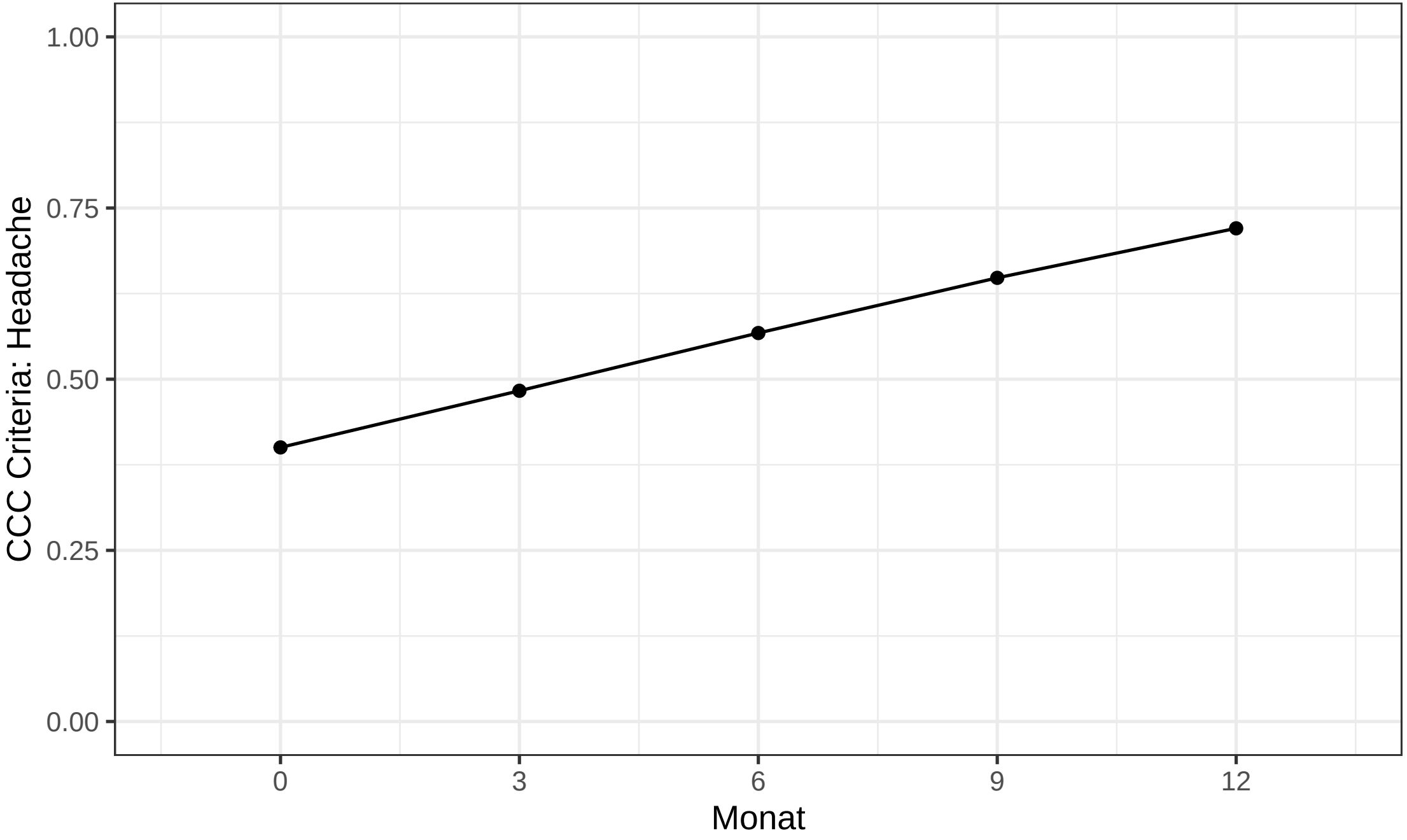
Adjustierter Mittelwert



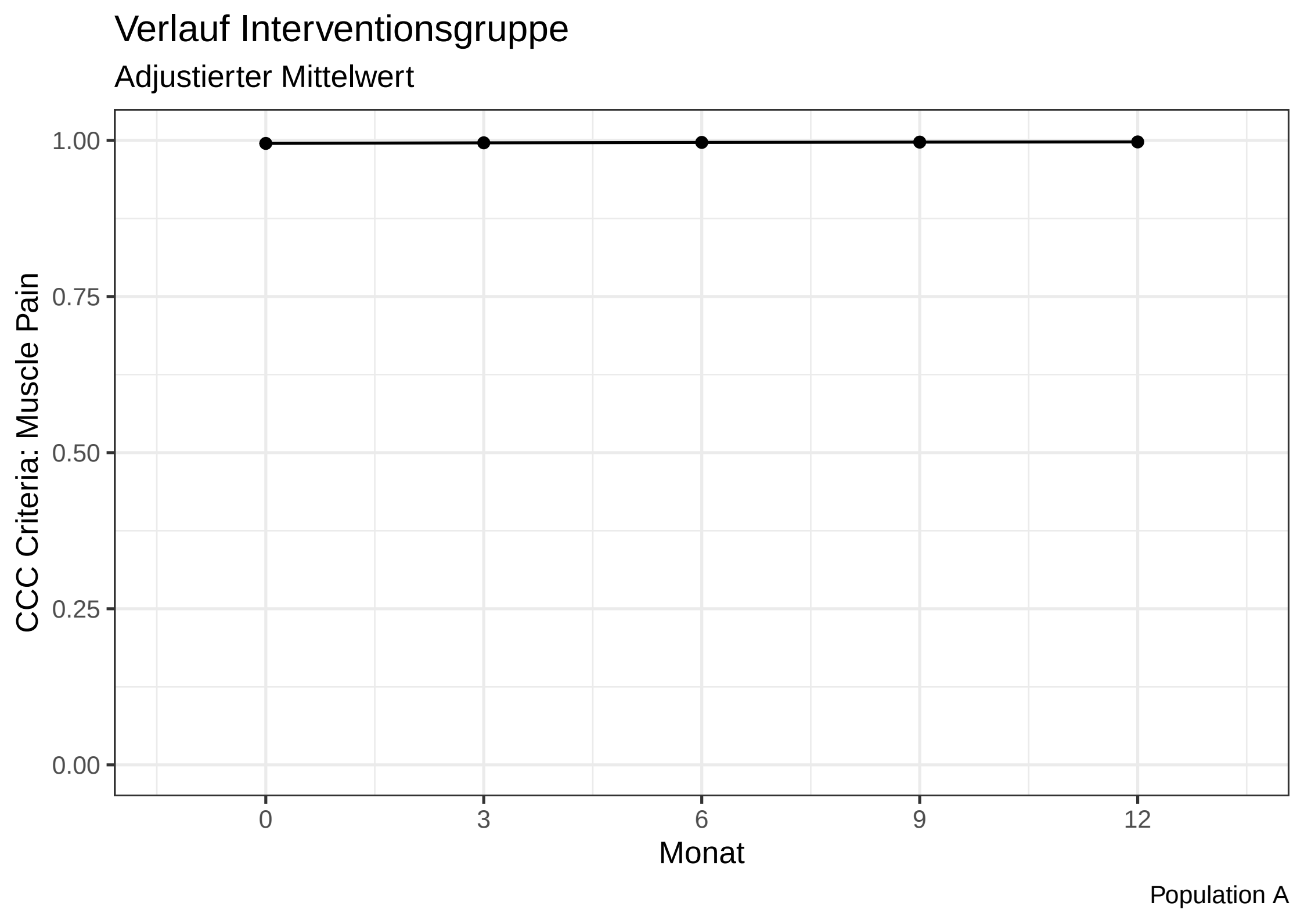
Population A

Verlauf Interventionsgruppe

Adjustierter Mittelwert

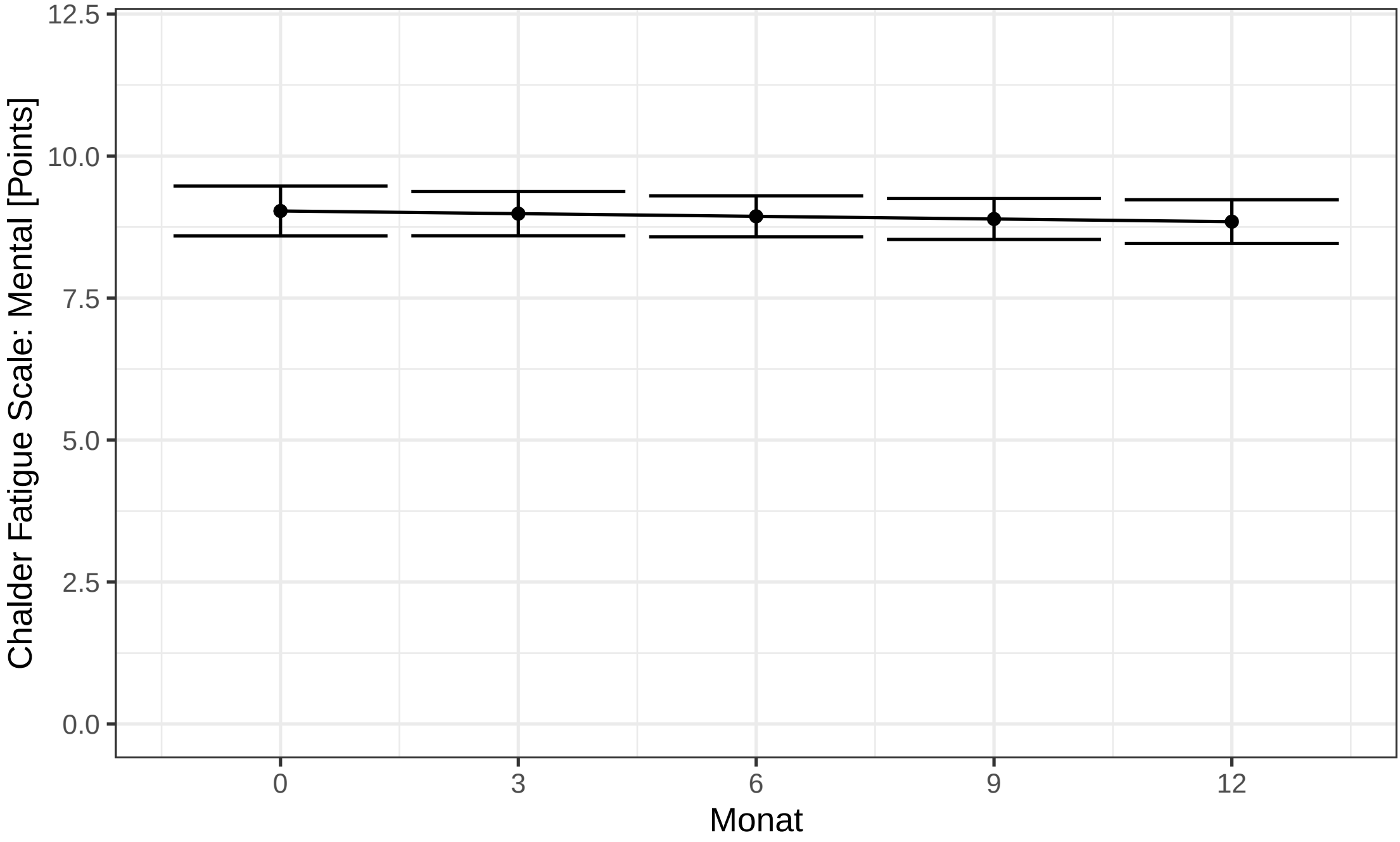


Population A


Verlauf Interventionsgruppe
Adjustierter Mittelwert
CCC Criteria: Muscle Pain
1.00
0.75
0.50
0.25
0.00
0
3
6
9
12
Monat
Population A

Verlauf Interventionsgruppe

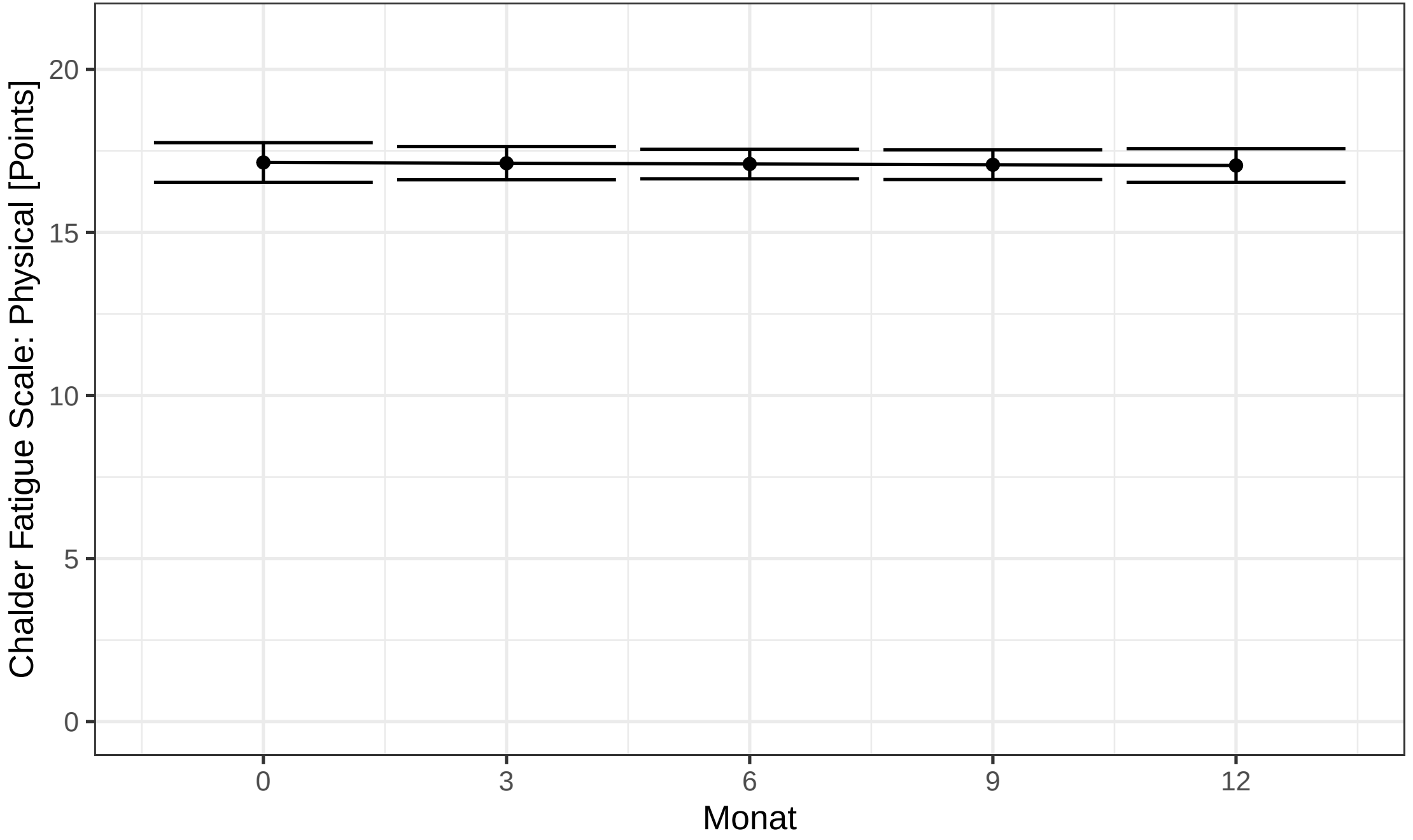
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

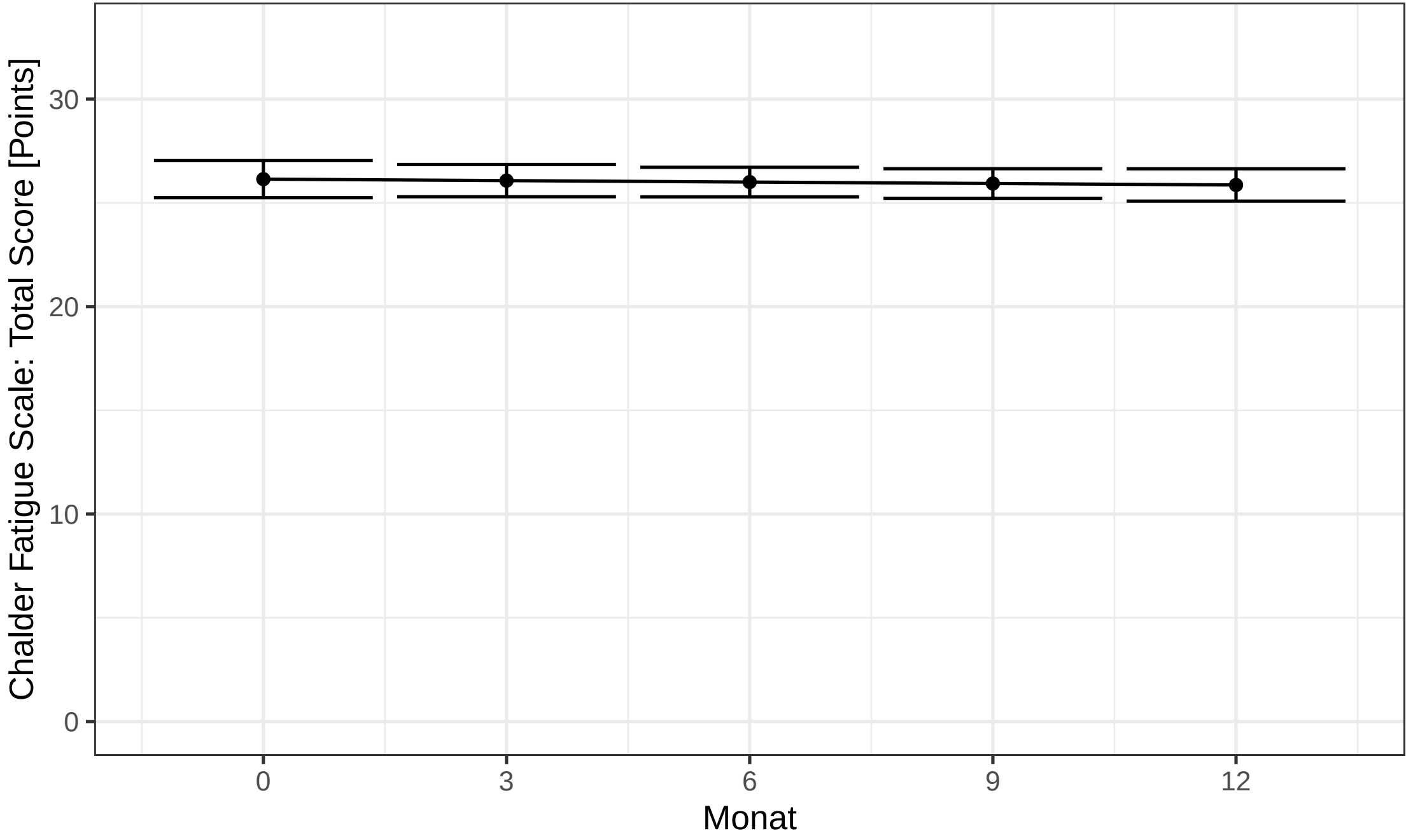
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

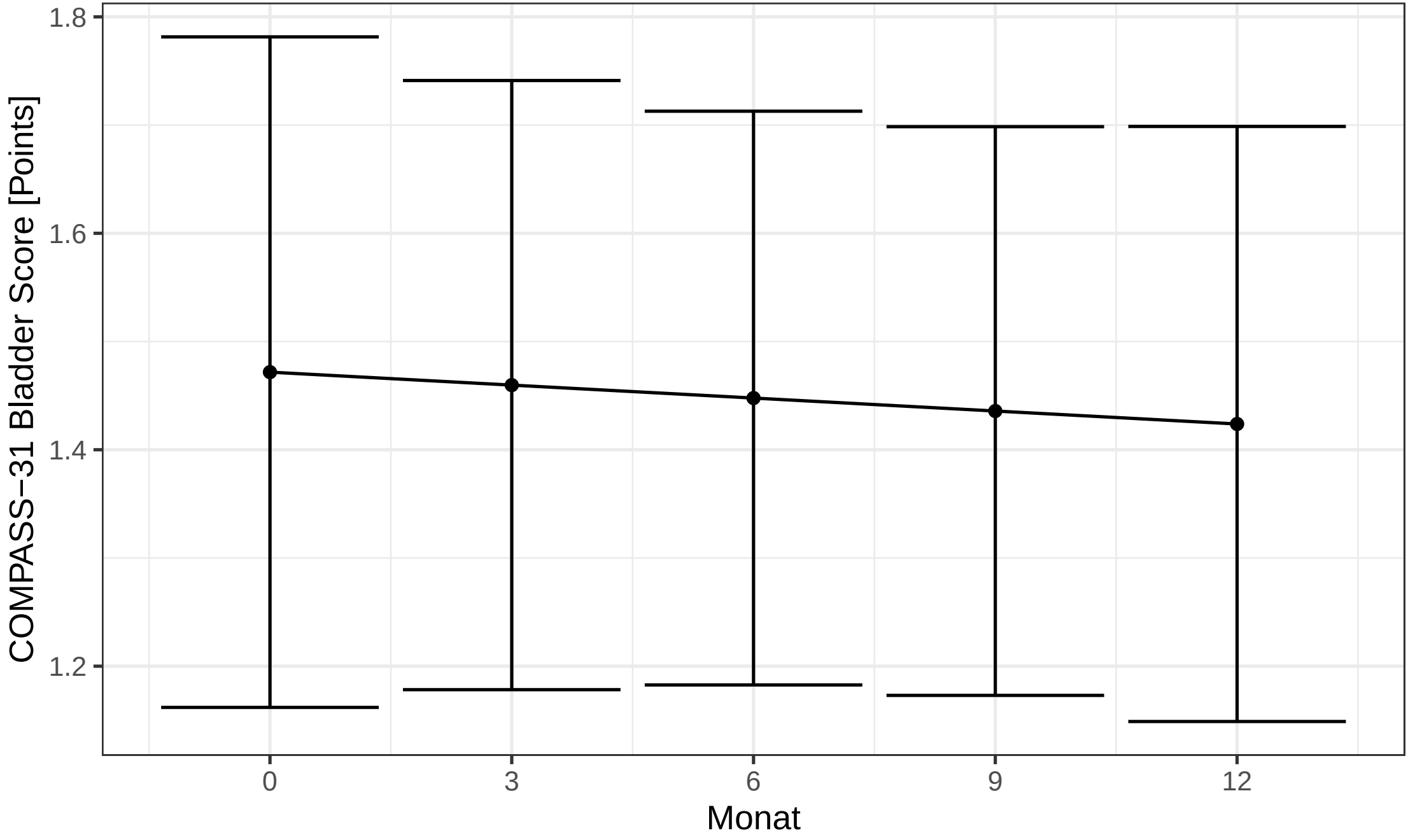
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

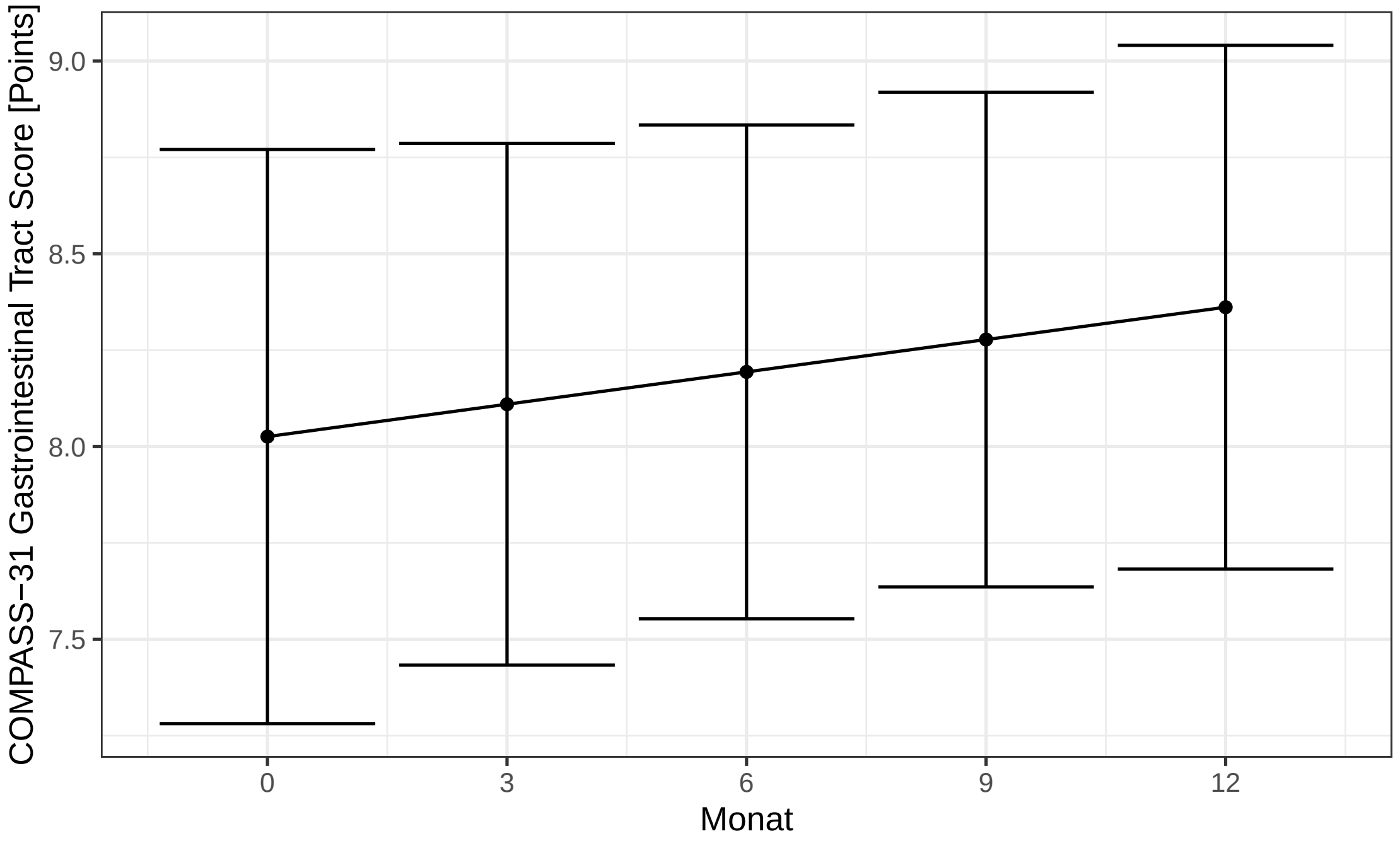
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

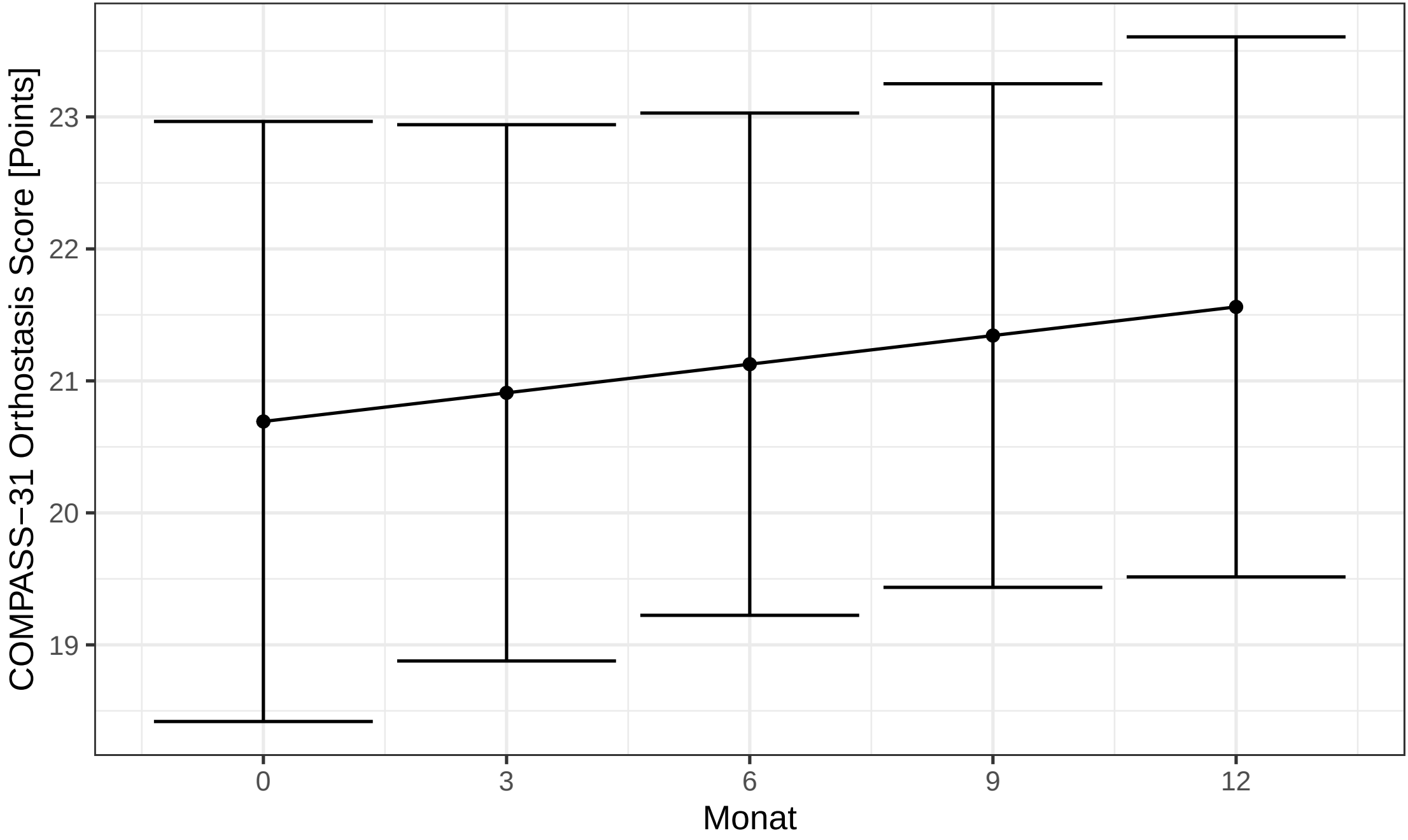
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

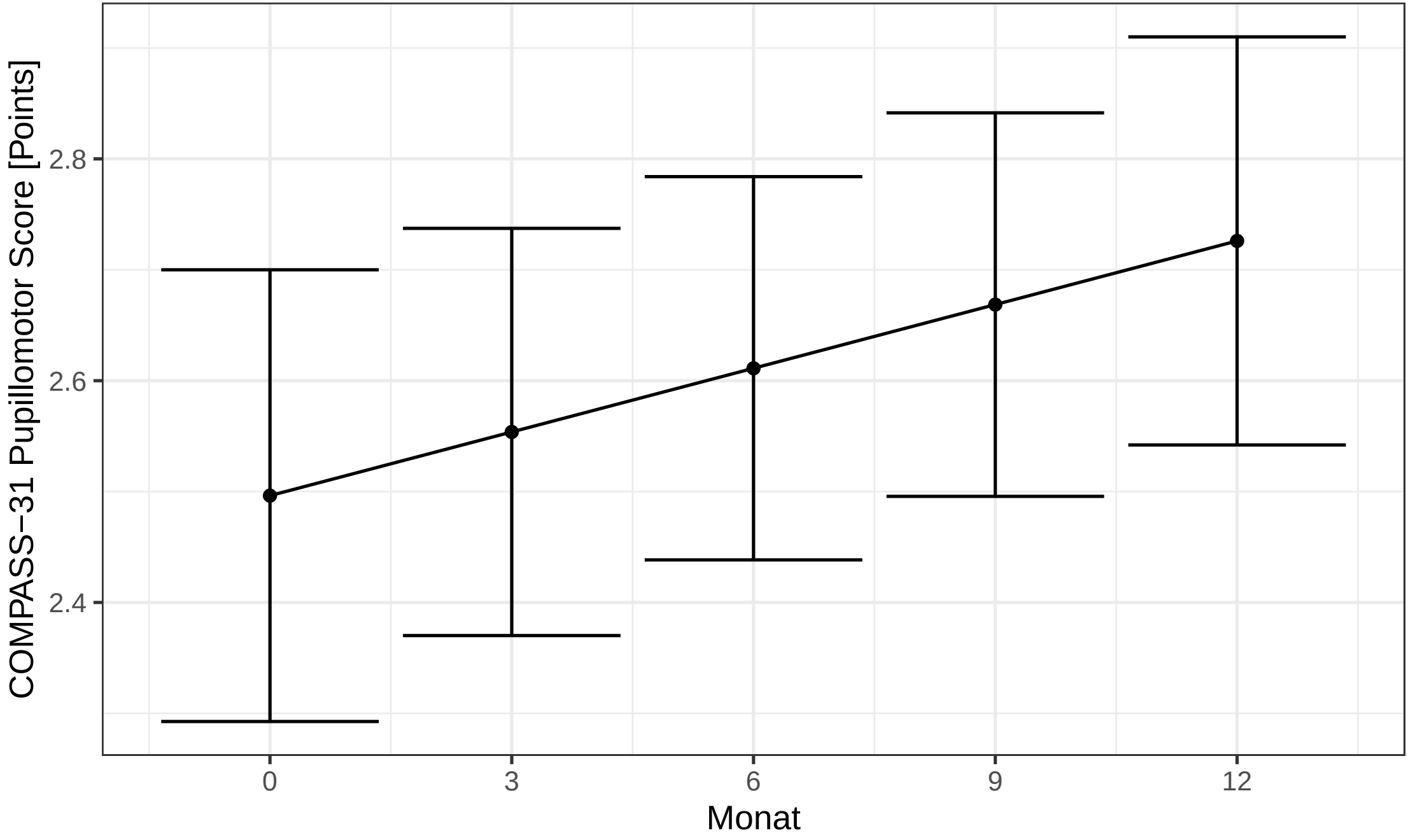
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

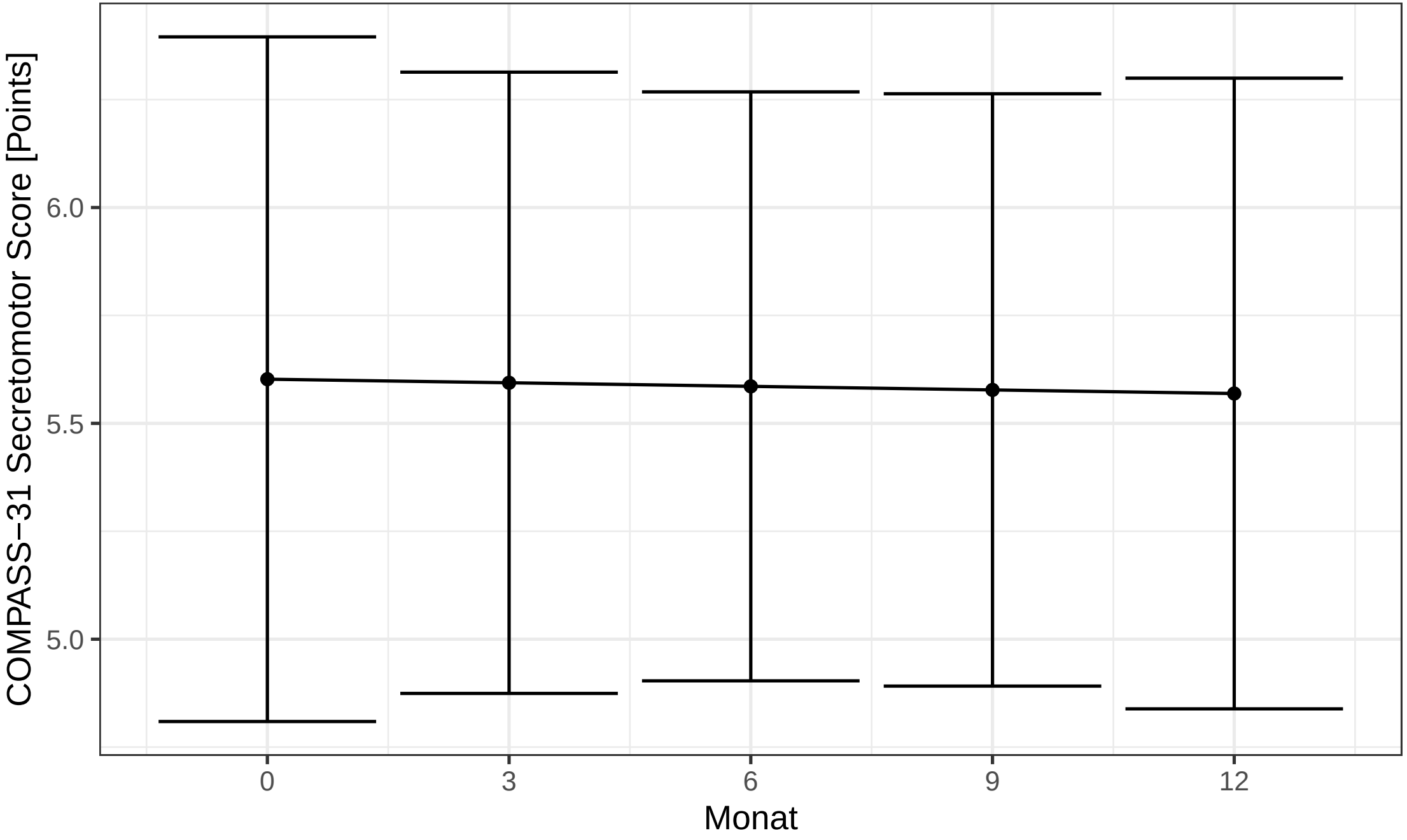
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

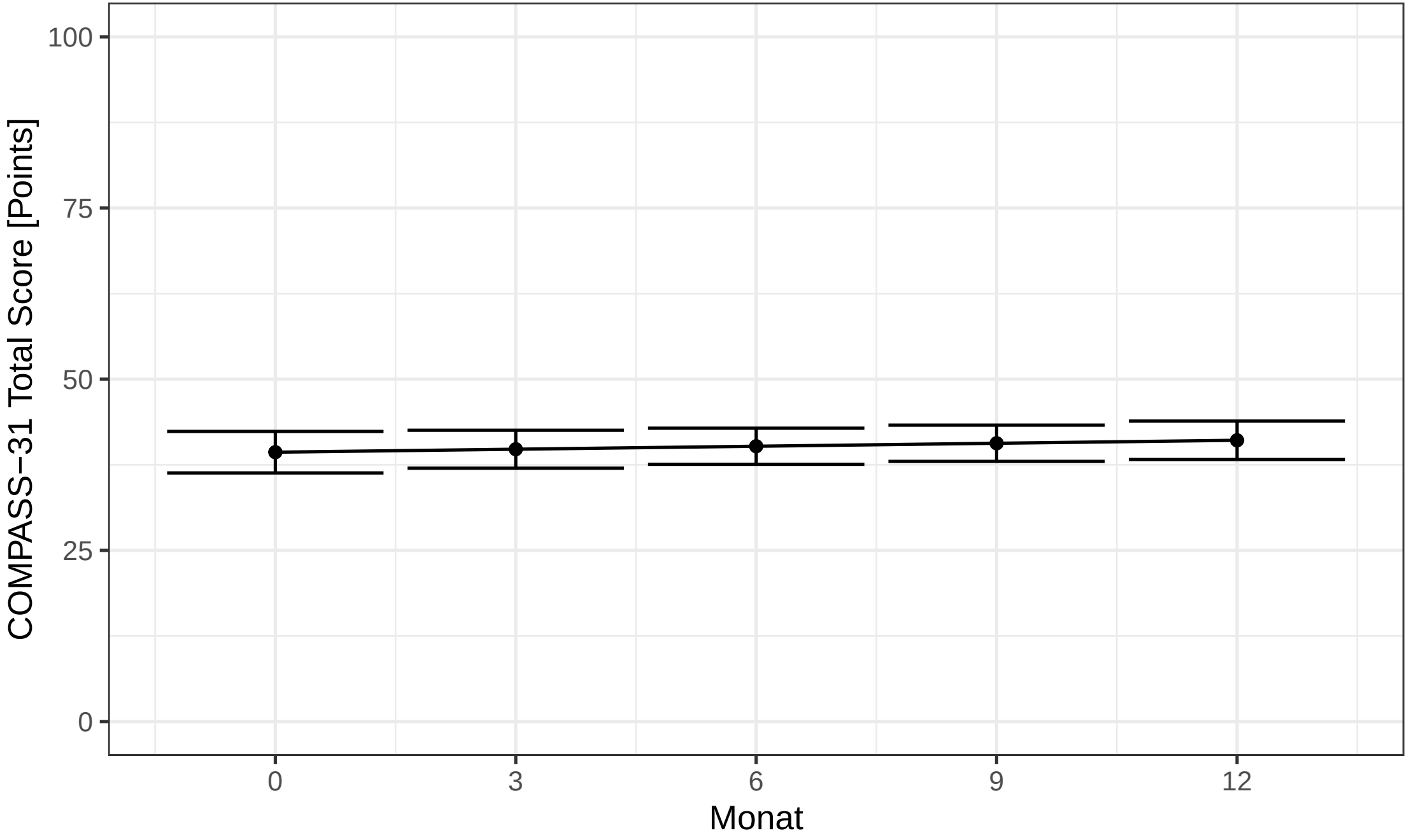
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

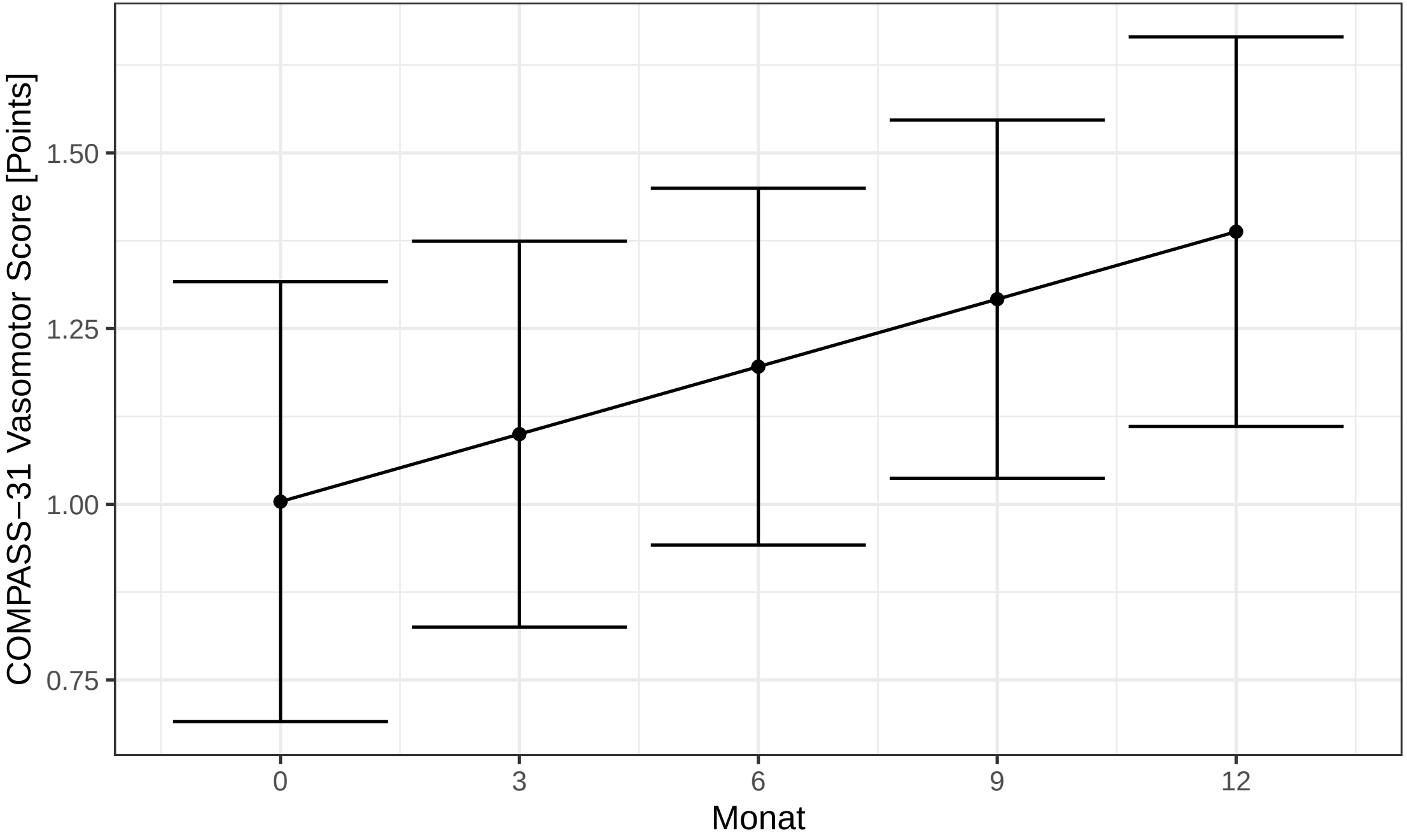
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

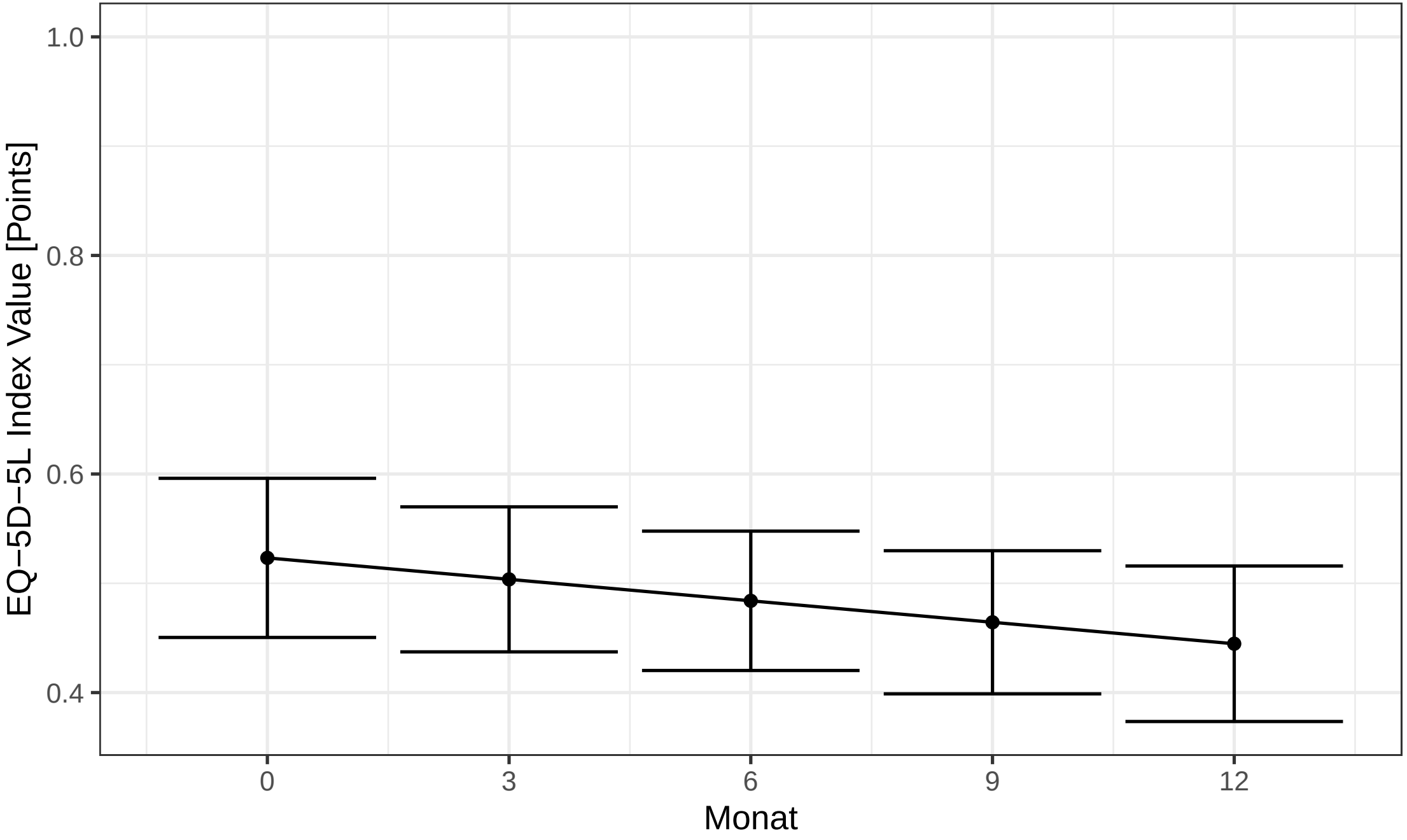
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

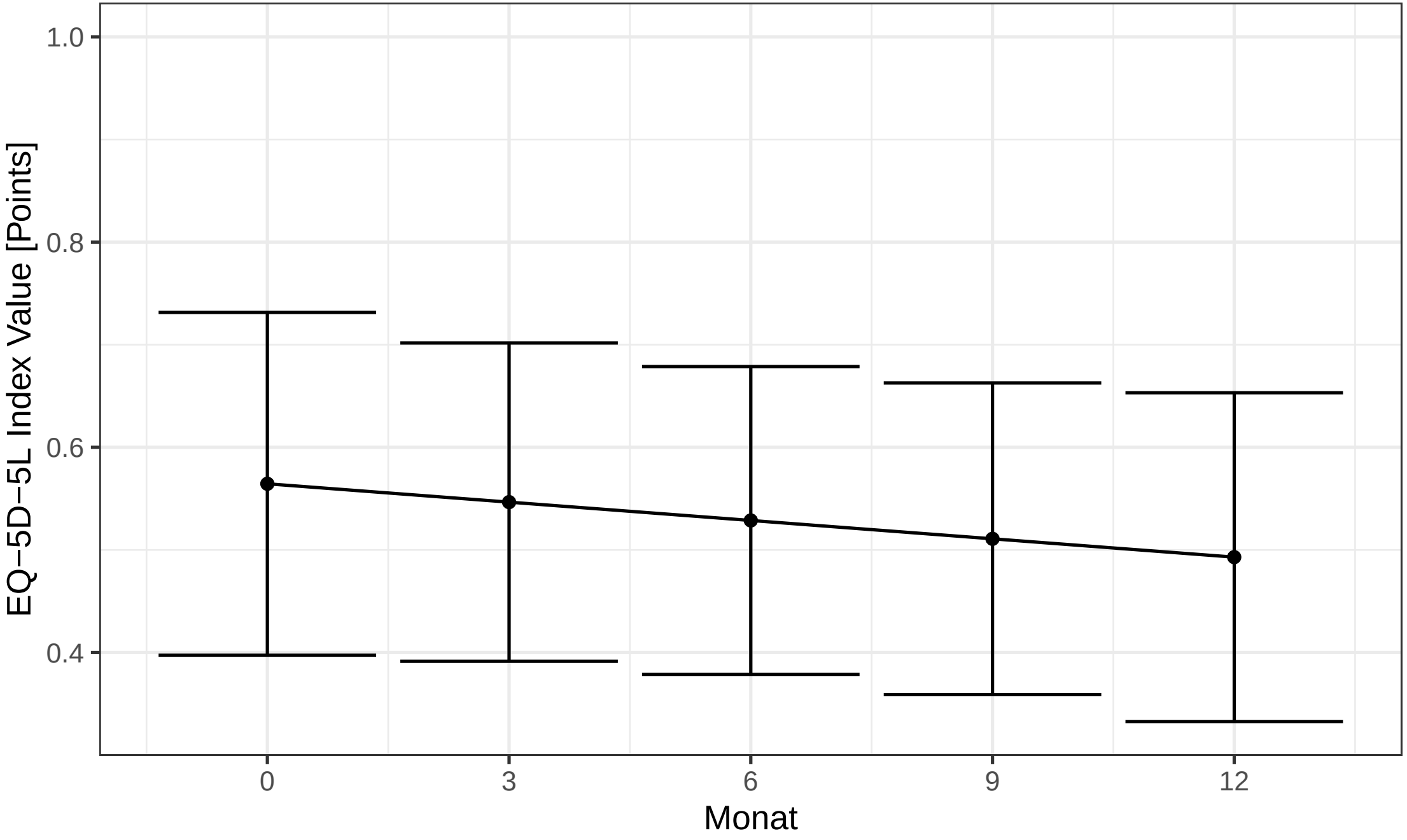
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

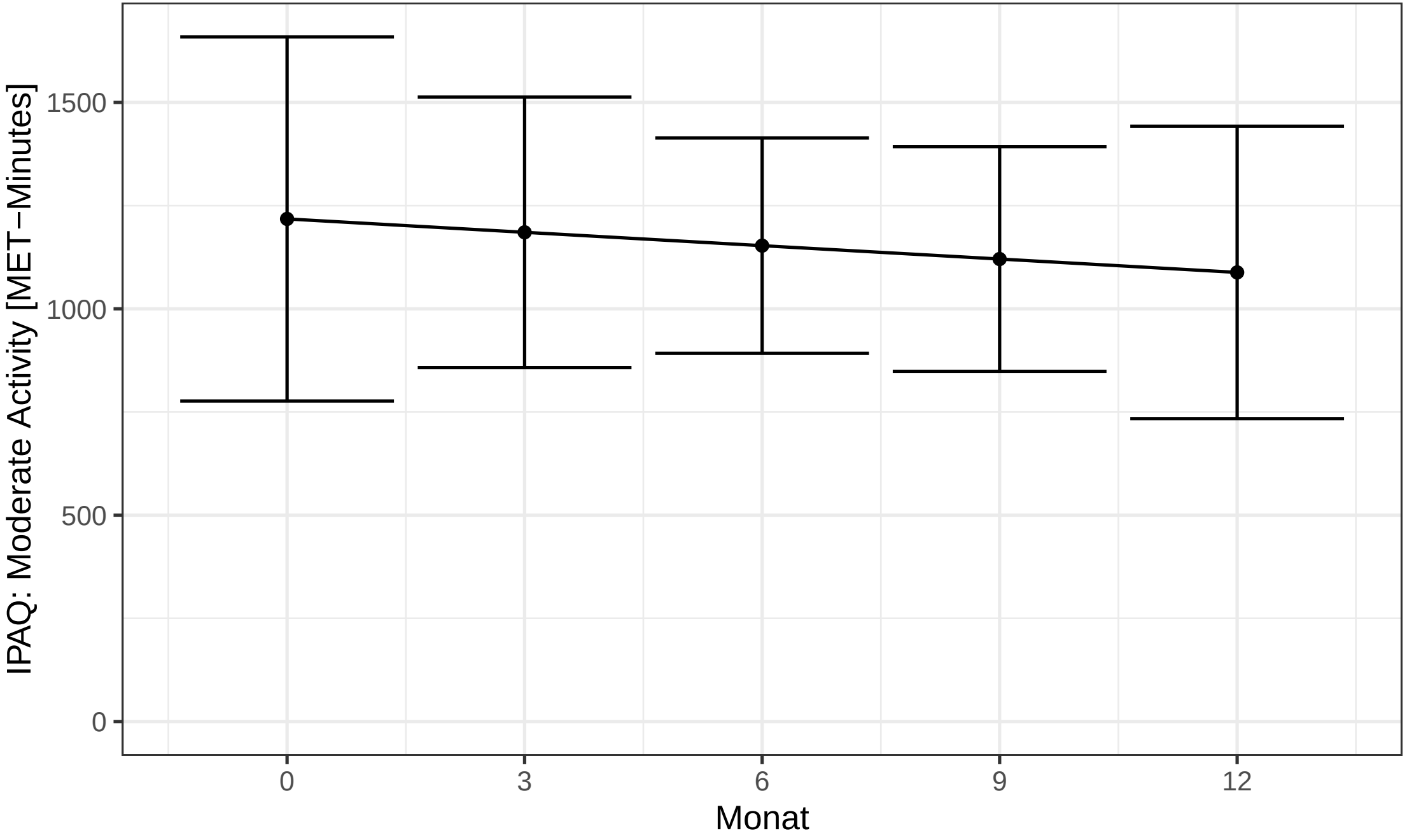
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

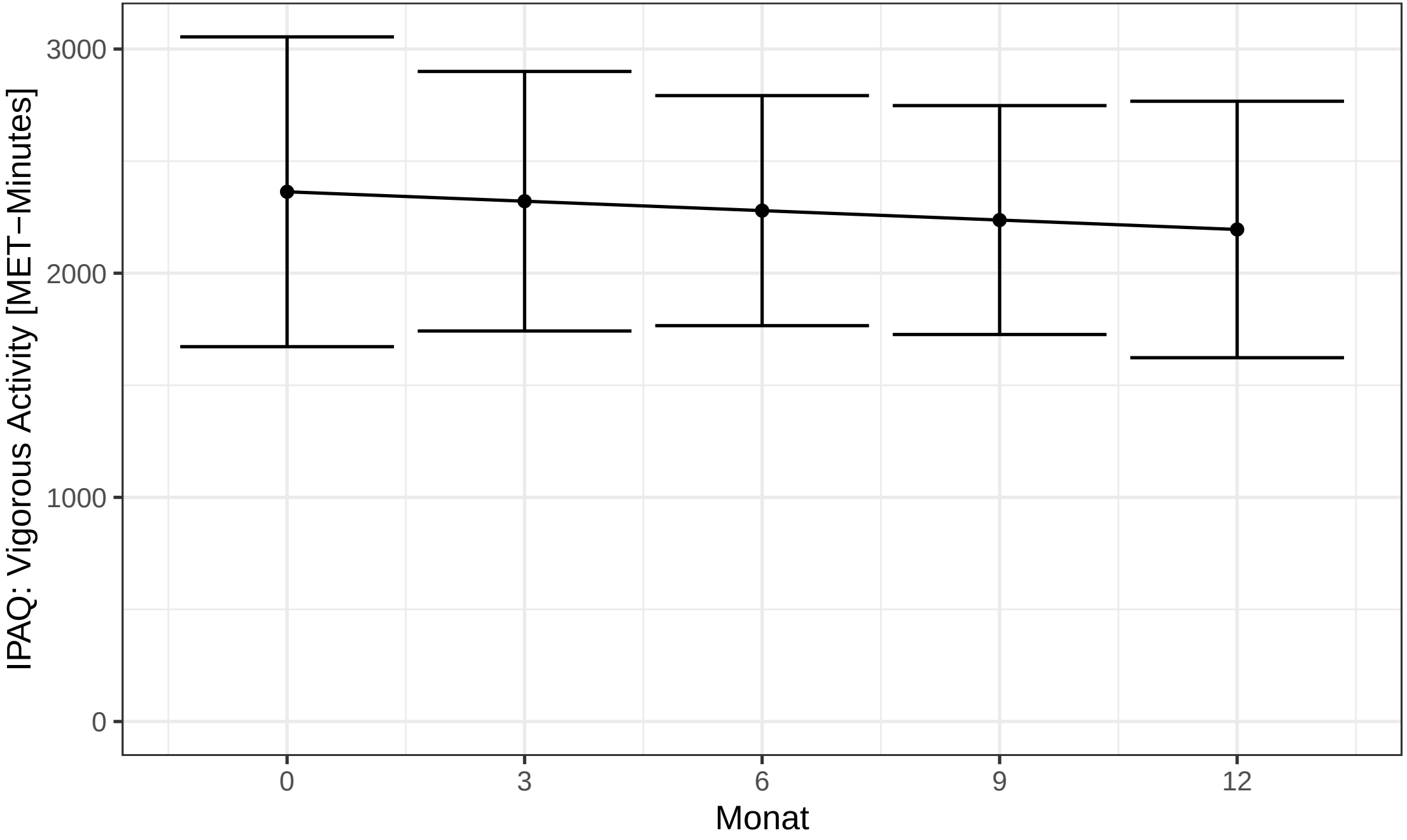
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

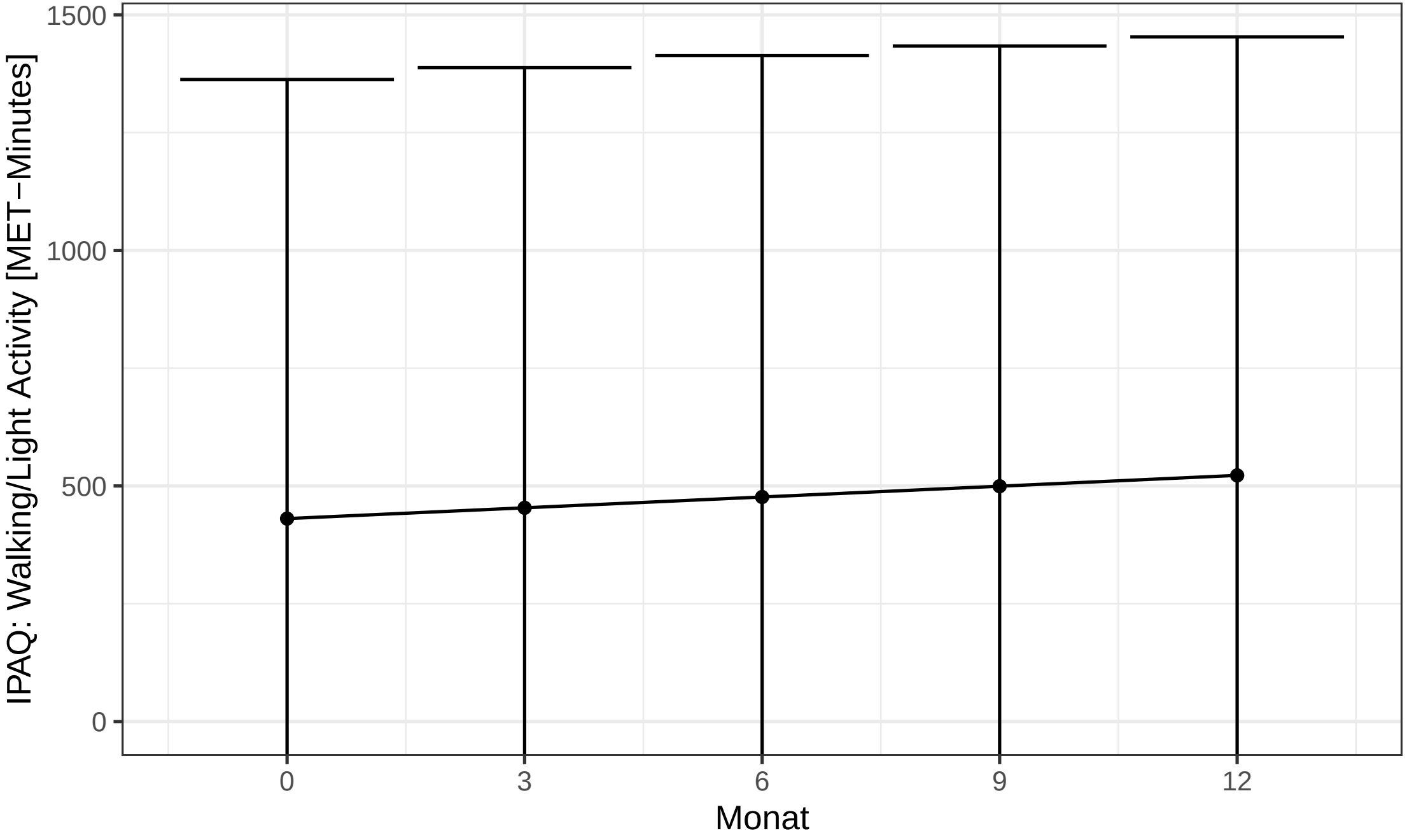
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

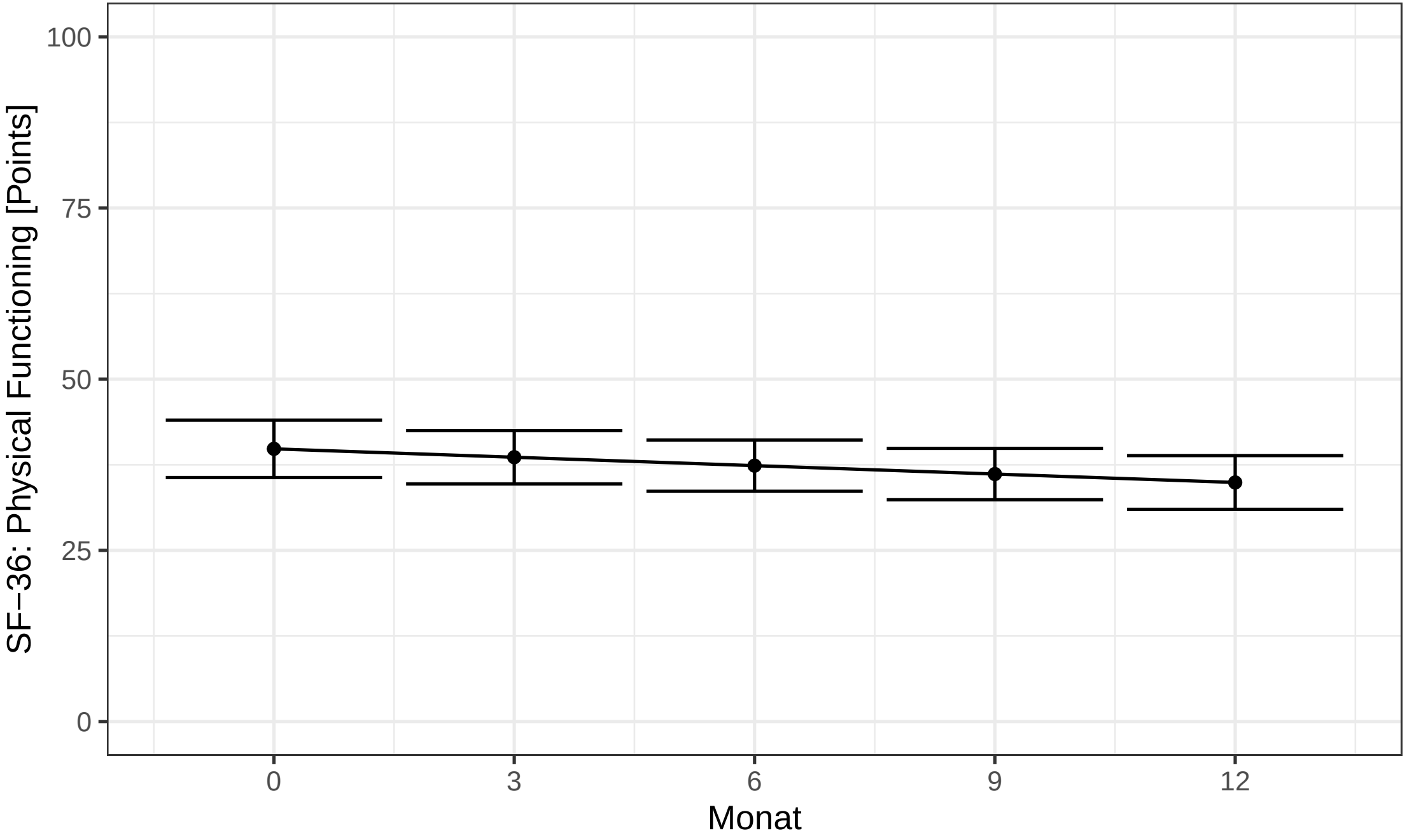
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

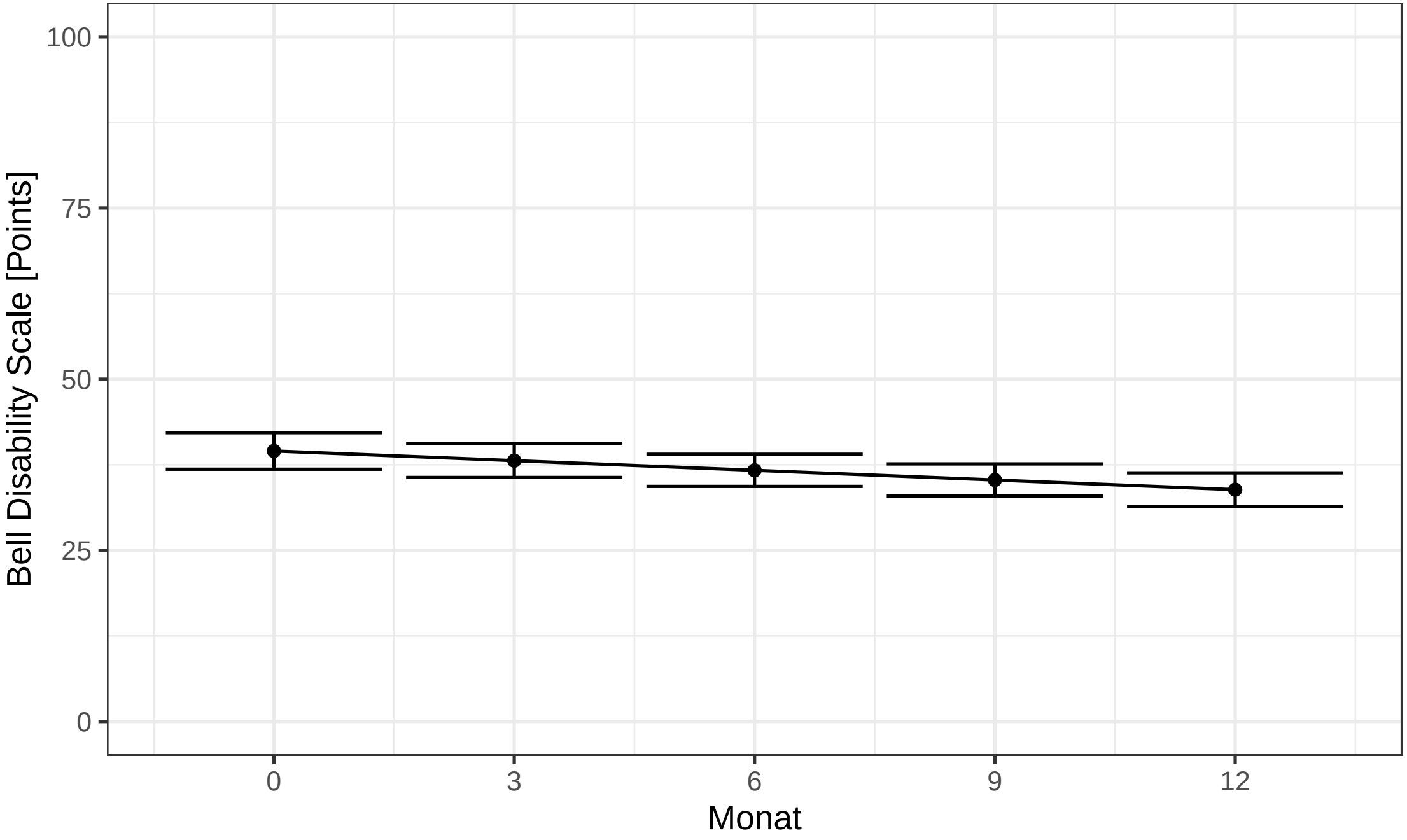
Adjustierter Mittelwert



Population A

Verlauf Interventionsgruppe

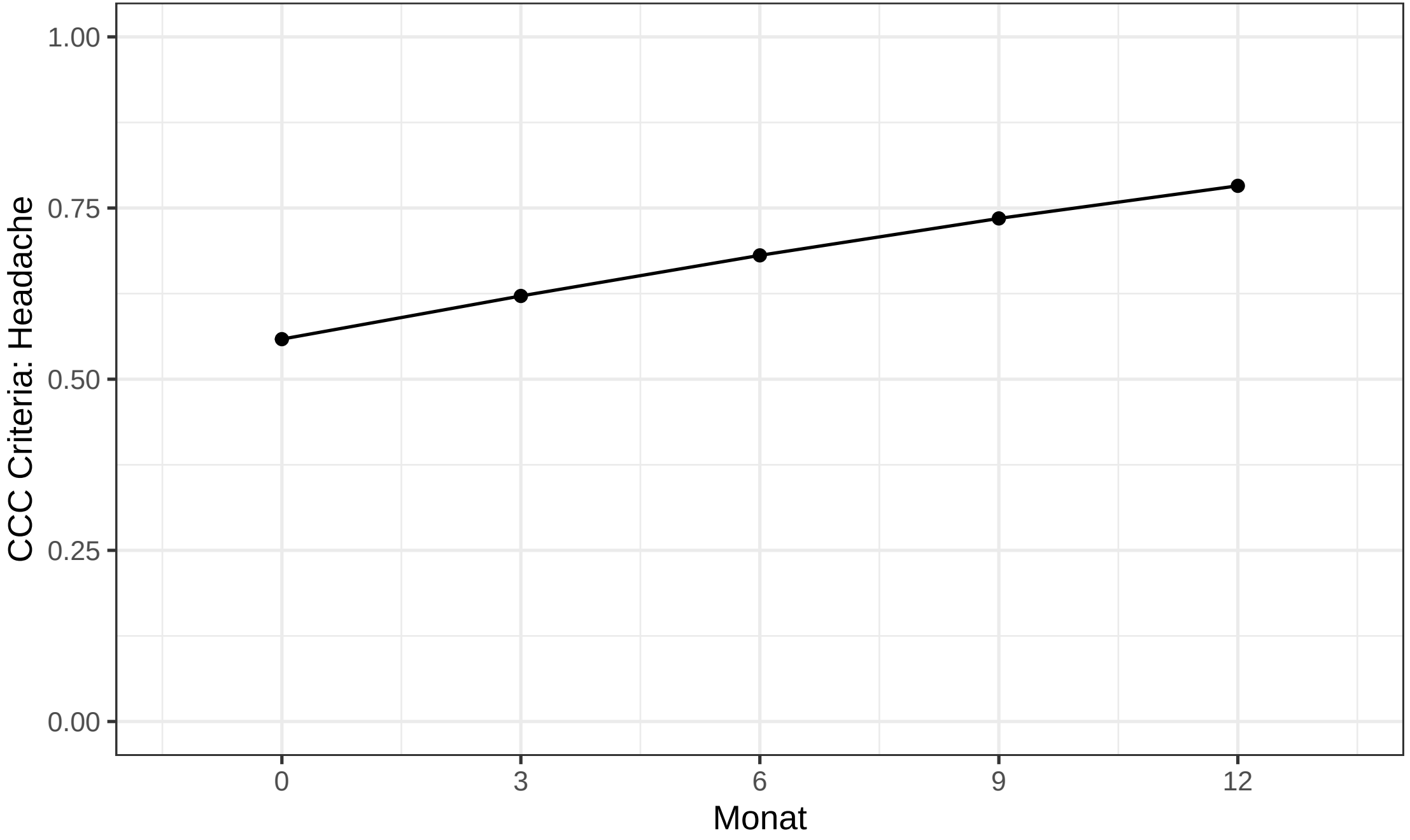
Adjustierter Mittelwert



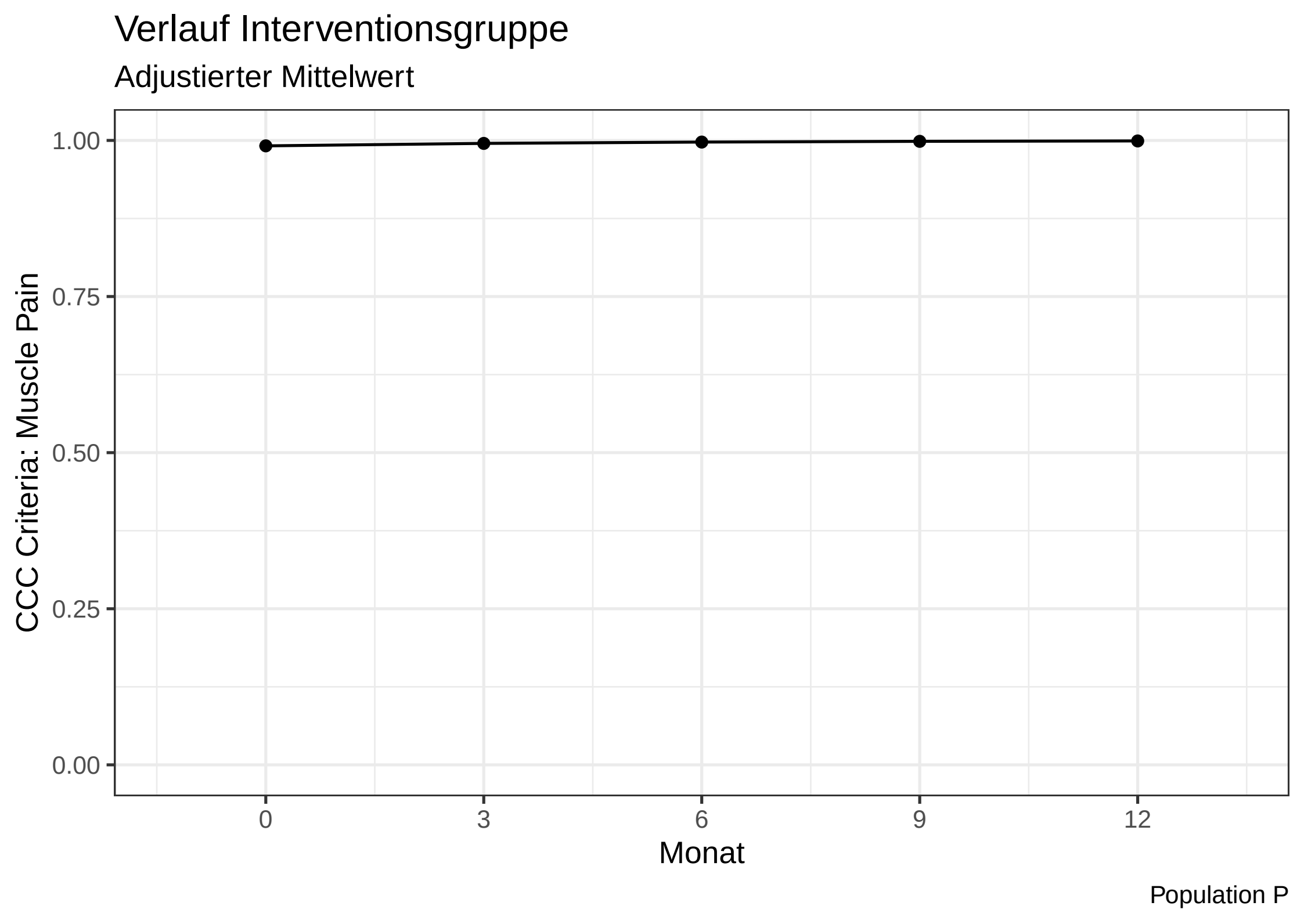
Population P

Verlauf Interventionsgruppe

Adjustierter Mittelwert

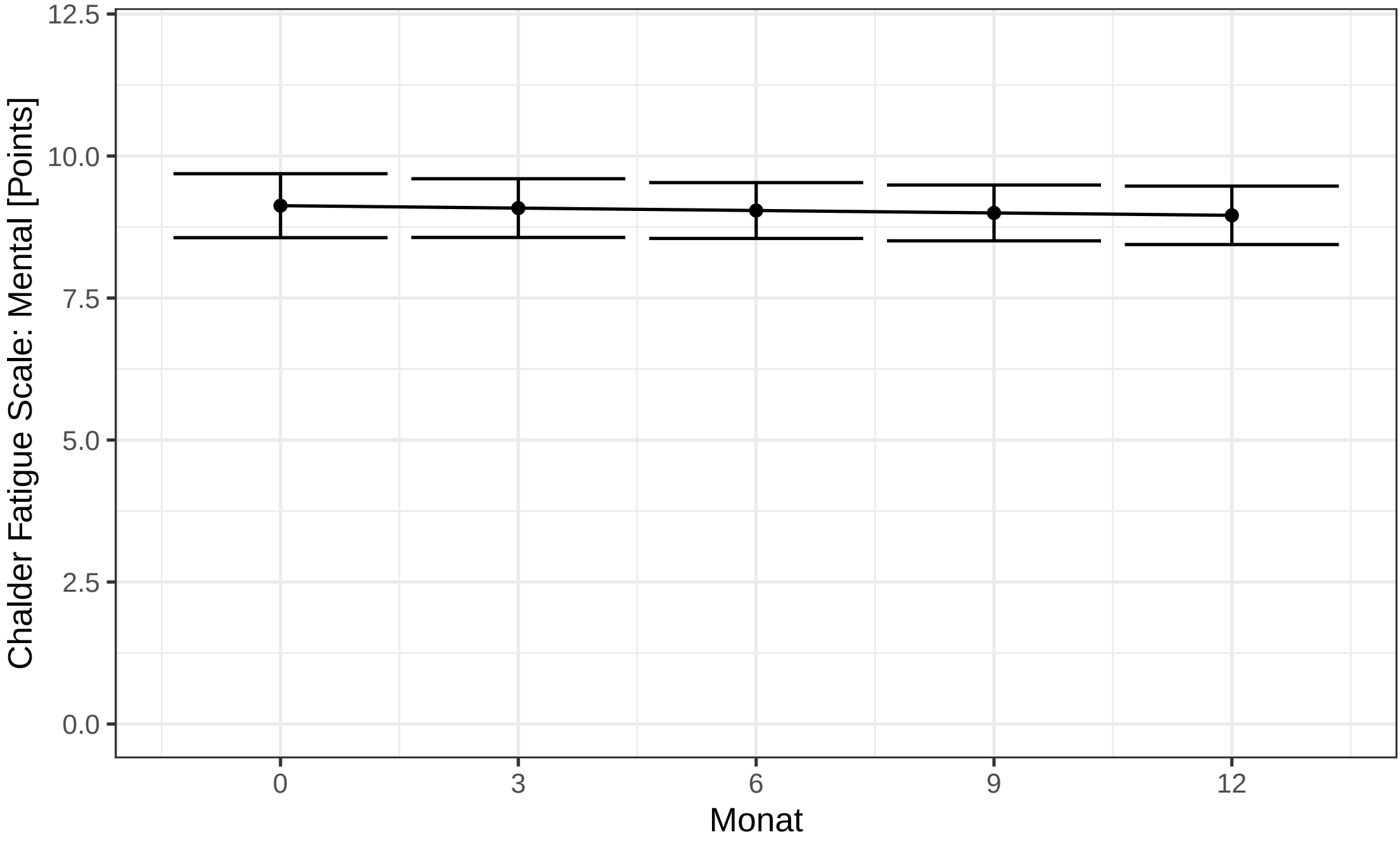


Population P


Verlauf Interventionsgruppe
Adjustierter Mittelwert
CCC Criteria: Muscle Pain
1.00
0.75
0.50
0.25
0.00
0
3
6
9
12
Monat
Population P

Verlauf Interventionsgruppe

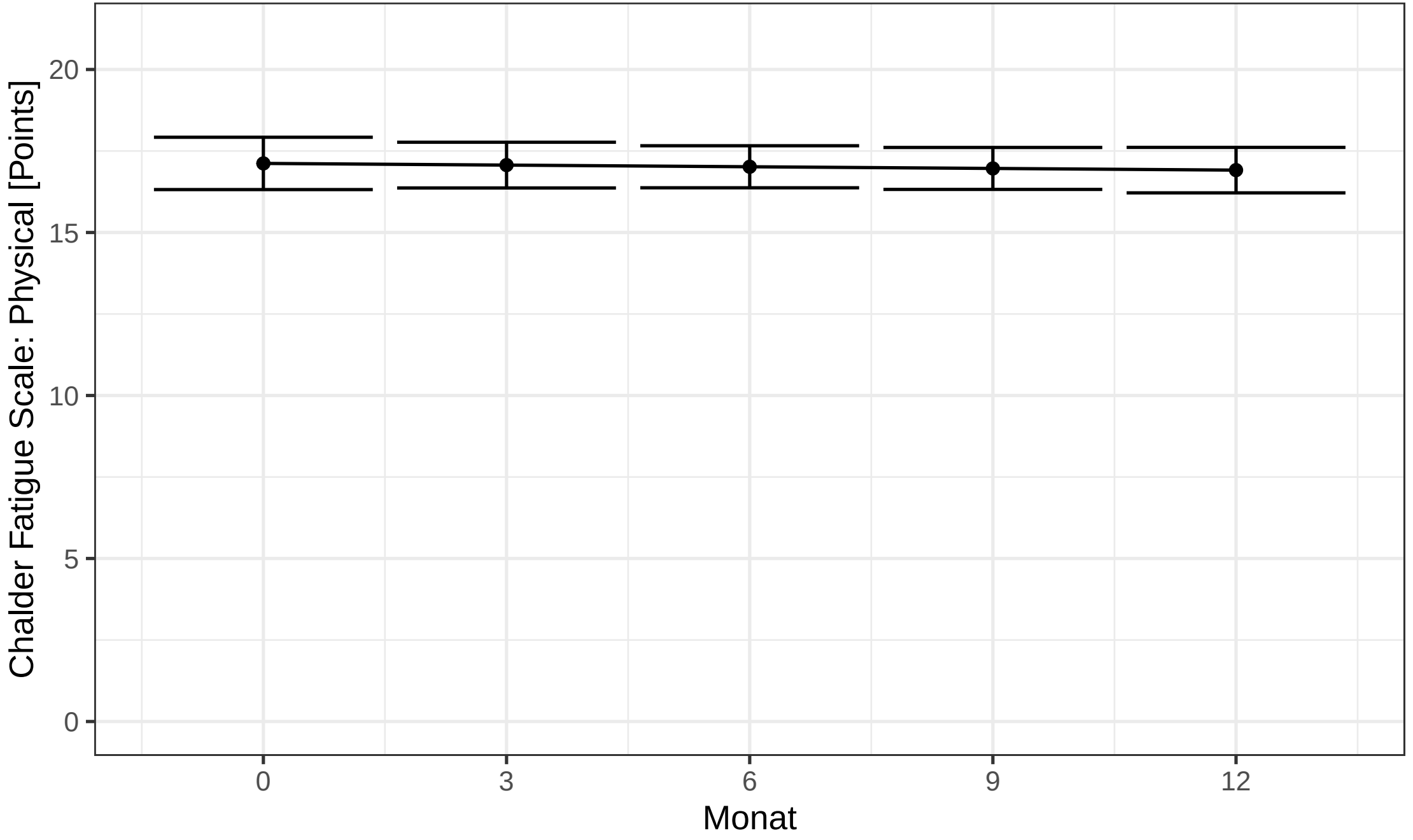
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

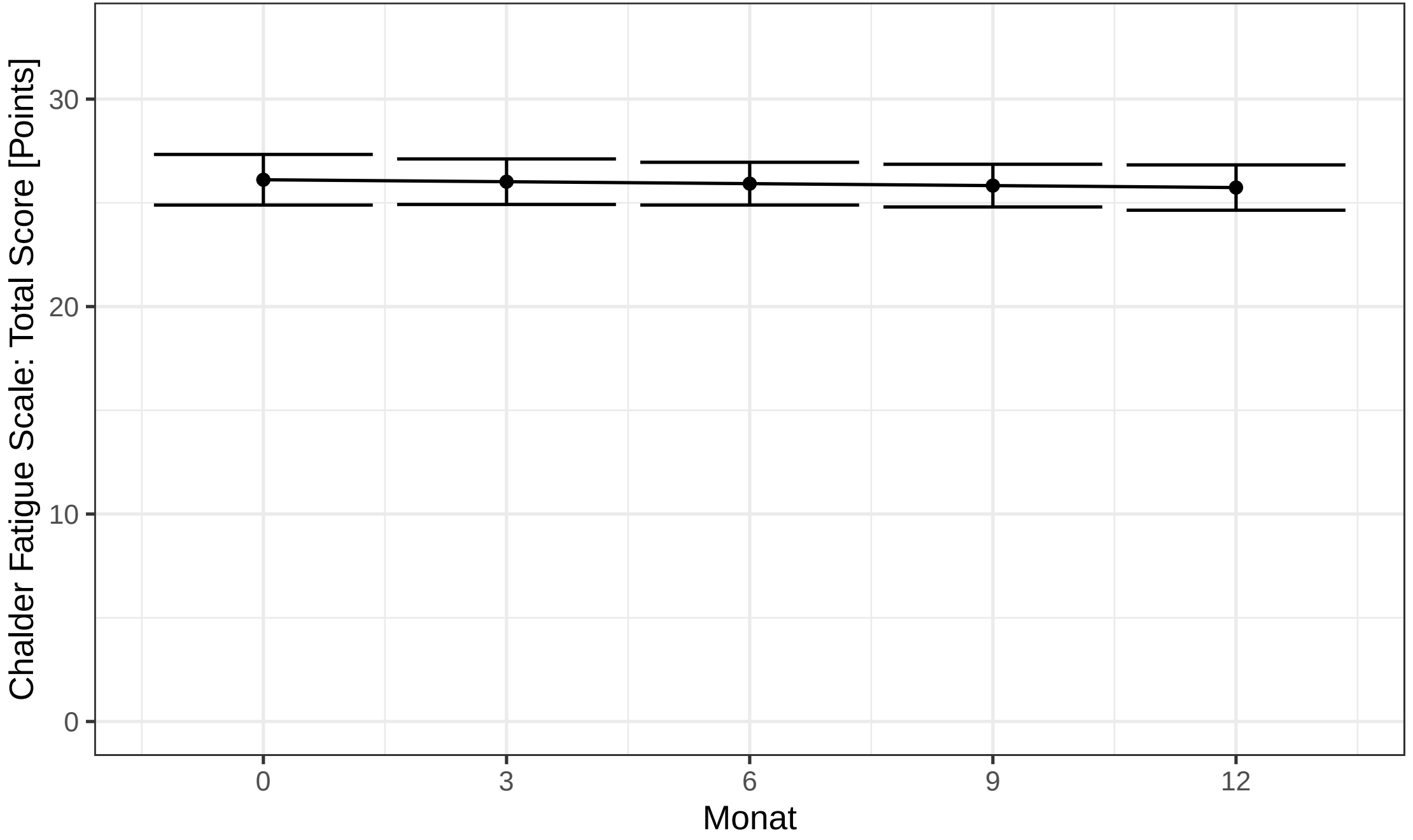
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

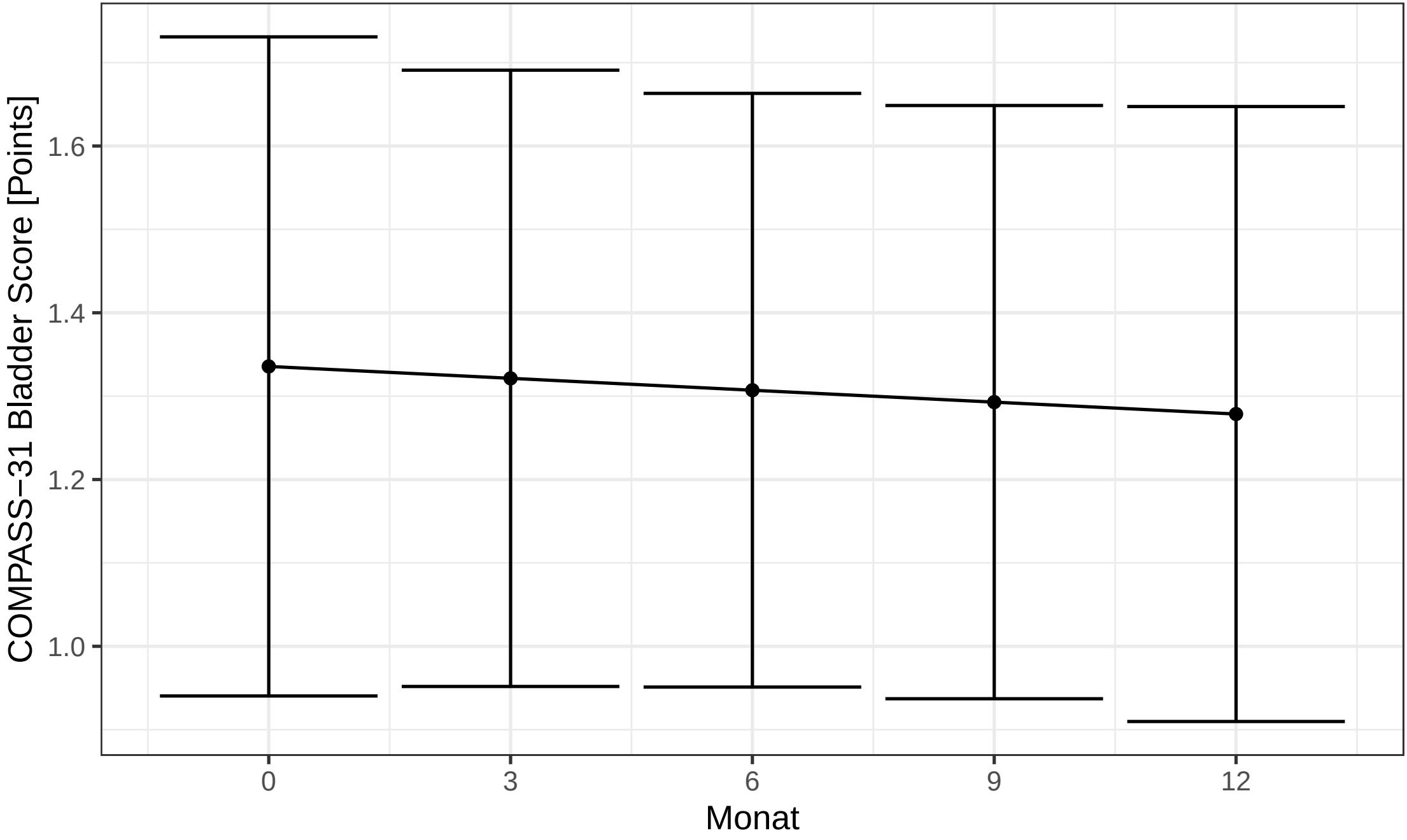
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

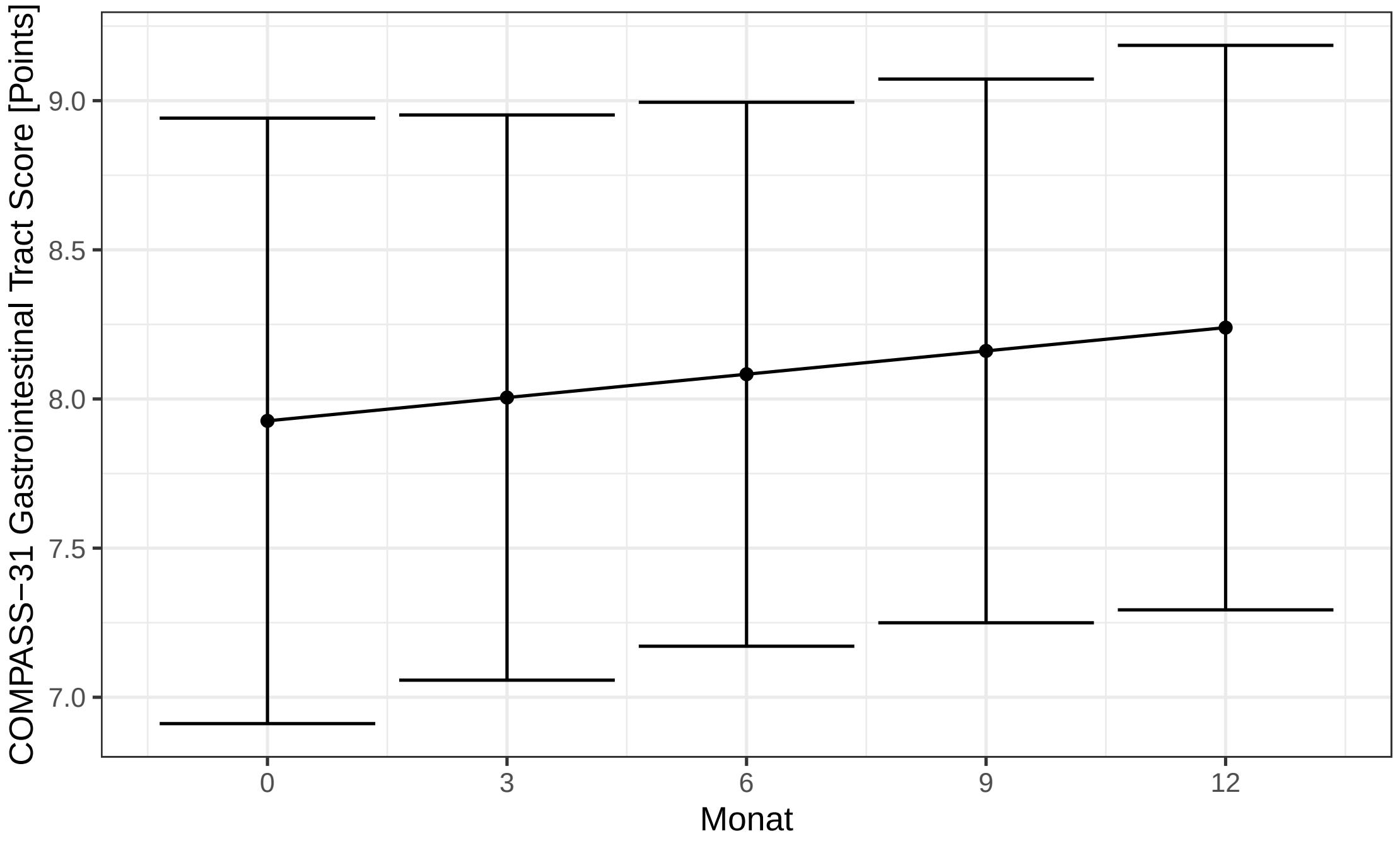
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

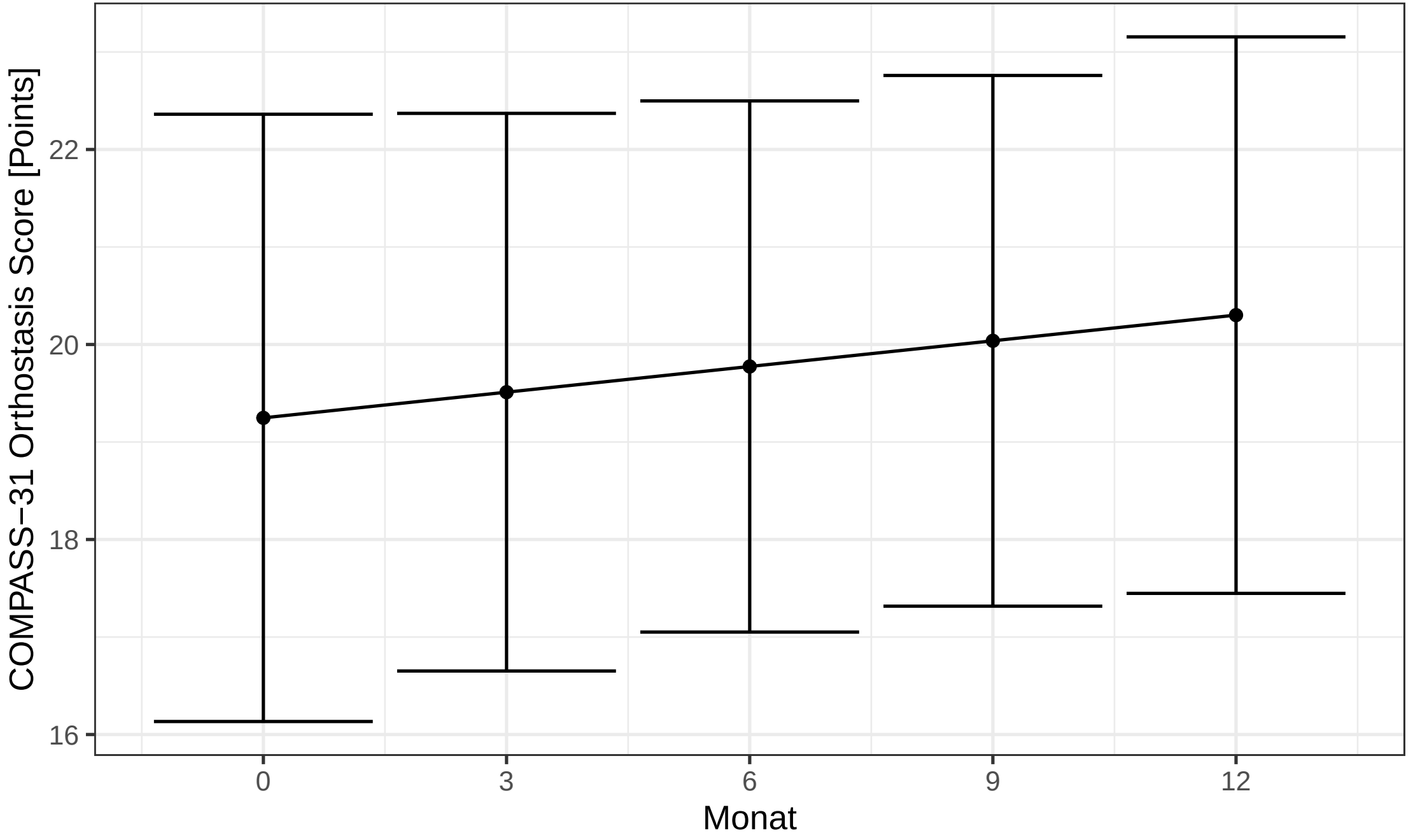
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

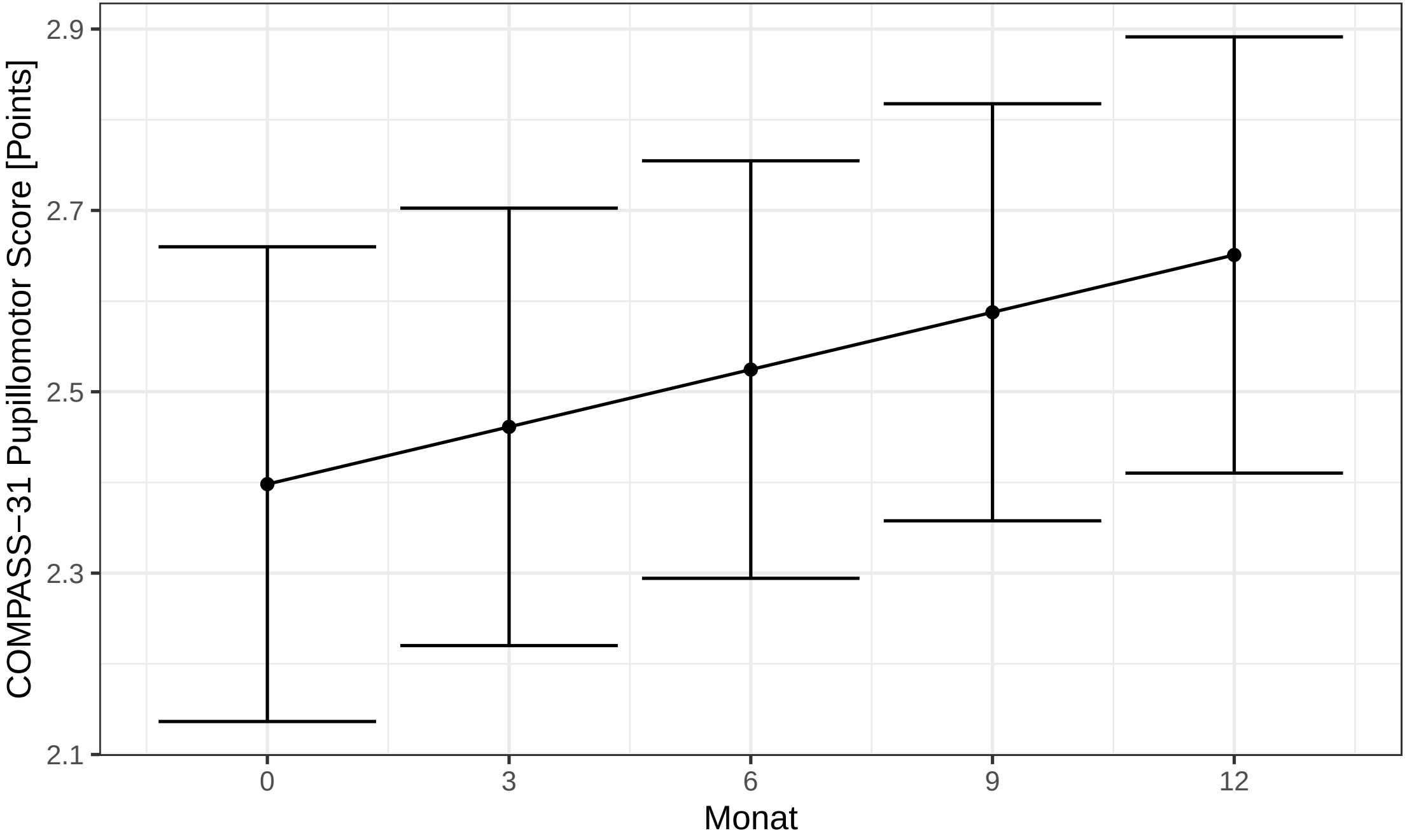
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

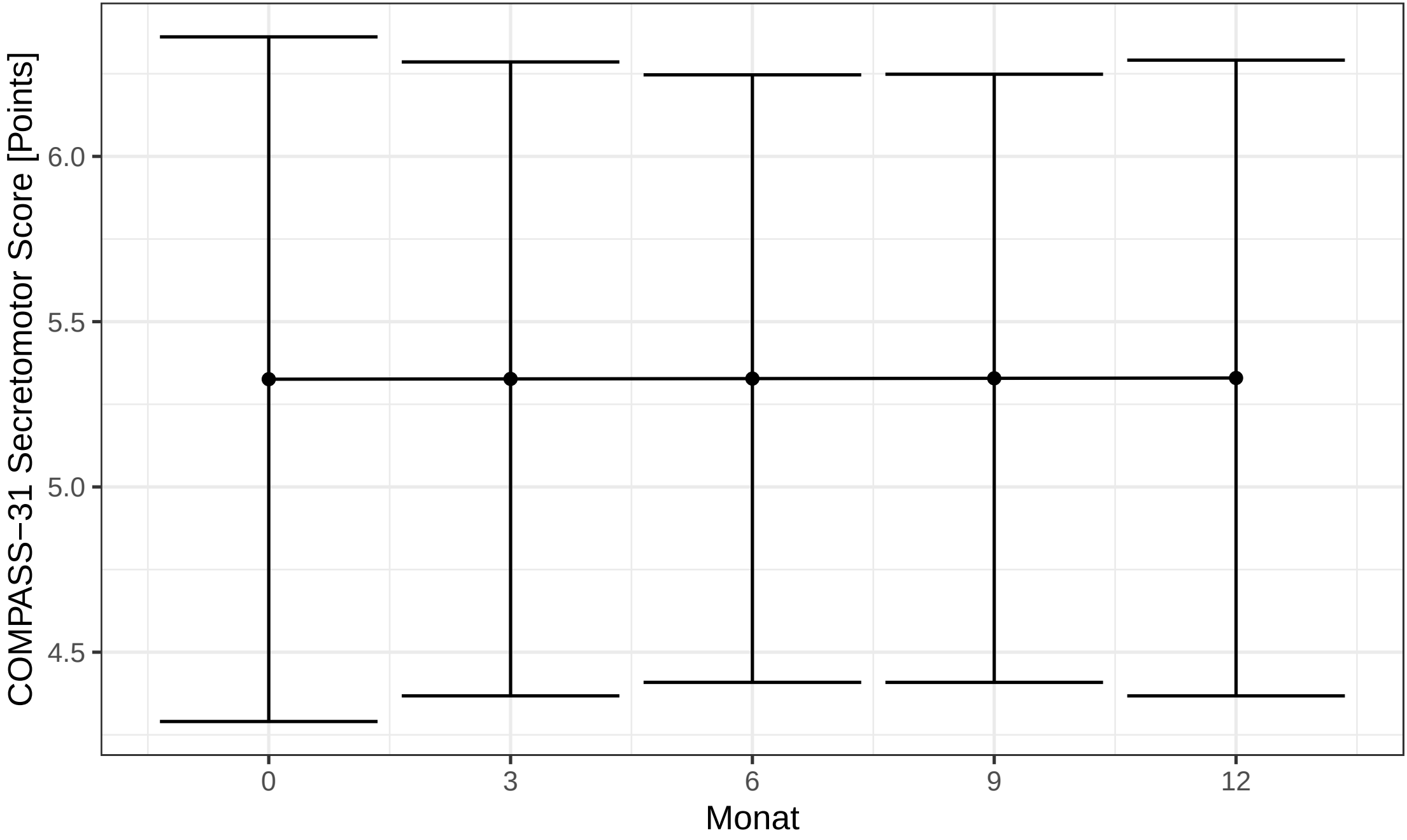
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

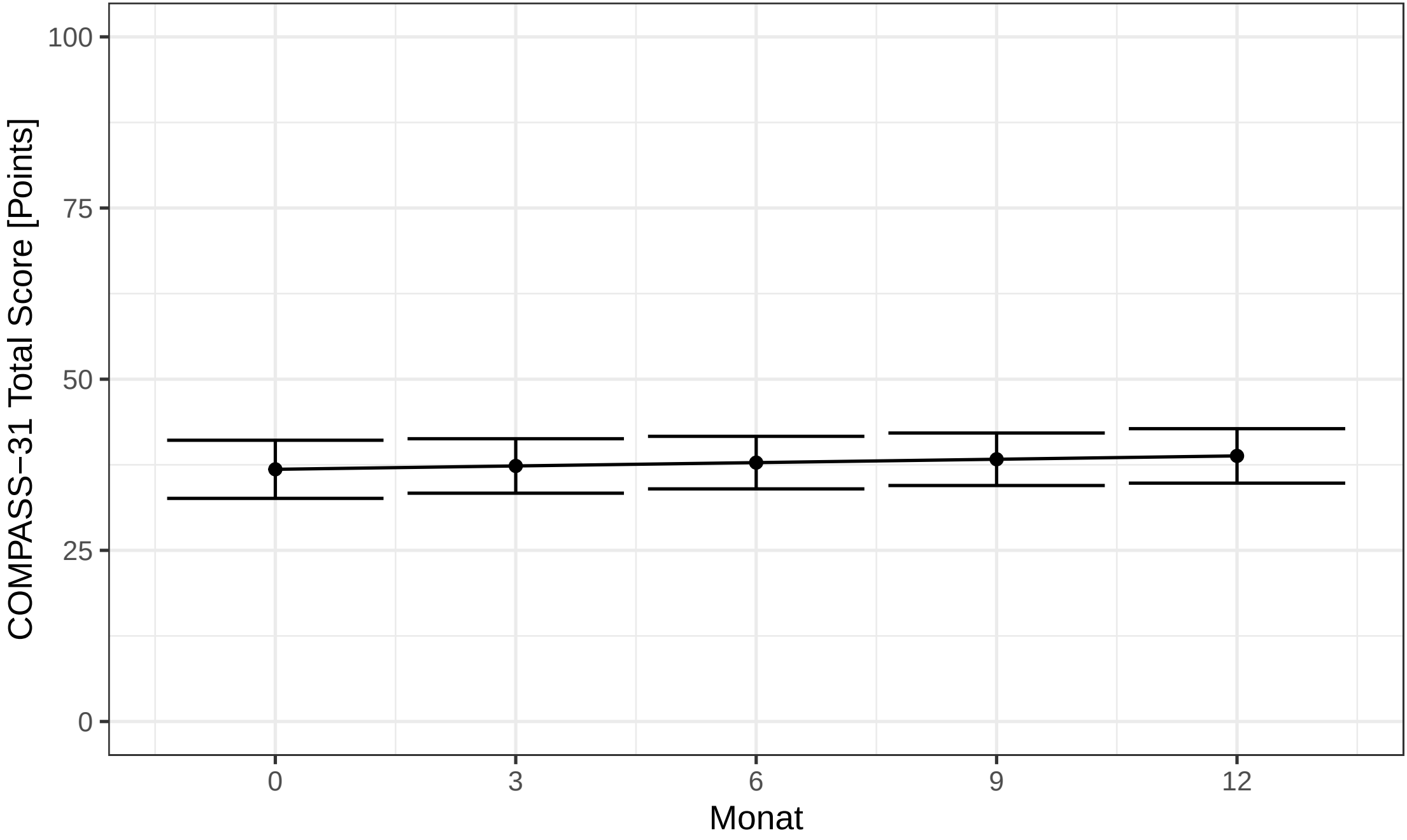
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

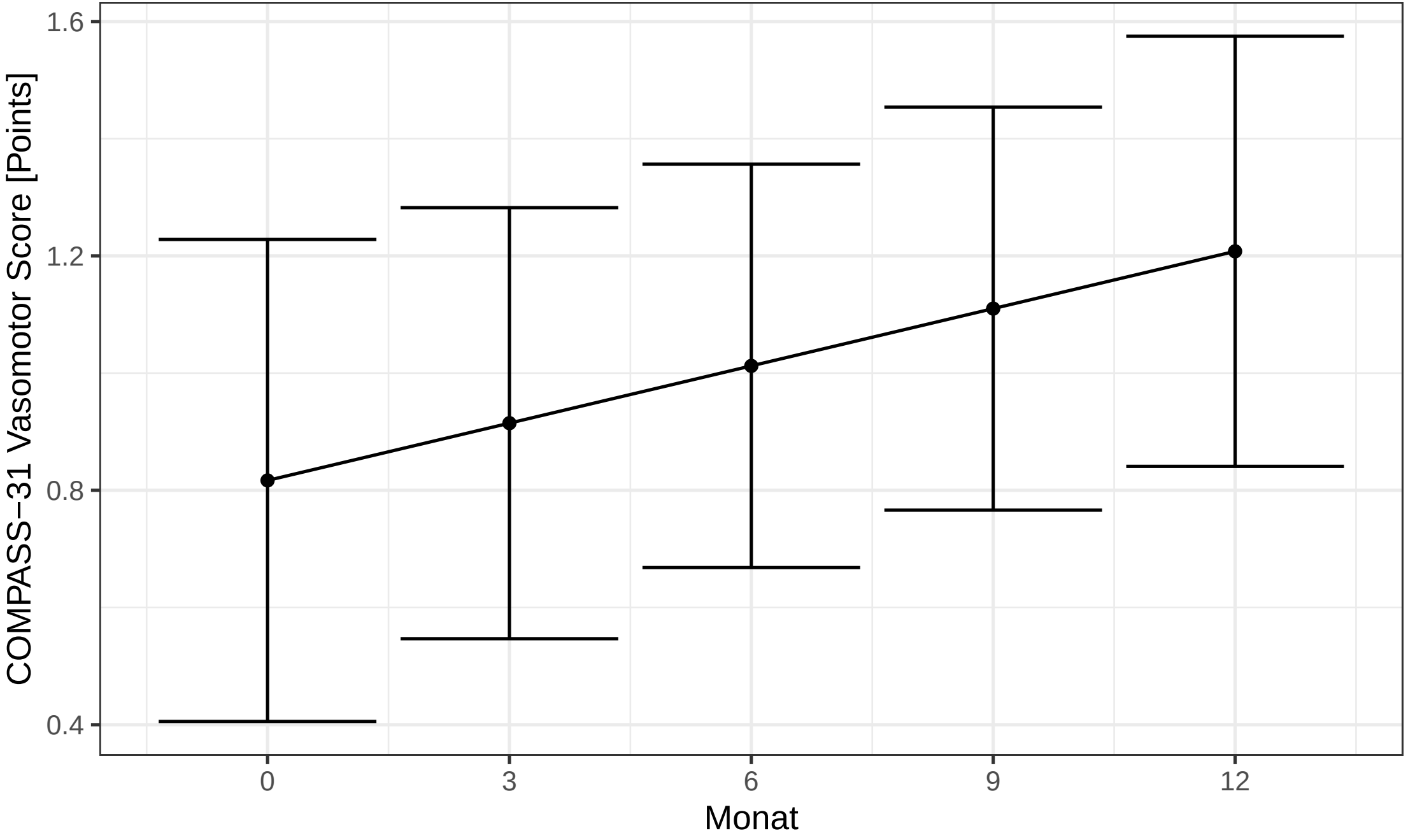
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

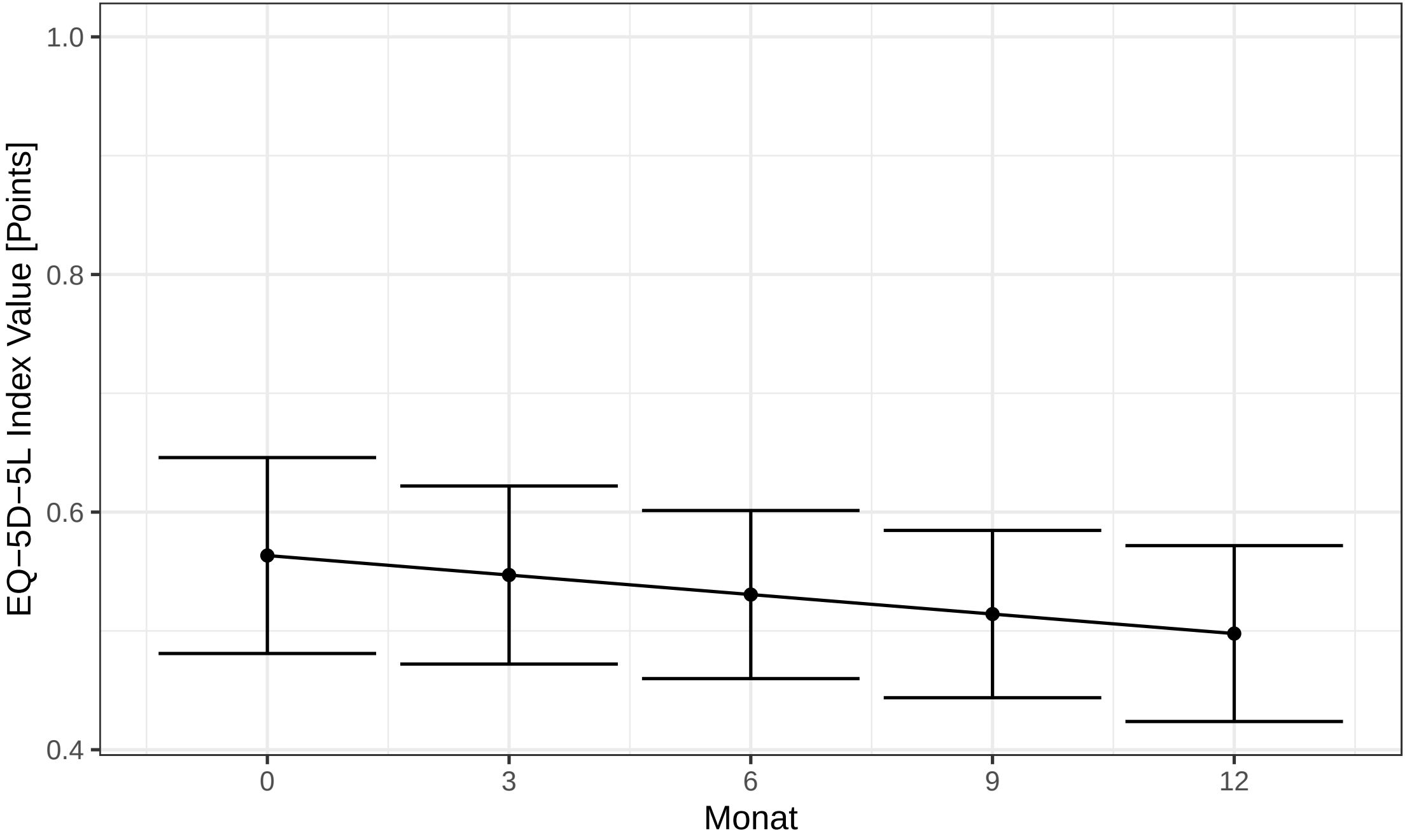
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

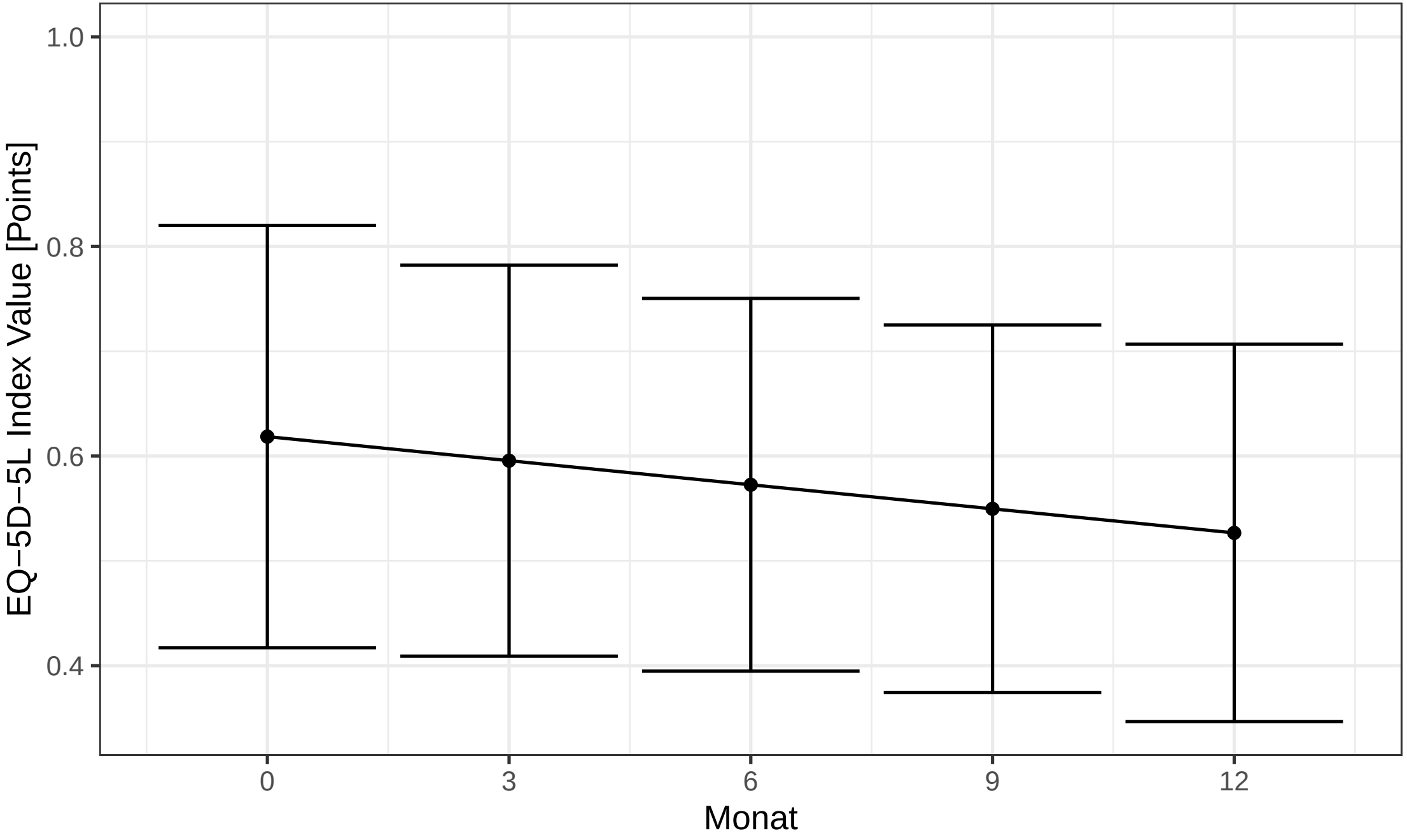
Adjustierter Mittelwert



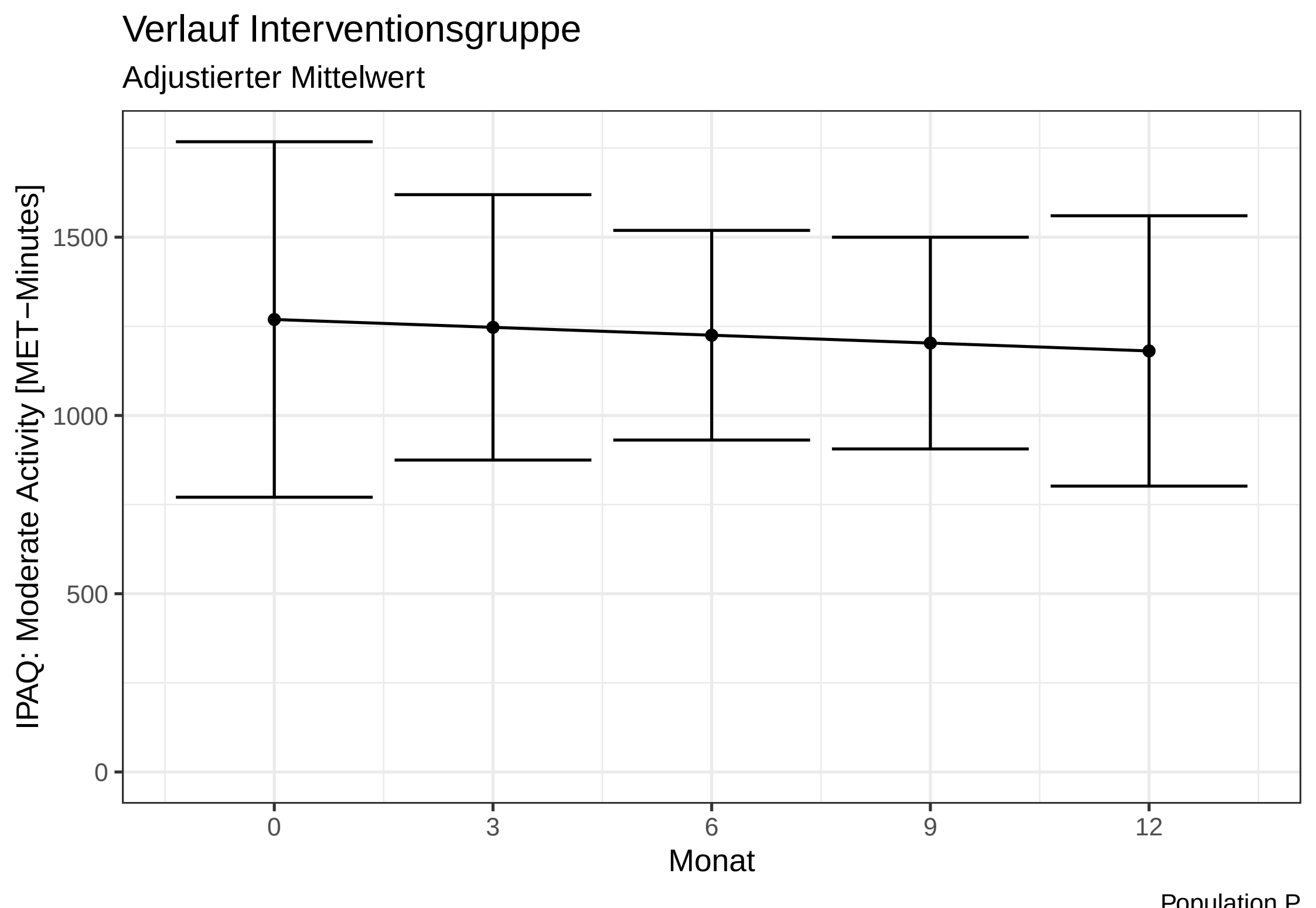
Population P

Verlauf Interventionsgruppe

Adjustierter Mittelwert

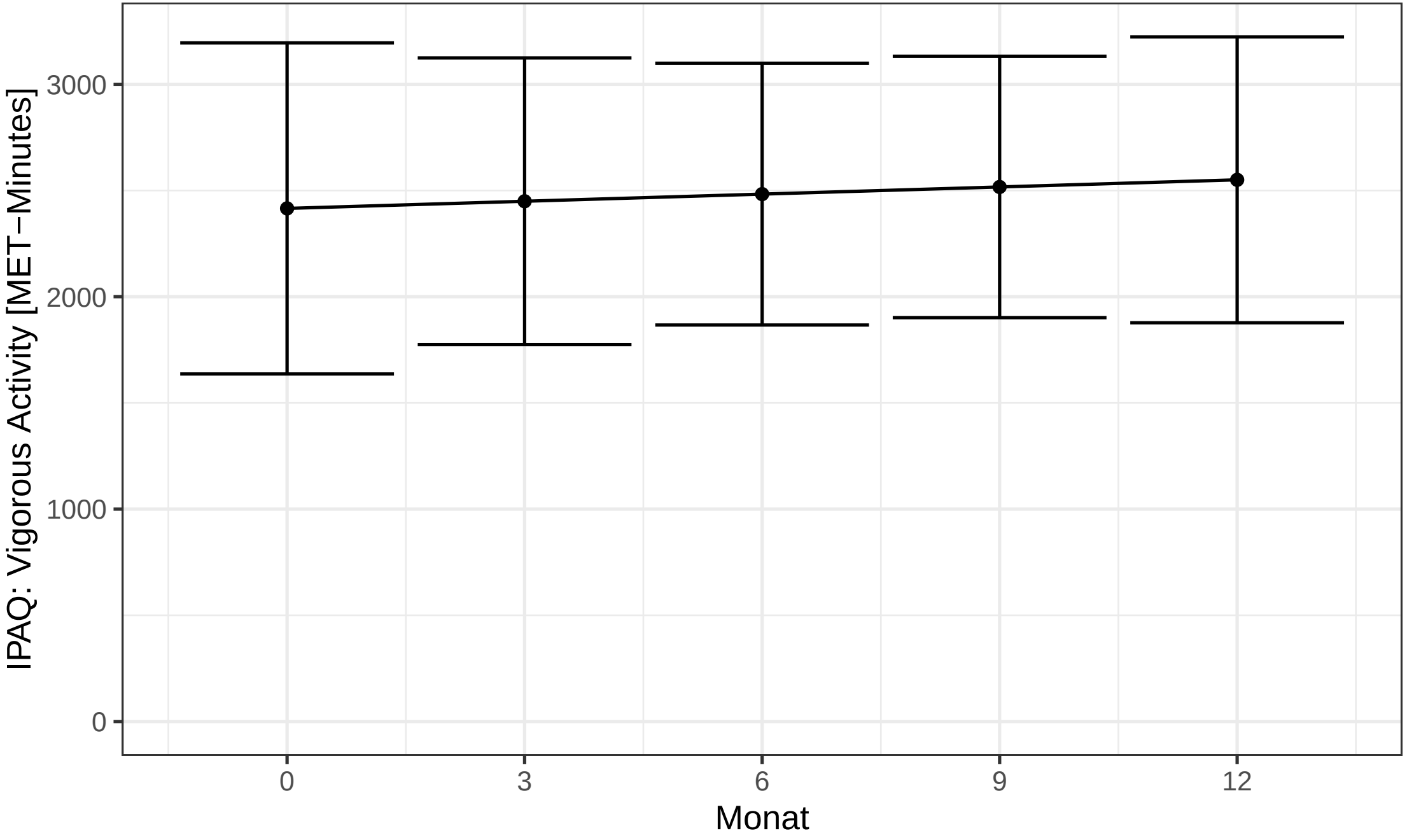


Population P


Verlauf Interventionsgruppe
Adjustierter Mittelwert
IPAQ: Moderate Activity [MET−Minutes]
1500
1000
500
0
0
3
6
9
12
Monat
Population P

Verlauf Interventionsgruppe

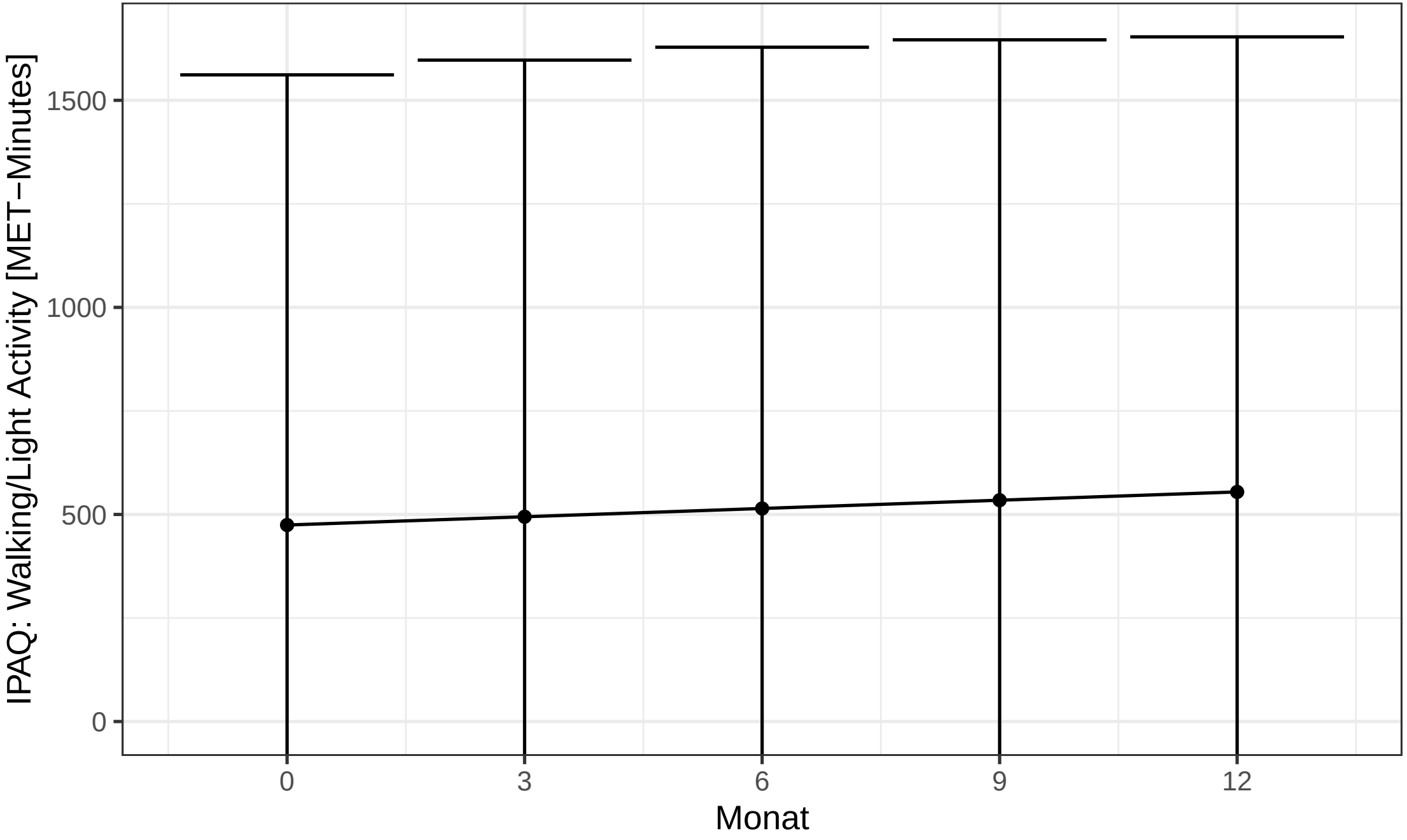
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

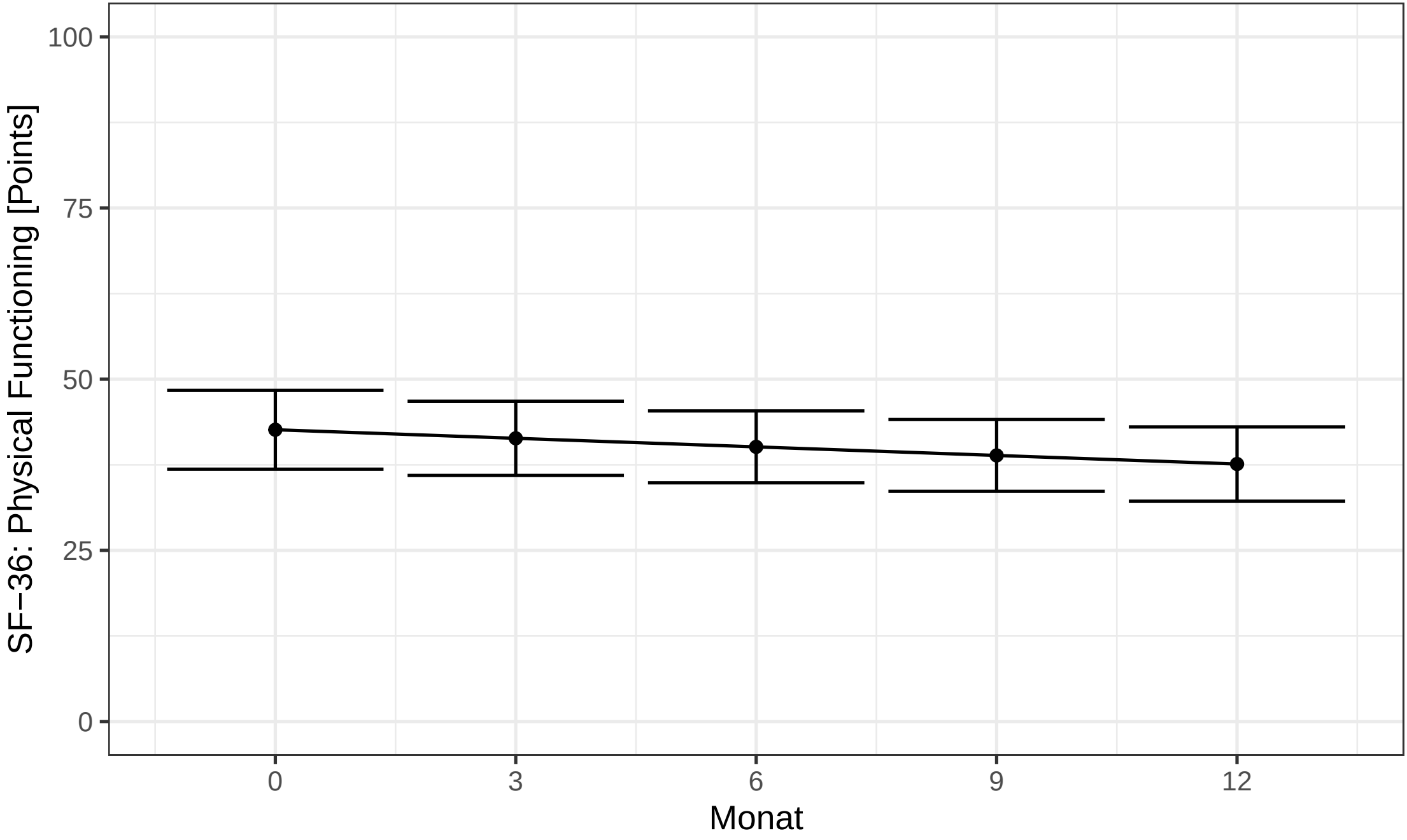
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

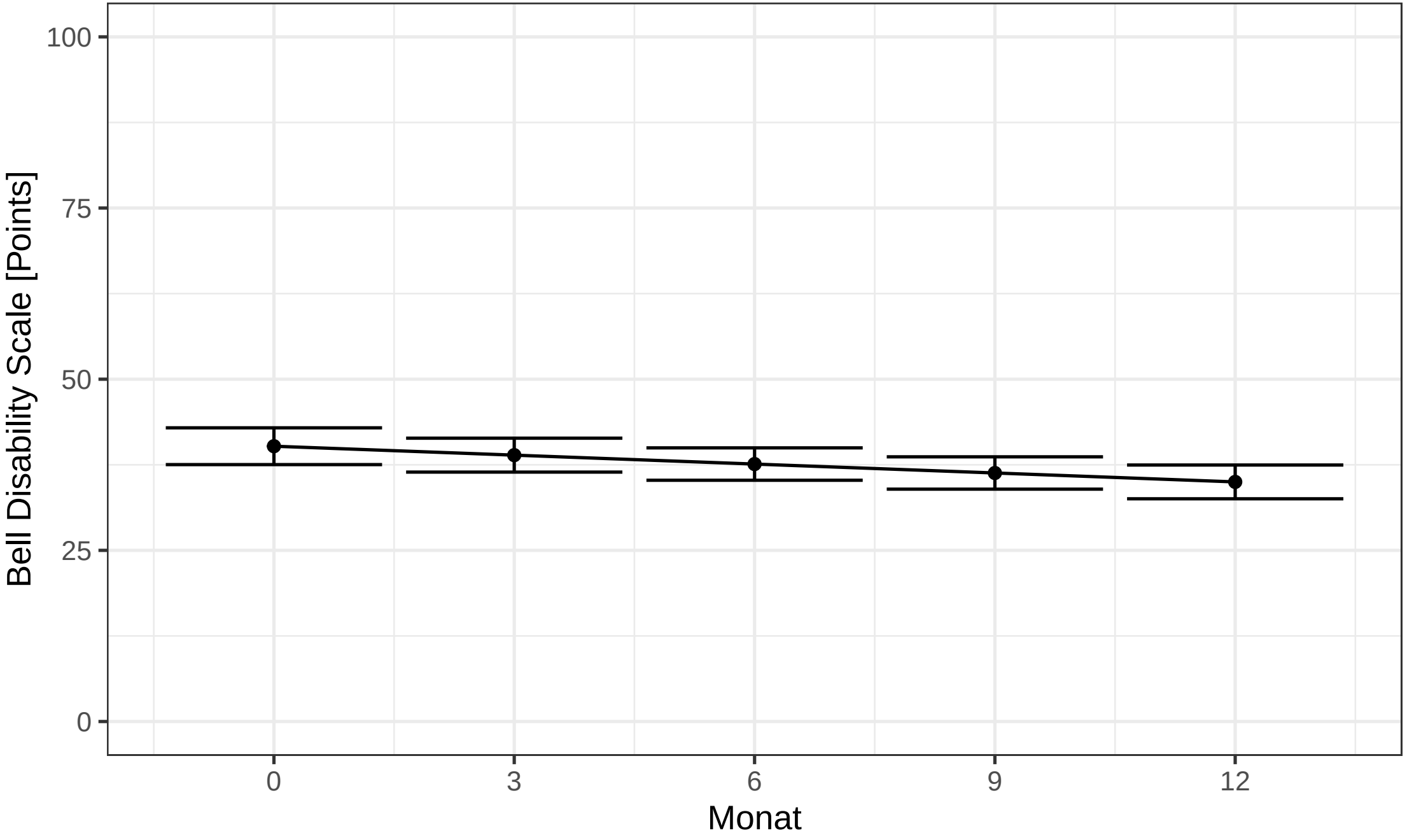
Adjustierter Mittelwert



Population P

Verlauf Interventionsgruppe

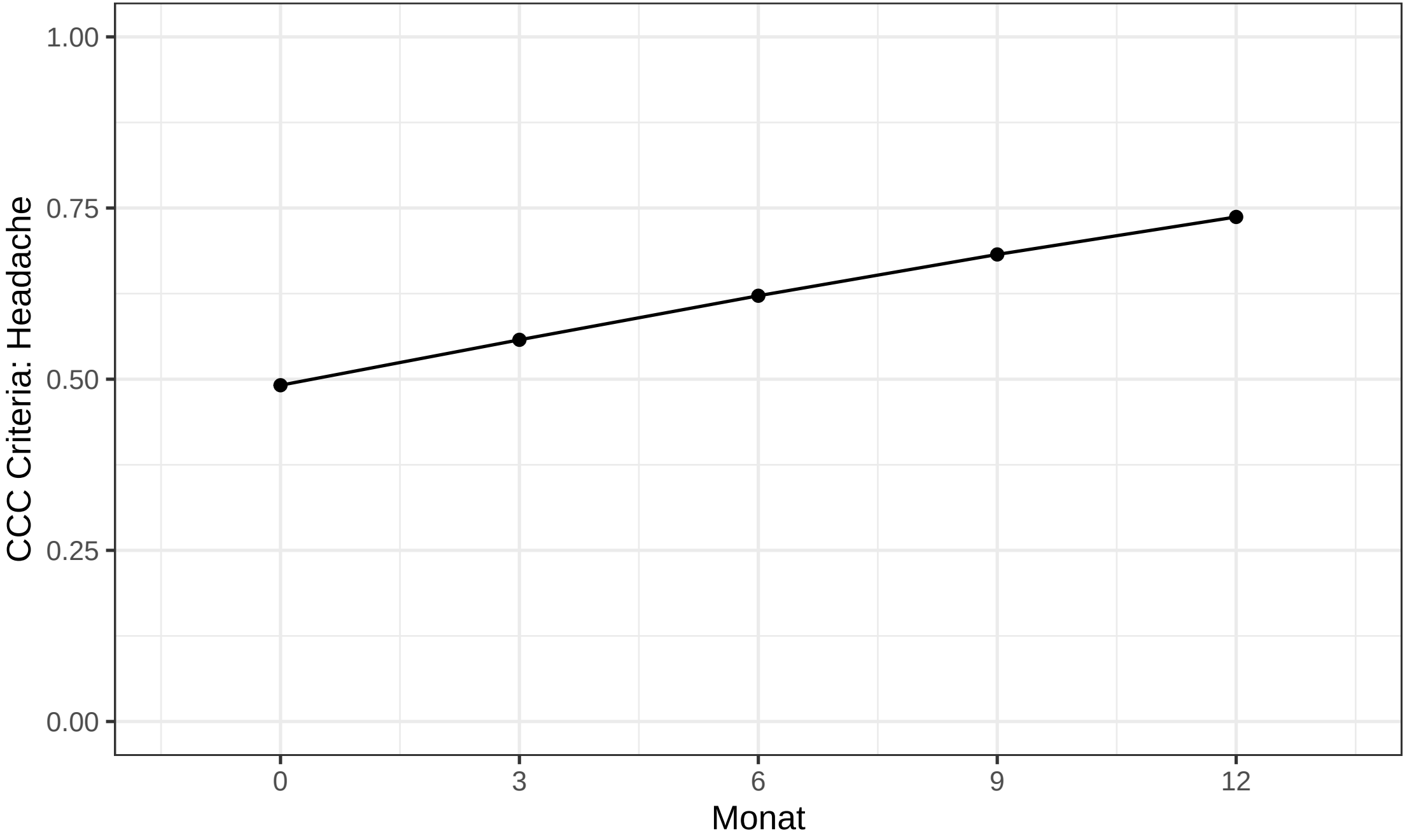
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

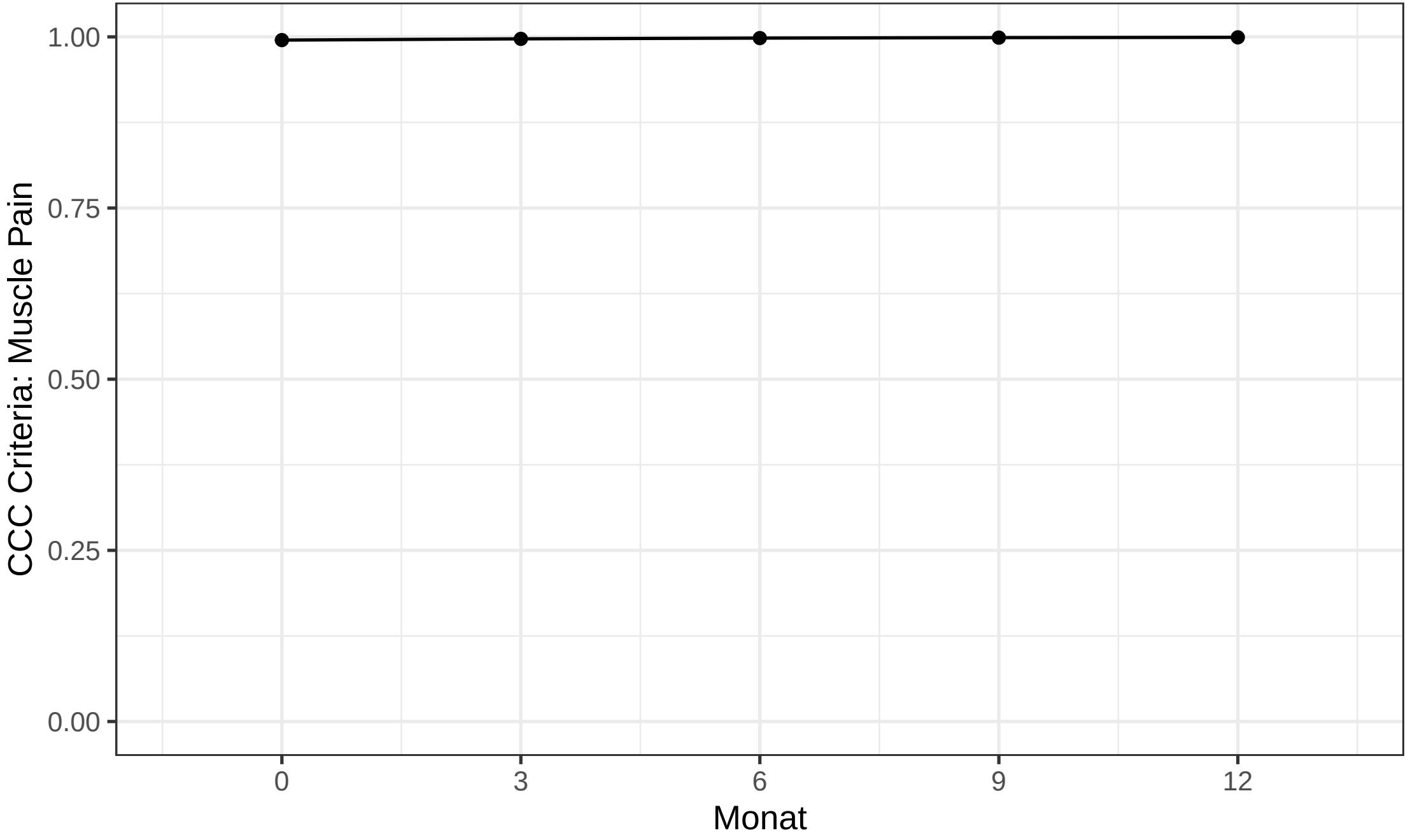
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

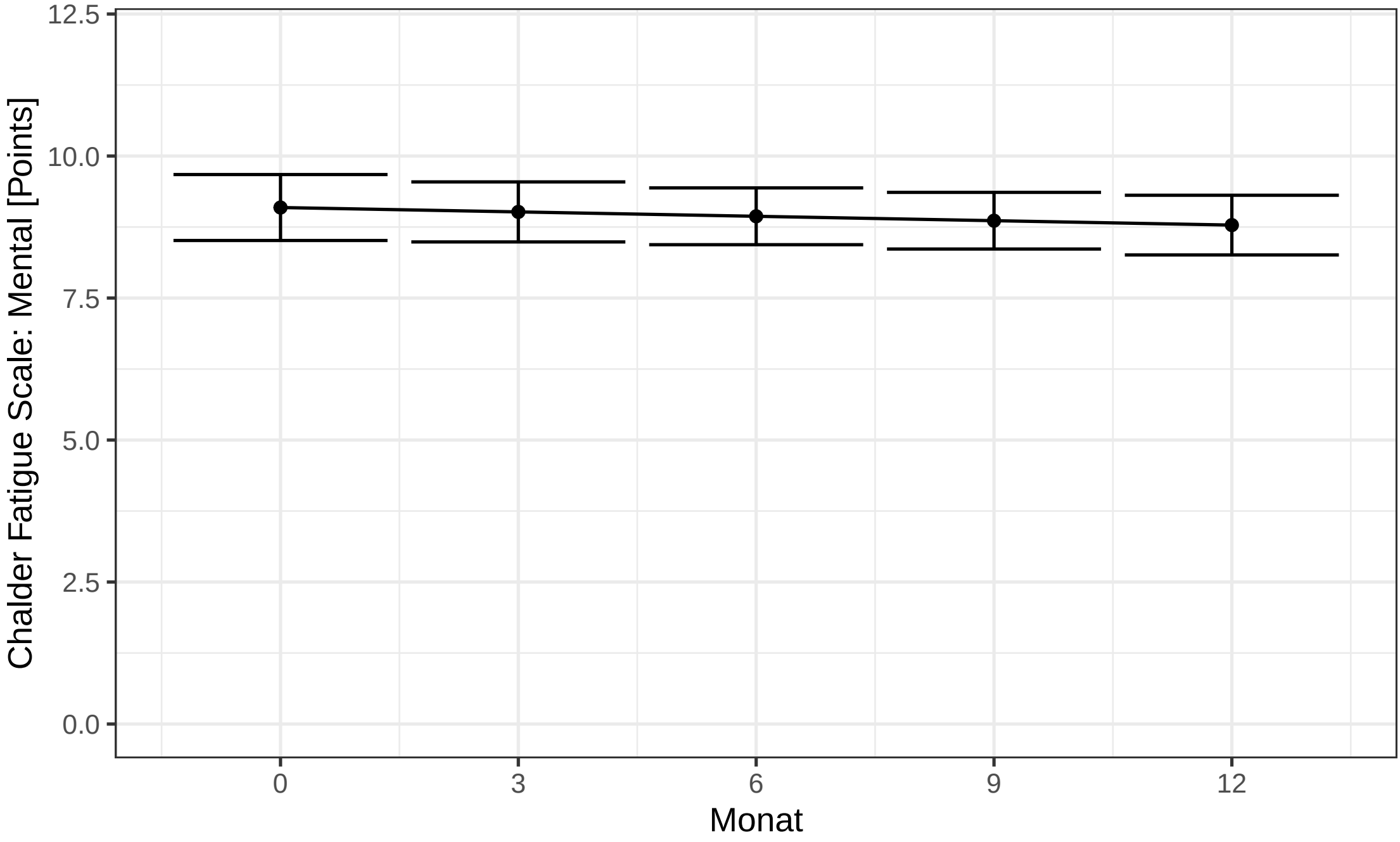
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

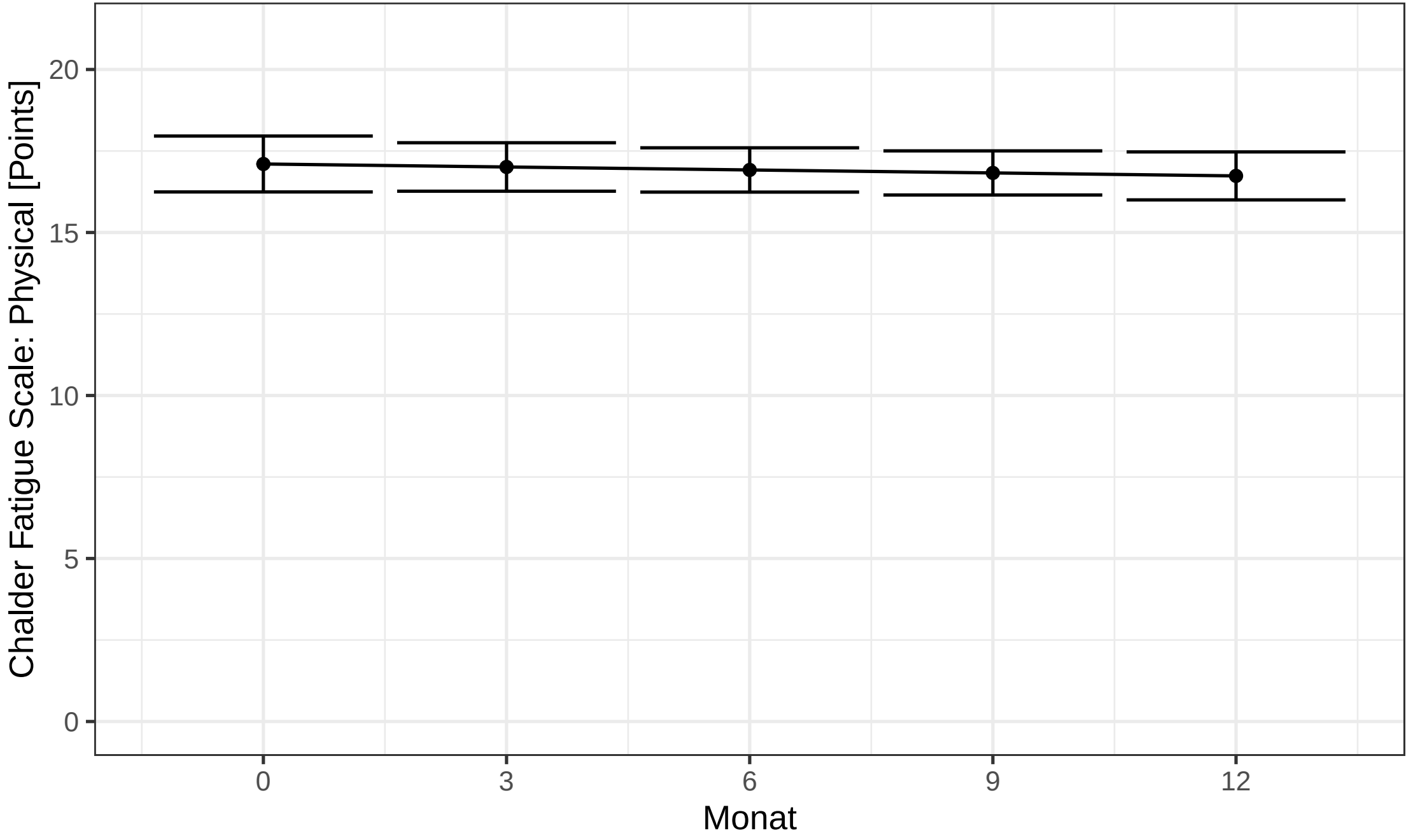
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

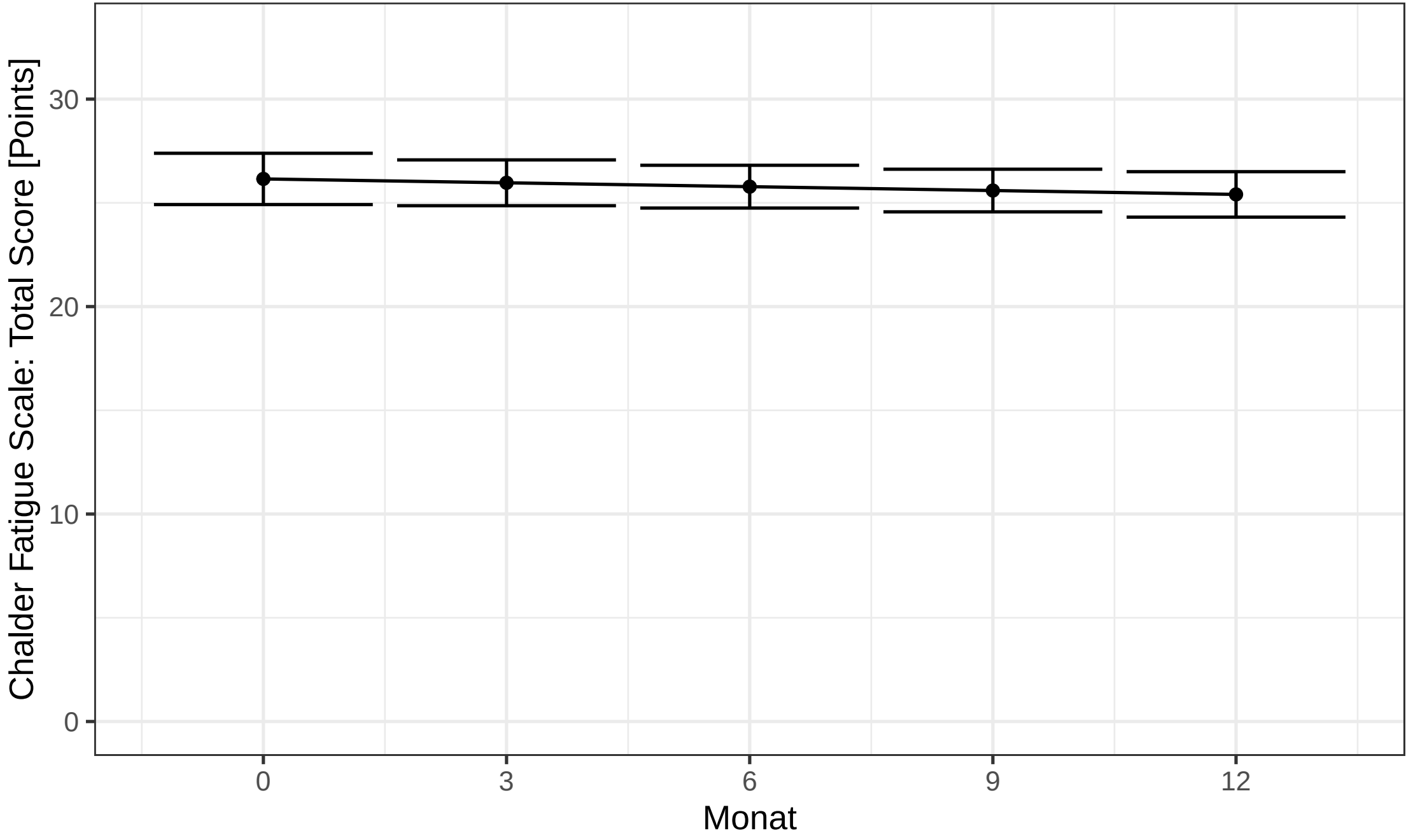
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

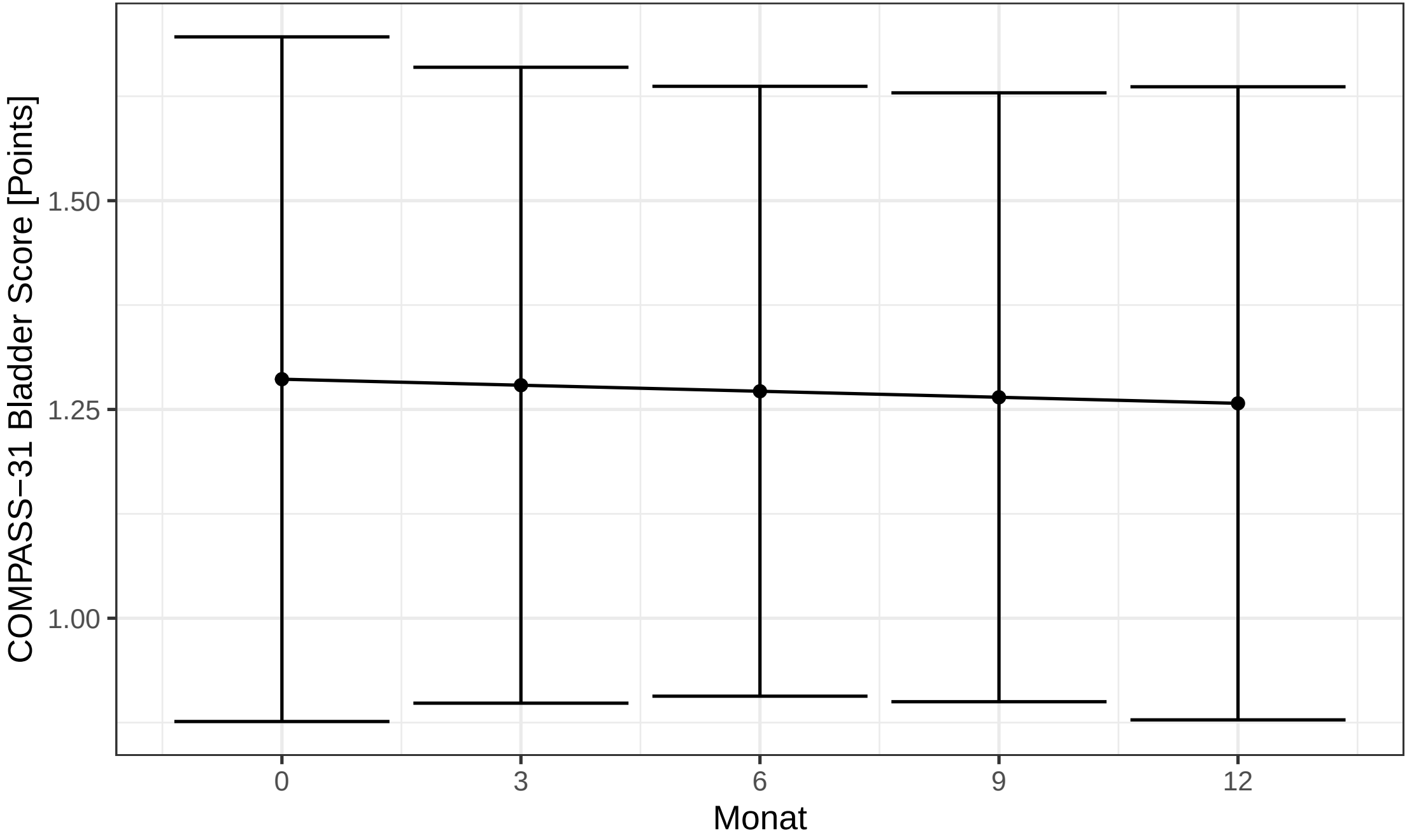
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

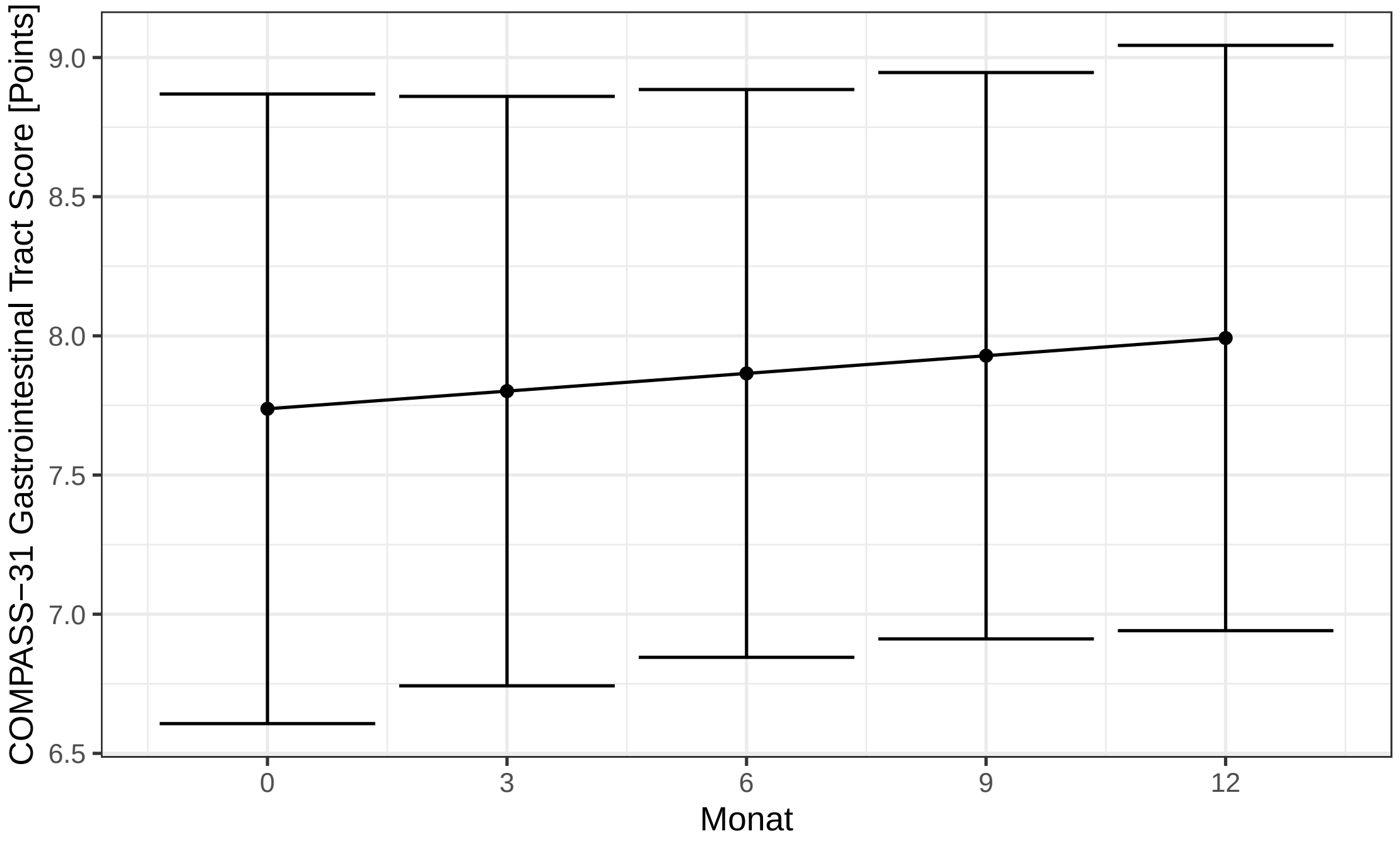
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

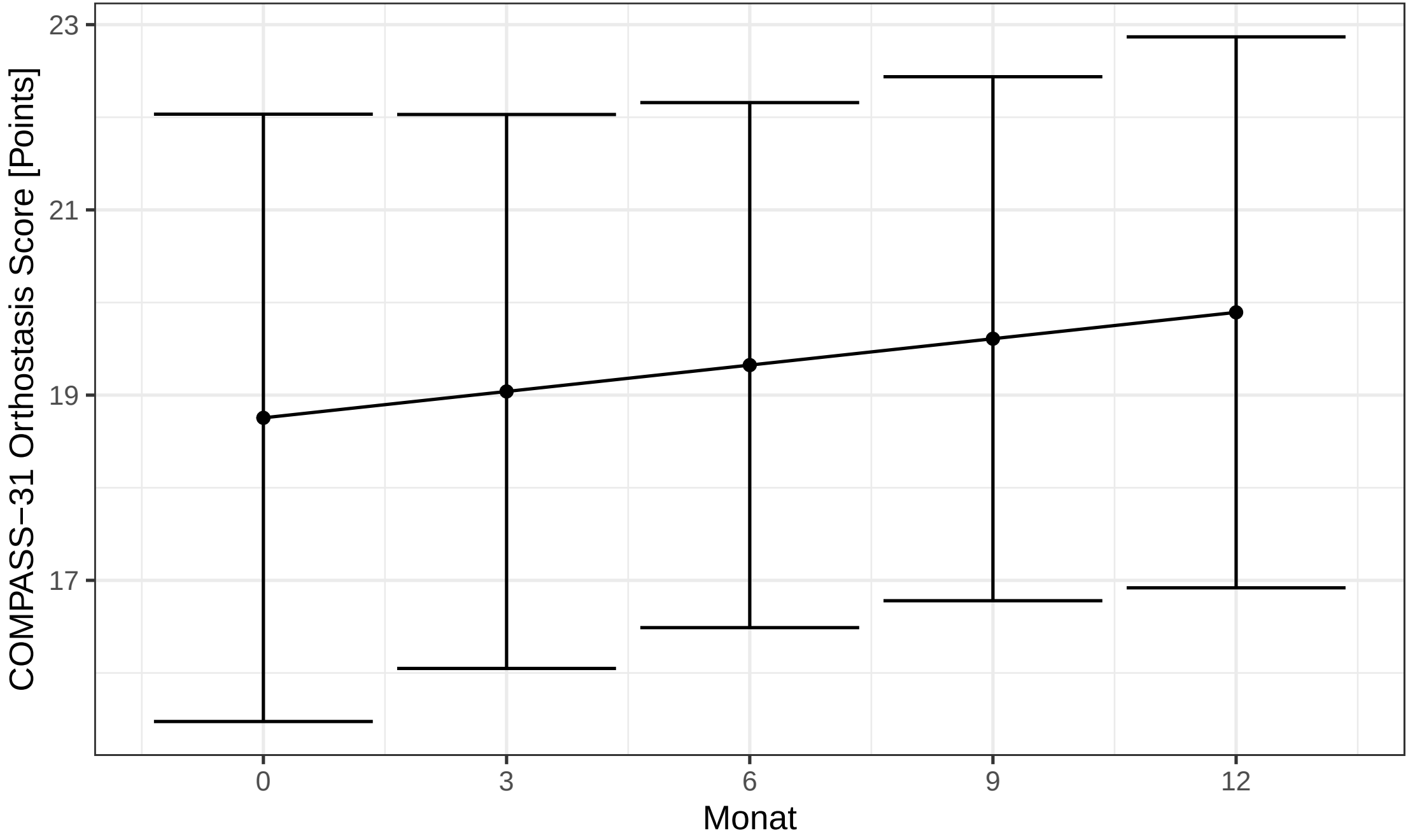
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

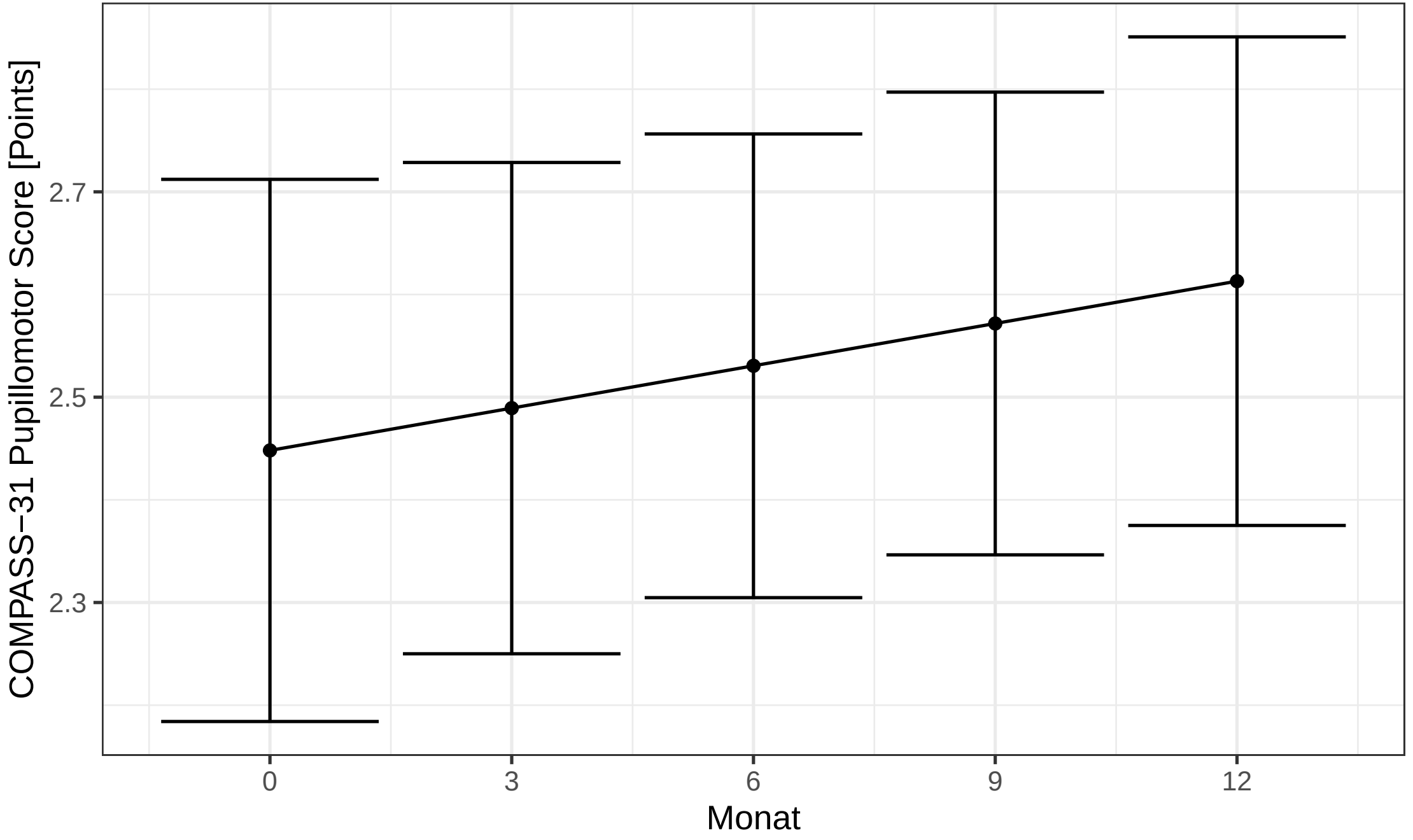
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

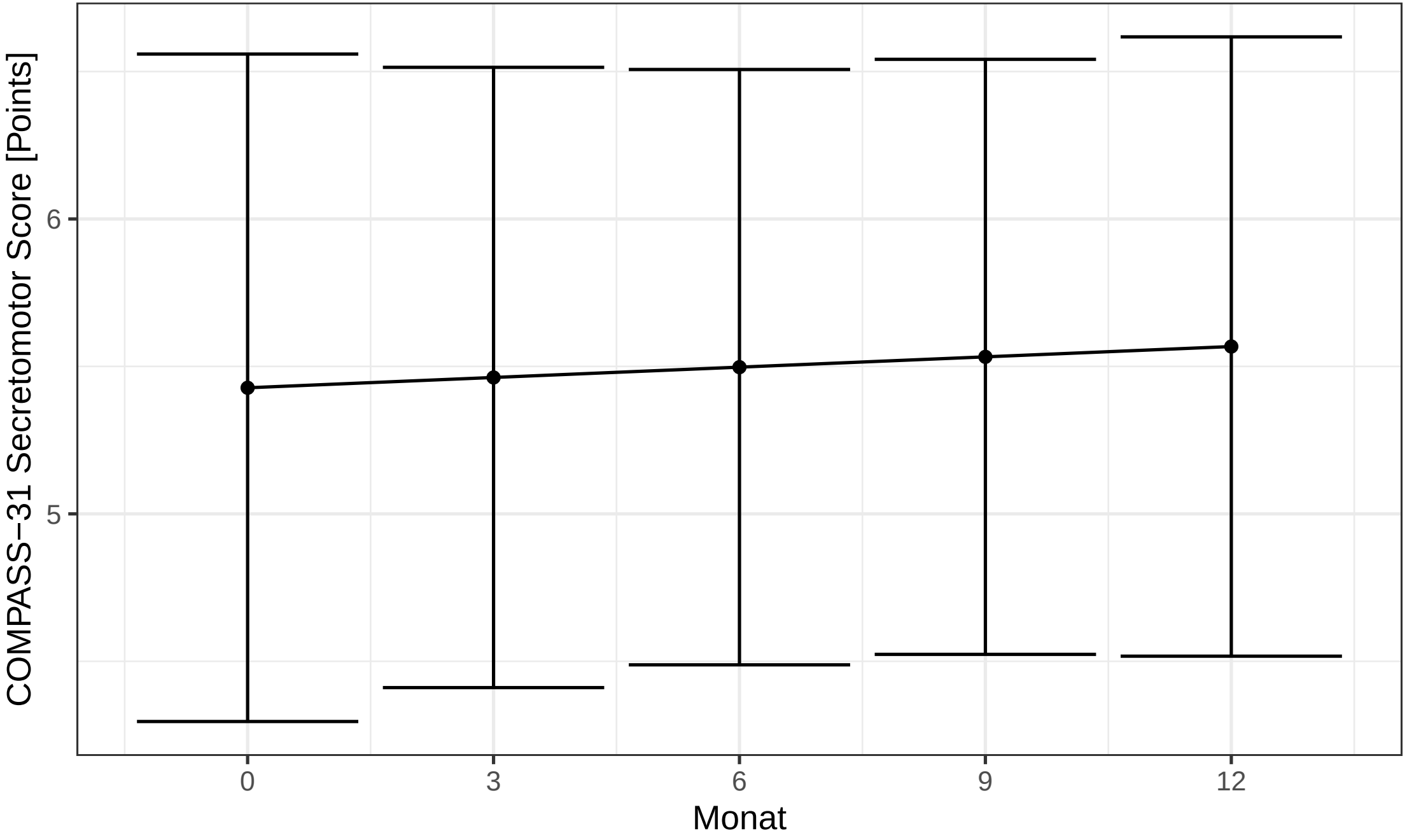
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

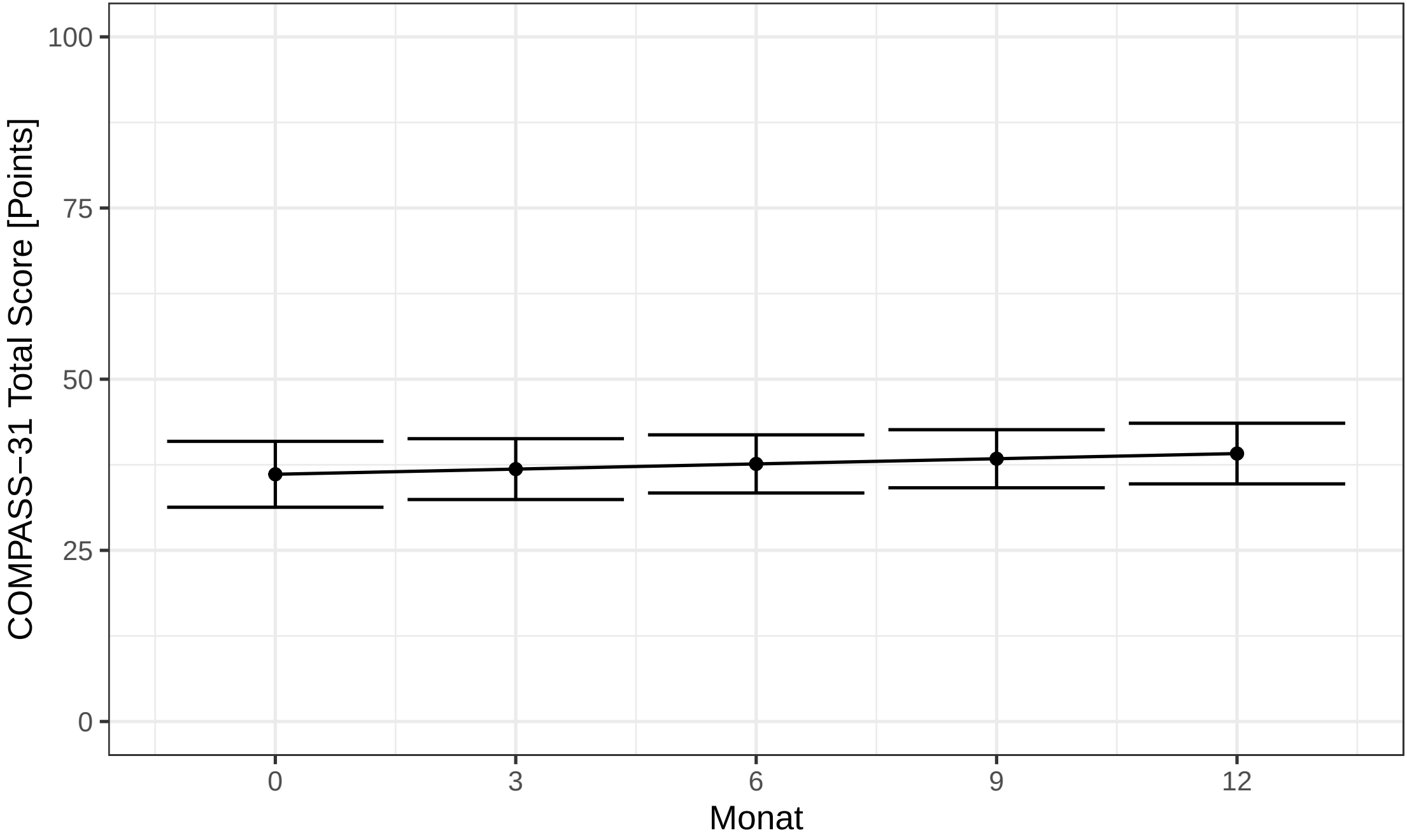
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

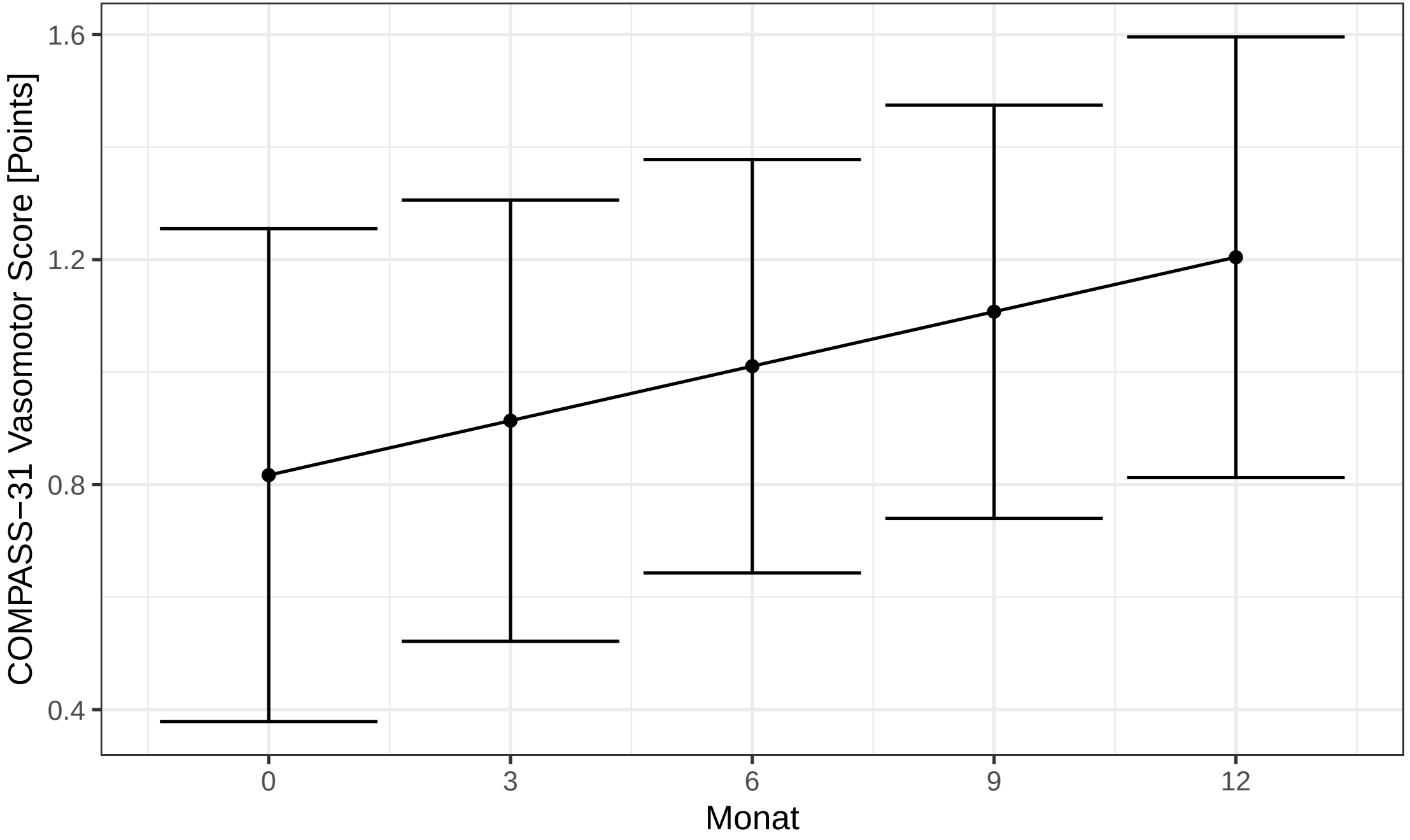
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

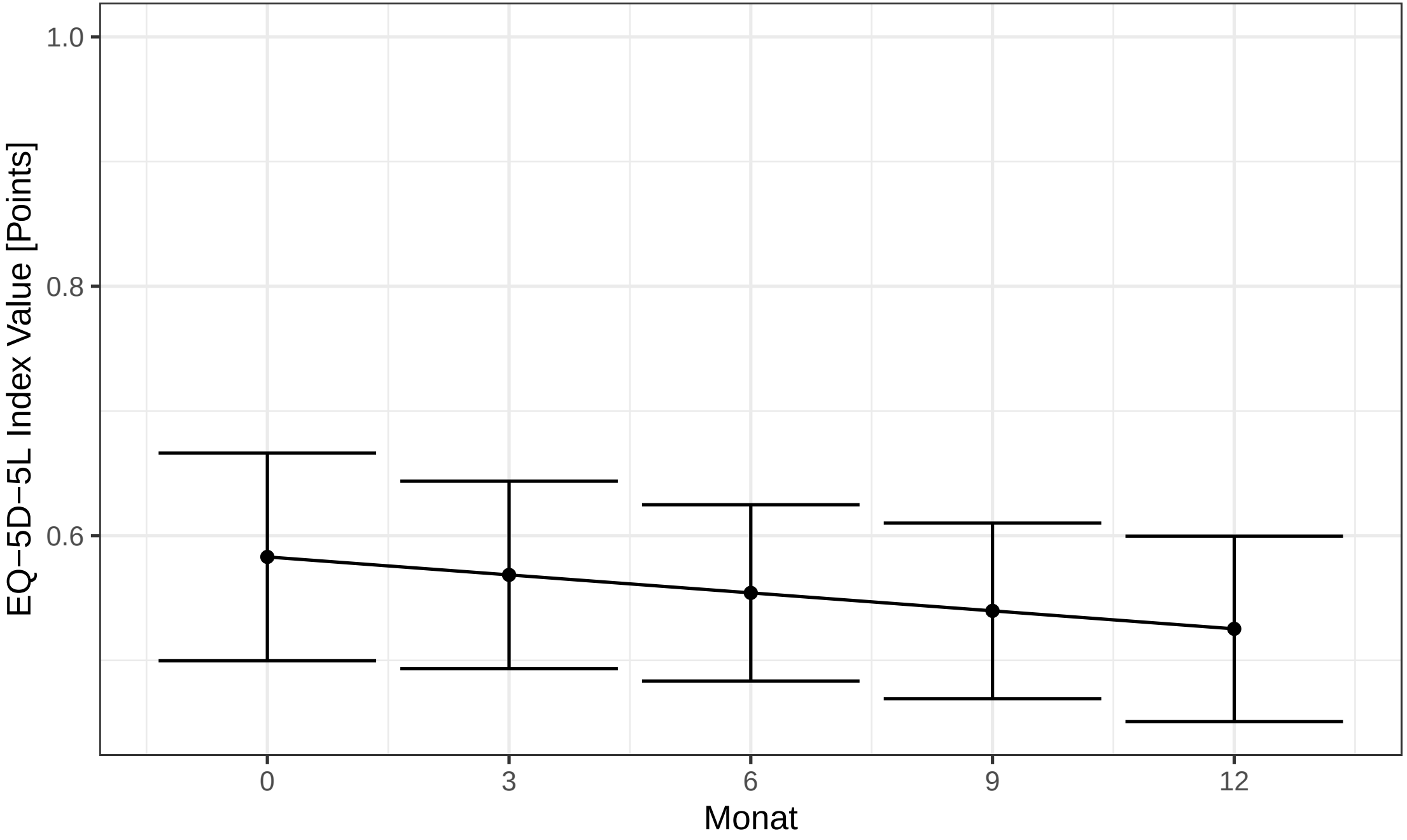
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

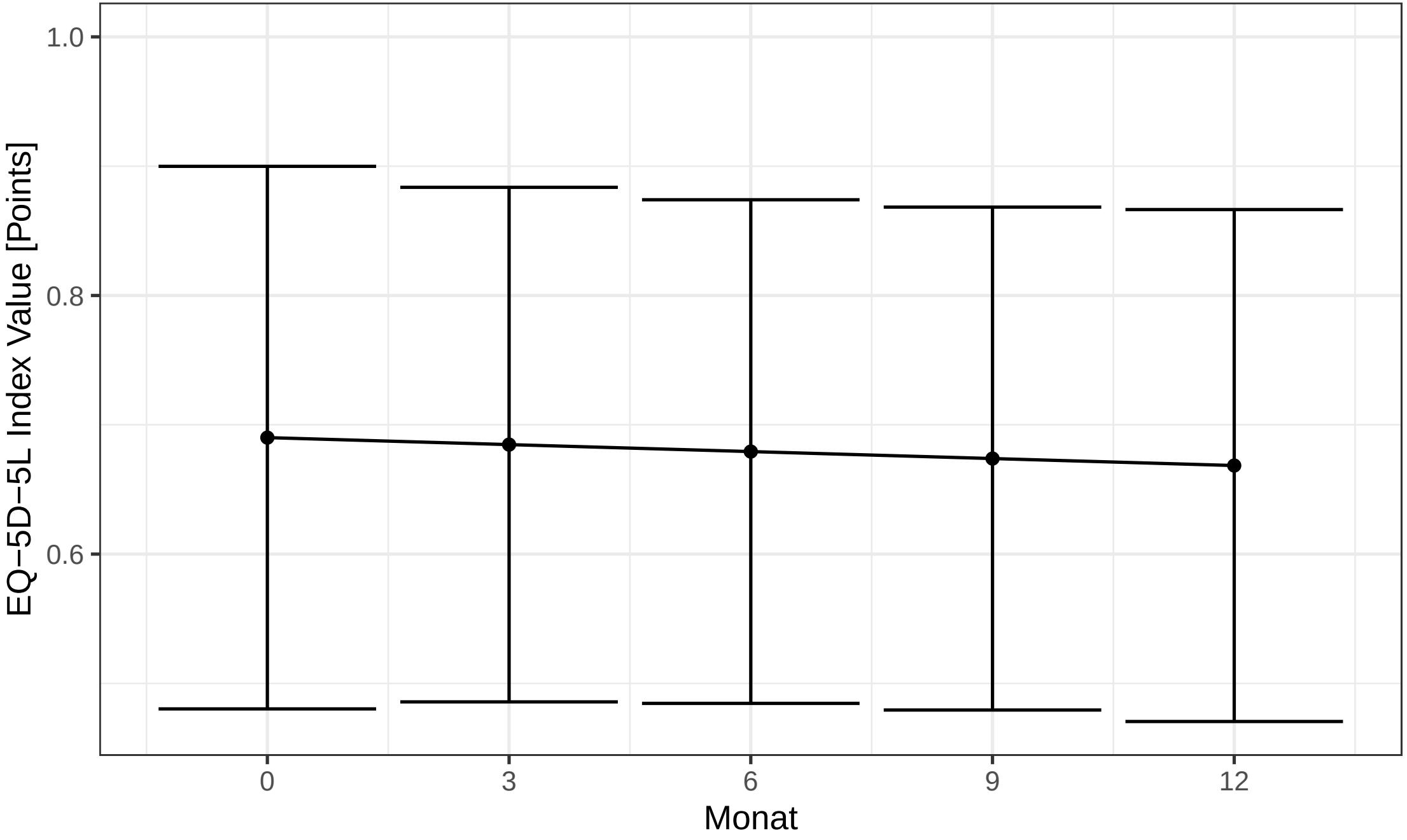
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

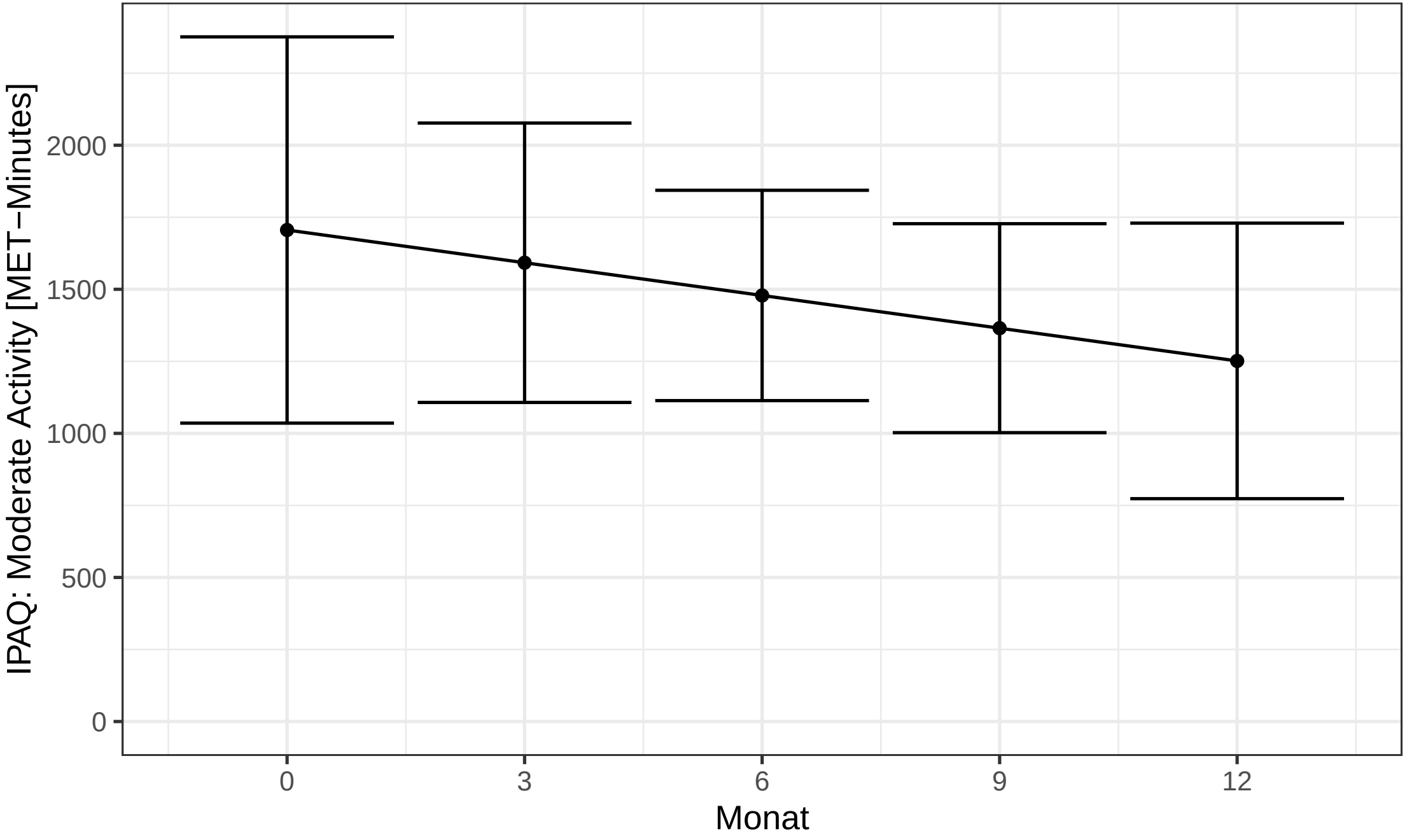
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

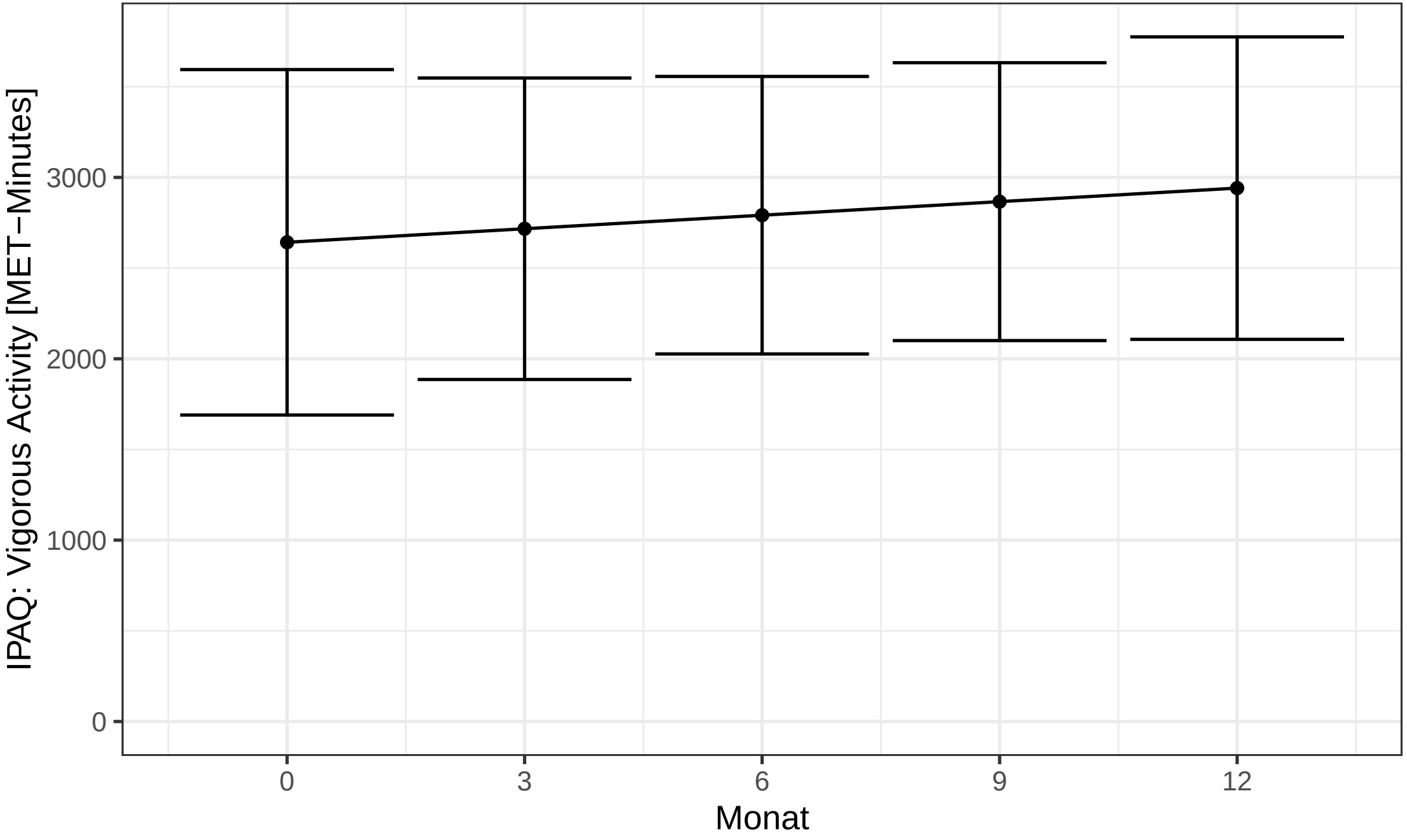
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

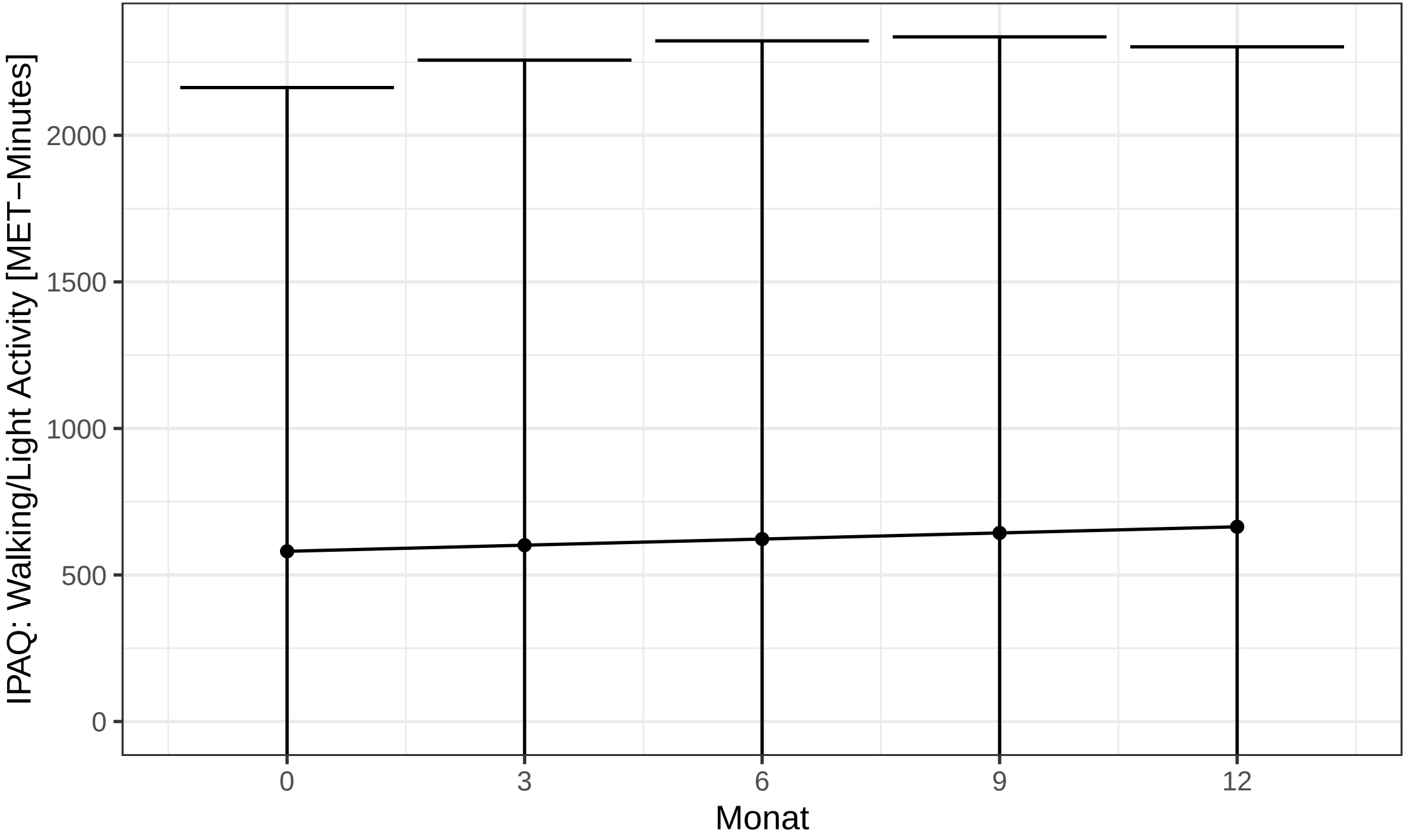
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

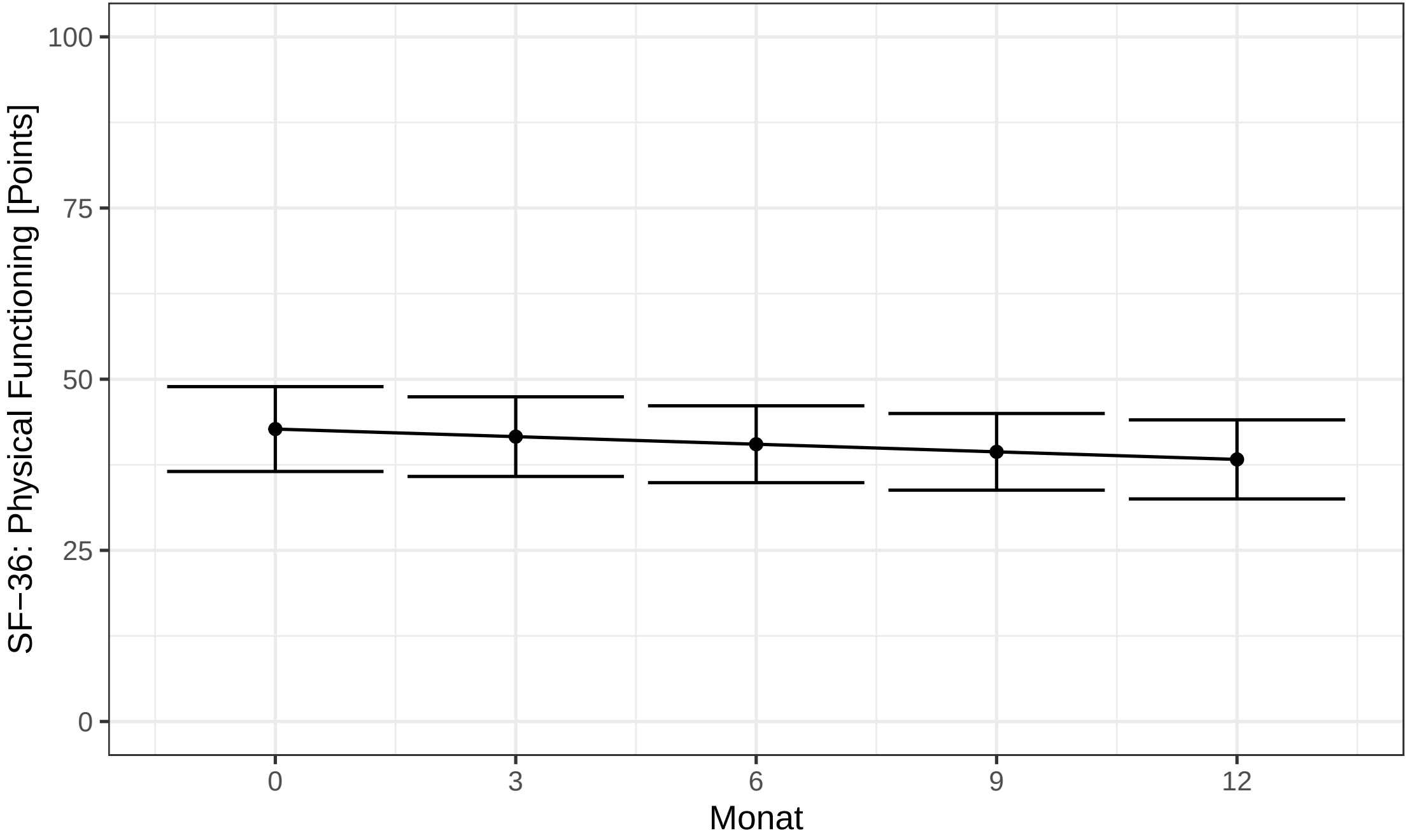
Adjustierter Mittelwert



Population C

Verlauf Interventionsgruppe

Adjustierter Mittelwert



Population C

Anlage 2: Ergebnisse aller Auswertepopulationen

Übersicht über die Auswertepopulationen:

- In Auswertepopulation A: Diese Auswertepopulation schloss alle Patient:innen ein, die zu einem Behandlungsarm allokiert wurden, unabhängig von der tatsächlichen Adhärenz. Fehlende Werte wurden imputiert.
- In Auswertepopulation B: Diese Auswertepopulation schloss alle Patient:innen, die der Kontrollgruppe zugewiesen wurden, und alle Patient:innen in der Interventionsgruppe, die die Reha wie geplant absolvierten, ein. Fehlende Werte wurden imputiert.
- Auswertepopulation P: Diese Auswertepopulation schloss alle Patient:innen aus, die die Studie nicht wie geplant abgeschlossen hatten. Fehlende Werte wurde imputiert.
- Auswertepopulation C: Diese Auswertepopulation schloss alle Patient:innen aus, die die Studie nicht wie geplant abgeschlossen hatten und fehlende Werte im primären Endpunkt aufwiesen. Fehlende Werte wurden nur in den Kovariaten imputiert.
- Complete-Case Analyse.

Tabelle 1: Ergebnisse der Analyse der Bell Disability Scale für alle Auswertepopulationen.

Bell Disability Scale [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	93	36,77 (7,96)	87	37,01 (7,79)		
Monat 12	93	38,06 (16,17)	87	33,33 (9,60)	-4,73 [-8,68; -0,79] η^2 = 0,030 P = 0,019	-3,78 [-7,67; 0,11] η^2 = 0,021 P = 0,057
Auswertepopulation B						
Baseline	100	36,50 (7,83)	97	36,80 (7,58)		
Monat 12	95	37,89 (16,04)	94	32,98 (9,48)	-4,92 [-8,70; -1,13] η^2 = 0,034 P = 0,011	-4,38 [-8,13; -0,64] η^2 = 0,029 P = 0,022
Auswertepopulation A						
Baseline	100	36,50 (7,83)	137	36,57 (7,32)		
Monat 12	95	37,89 (16,04)	113	33,45 (10,16)	-4,44 [-8,06; -0,83] η^2 = 0,028 P = 0,016	-4,19 [-7,58; -0,80] η^2 = 0,027 P = 0,016
Auswertepopulation P						
Baseline	99	36,46 (7,87)	95	36,84 (7,62)		
Monat 12	95	37,89 (16,04)	94	32,98 (9,48)	-4,92 [-8,70; -1,13] η^2 = 0,034 P = 0,011	-4,10 [-7,87; -0,34] η^2 = 0,025 P = 0,033
Auswertepopulation C						
Baseline	93	36,77 (7,96)	89	36,52 (7,25)		
Monat 12	93	37,96 (16,05)	89	33,15 (9,49)	-4,81 [-8,69; -0,93] η^2 = 0,032 P = 0,015	-3,70 [-7,53; 0,13] η^2 = 0,021 P = 0,058

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 2: Ergebnisse der Analyse der Chalder Fatige Scale (Gesamtscore) für alle Auswertepopulationen.

Chalder Fatigue Scale: Gesamtscore [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	88	27,65 (3,49)	83	27,95 (2,94)		
Monat 12	88	25,80 (4,45)	83	26,16 (4,87)	0,36 [-1,05; 1,77] η^2 = 0,002 P = 0,613	-0,06 [-1,47; 1,34] η^2 < 0,001 P = 0,928
Auswertepopulation B						
Baseline	95	27,86 (3,49)	93	27,76 (2,98)		
Monat 12	94	25,80 (4,52)	93	26,00 (4,88)	0,20 [-1,16; 1,56] η^2 < 0,001 P = 0,769	-0,17 [-1,47; 1,13] η^2 < 0,001 P = 0,796
Auswertepopulation A						
Baseline	95	27,86 (3,49)	132	27,53 (3,16)		
Monat 12	94	25,80 (4,52)	112	25,98 (4,87)	0,18 [-1,12; 1,49] η^2 < 0,001 P = 0,780	-0,19 [-1,35; 0,97] η^2 < 0,001 P = 0,746
Auswertepopulation P						
Baseline	94	27,82 (3,48)	91	27,78 (2,97)		
Monat 12	94	25,80 (4,52)	93	26,00 (4,88)	0,20 [-1,16; 1,56] η^2 < 0,001 P = 0,769	-0,14 [-1,46; 1,18] η^2 < 0,001 P = 0,833
Auswertepopulation C						
Baseline	89	27,71 (3,47)	86	27,73 (2,92)		
Monat 12	93	25,87 (4,49)	88	25,88 (4,88)	0,00 [-1,37; 1,38] η^2 < 0,001 P = 0,995	-0,29 [-1,68; 1,11] η^2 < 0,001 P = 0,686

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 3: Ergebnisse der Analyse der Chalder Fatige Scale (körperlich) für alle Auswertepopulationen.

Chalder Fatigue Scale: körperlich [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	91	18,47 (2,31)	86	18,37 (2,05)		
Monat 12	91	16,97 (3,13)	86	17,15 (3,04)	0,18 [-0,73; 1,10] η^2 < 0,001 P = 0,692	-0,00 [-0,94; 0,94] η^2 < 0,001 P = 0,995
Auswertepopulation B						
Baseline	98	18,57 (2,32)	95	18,28 (2,07)		
Monat 12	94	16,95 (3,12)	94	17,12 (3,04)	0,17 [-0,72; 1,06] η^2 < 0,001 P = 0,705	-0,03 [-0,90; 0,85] η^2 < 0,001 P = 0,951
Auswertepopulation A						
Baseline	98	18,57 (2,32)	135	18,26 (2,06)		
Monat 12	94	16,95 (3,12)	113	17,15 (3,04)	0,20 [-0,64; 1,05] η^2 = 0,001 P = 0,635	-0,00 [-0,79; 0,78] η^2 < 0,001 P = 0,993
Auswertepopulation P						
Baseline	97	18,55 (2,32)	93	18,27 (2,08)		
Monat 12	94	16,95 (3,12)	94	17,12 (3,04)	0,17 [-0,72; 1,06] η^2 < 0,001 P = 0,705	-0,02 [-0,91; 0,87] η^2 < 0,001 P = 0,966
Auswertepopulation C						
Baseline	92	18,48 (2,32)	88	18,26 (2,08)		
Monat 12	93	16,99 (3,11)	89	17,03 (3,05)	0,04 [-0,86; 0,95] η^2 < 0,001 P = 0,923	-0,15 [-1,09; 0,79] η^2 < 0,001 P = 0,751

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 4: Ergebnisse der Analyse der Chalder Fatige Scale (mental) für alle Auswertepopulationen.

Chalder Fatigue Scale: Mental [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	90	9,14 (1,92)	84	9,62 (1,67)		
Monat 12	90	8,89 (2,03)	84	9,00 (2,20)	0,11 [-0,52; 0,74] η^2 < 0,001 P = 0,729	-0,18 [-0,78; 0,43] η^2 = 0,002 P = 0,570
Auswertepopulation B						
Baseline	97	9,26 (1,91)	94	9,50 (1,75)		
Monat 12	95	8,87 (2,03)	93	8,89 (2,16)	0,02 [-0,58; 0,62] η^2 < 0,001 P = 0,951	-0,21 [-0,77; 0,36] η^2 = 0,003 P = 0,472
Auswertepopulation A						
Baseline	97	9,26 (1,91)	133	9,31 (1,93)		
Monat 12	95	8,87 (2,03)	112	8,84 (2,19)	-0,03 [-0,62; 0,55] η^2 < 0,001 P = 0,907	-0,24 [-0,76; 0,27] η^2 = 0,004 P = 0,354
Auswertepopulation P						
Baseline	96	9,24 (1,91)	92	9,53 (1,73)		
Monat 12	95	8,87 (2,03)	93	8,89 (2,16)	0,02 [-0,58; 0,62] η^2 < 0,001 P = 0,951	-0,19 [-0,76; 0,39] η^2 = 0,002 P = 0,527
Auswertepopulation C						
Baseline	90	9,20 (1,93)	87	9,49 (1,69)		
Monat 12	93	8,88 (2,02)	88	8,85 (2,15)	-0,03 [-0,64; 0,58] η^2 < 0,001 P = 0,924	-0,21 [-0,81; 0,38] η^2 = 0,003 P = 0,481

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 5: Ergebnisse der Analyse des Symptoms Kopfschmerzen nach CCC im Zeitverlauf für alle Auswertepopulationen.

CCC-Kriterien: Kopfschmerzen	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	Positiv (%)	n	Positiv (%)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	94	60 (63,83)	87	55 (63,22)		
Monat 12	94	57 (60,64)	87	54 (62,07)	1,06 [0,58; 1,94] P = 0,844	1,21 [0,58; 2,52] P = 0,617
Auswertepopulation B						
Baseline	100	65 (65,00)	96	59 (61,46)		
Monat 12	96	59 (61,46)	94	57 (60,64)	0,97 [0,54; 1,74] P = 0,908	1,13 [0,55; 2,31] P = 0,735
Auswertepopulation A						
Baseline	100	65 (65,00)	136	81 (59,56)		
Monat 12	96	59 (61,46)	113	70 (61,95)	1,02 [0,58; 1,79] P = 0,942	1,14 [0,58; 2,24] P = 0,704
Auswertepopulation P						
Baseline	99	64 (64,65)	94	59 (62,77)		
Monat 12	96	59 (61,46)	94	57 (60,64)	0,97 [0,54; 1,74] P = 0,908	1,11 [0,54; 2,26] P = 0,774
Auswertepopulation C						
Baseline	93	61 (65,59)	89	54 (60,67)		
Monat 12	93	57 (61,29)	89	52 (58,43)	0,89 [0,49; 1,61] P = 0,694	1,08 [0,52; 2,25] P = 0,839

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 6: Ergebnisse der Analyse des Symptoms Muskelschmerz nach CCC im Zeitverlauf.

CCC-Kriterien: Muskelschmerz	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	Positiv (%)	n	Positiv (%)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	94	81 (86,17)	87	79 (90,80)		
Monat 12	94	77 (81,91)	87	83 (95,40)	4,58 [1,47; 14,33] P = 0,009	3,27 [0,90; 11,86] P = 0,071
Auswertepopulation B						
Baseline	100	86 (86,00)	96	87 (90,62)		
Monat 12	96	79 (82,29)	94	90 (95,74)	4,84 [1,55; 15,10] P = 0,007	3,76 [1,01; 14,05] P = 0,049
Auswertepopulation A						
Baseline	100	86 (86,00)	136	122 (89,71)		
Monat 12	96	79 (82,29)	113	105 (92,92)	2,82 [1,15; 6,91] P = 0,023	2,73 [0,92; 8,07] P = 0,070
Auswertepopulation P						
Baseline	99	85 (85,86)	94	85 (90,43)		
Monat 12	96	79 (82,29)	94	90 (95,74)	4,84 [1,55; 15,10] P = 0,007	3,70 [1,01; 13,55] P = 0,048
Auswertepopulation C						
Baseline	93	80 (86,02)	89	82 (92,13)		
Monat 12	93	78 (83,87)	89	85 (95,51)	4,09 [1,29; 12,94] P = 0,017	2,65 [0,70; 10,00] P = 0,148

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 7: Ergebnisse der Analyse des COMPASS-31 Scores (Blase) für alle Auswertepopulationen.

COMPASS-31 Blase Score [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	92	1,04 (1,33)	86	1,16 (1,40)		
Monat 12	92	1,11 (1,38)	86	1,16 (1,35)	0,05 [-0,35; 0,46] η^2 < 0,001 P = 0,801	-0,13 [-0,45; 0,20] η^2 = 0,003 P = 0,444
Auswertepopulation B						
Baseline	99	1,25 (1,66)	95	1,24 (1,48)		
Monat 12	95	1,12 (1,41)	94	1,19 (1,40)	0,07 [-0,33; 0,47] η^2 < 0,001 P = 0,729	-0,10 [-0,42; 0,22] η^2 = 0,002 P = 0,537
Auswertepopulation A						
Baseline	99	1,25 (1,66)	134	1,37 (1,61)		
Monat 12	95	1,12 (1,41)	112	1,26 (1,47)	0,14 [-0,26; 0,53] η^2 = 0,002 P = 0,497	-0,05 [-0,35; 0,25] η^2 < 0,001 P = 0,725
Auswertepopulation P						
Baseline	98	1,20 (1,61)	93	1,23 (1,49)		
Monat 12	95	1,12 (1,41)	94	1,19 (1,40)	0,07 [-0,33; 0,47] η^2 < 0,001 P = 0,729	-0,11 [-0,43; 0,21] η^2 = 0,002 P = 0,507
Auswertepopulation C						
Baseline	92	1,01 (1,32)	88	1,19 (1,48)		
Monat 12	93	1,12 (1,41)	89	1,17 (1,41)	0,05 [-0,36; 0,46] η^2 < 0,001 P = 0,809	-0,15 [-0,47; 0,17] η^2 = 0,005 P = 0,348

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 8: Ergebnisse der Analyse des COMPASS-31 Scores (Gastroinestinaltrakt) für alle Auswertepopulationen.

COMPASS-31 Gastroinestinaltrakt Score [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	88	7,73 (3,68)	82	7,76 (3,68)		
Monat 12	88	8,46 (3,76)	82	8,02 (4,23)	-0,44 [-1,65; 0,77] η^2 = 0,003 P = 0,477	-0,24 [-1,26; 0,79] η^2 = 0,001 P = 0,650
Auswertepopulation B						
Baseline	98	7,68 (3,76)	95	7,62 (3,69)		
Monat 12	92	8,46 (3,88)	90	8,10 (4,23)	-0,37 [-1,55; 0,82] η^2 = 0,002 P = 0,542	-0,15 [-1,09; 0,79] η^2 < 0,001 P = 0,759
Auswertepopulation A						
Baseline	98	7,68 (3,76)	135	7,88 (3,74)		
Monat 12	92	8,46 (3,88)	108	8,18 (4,19)	-0,29 [-1,42; 0,85] η^2 = 0,001 P = 0,619	-0,22 [-1,09; 0,65] η^2 = 0,001 P = 0,613
Auswertepopulation P						
Baseline	97	7,65 (3,77)	93	7,58 (3,71)		
Monat 12	92	8,46 (3,88)	90	8,10 (4,23)	-0,37 [-1,55; 0,82] η^2 = 0,002 P = 0,542	-0,20 [-1,15; 0,75] η^2 < 0,001 P = 0,683
Auswertepopulation C						
Baseline	91	7,54 (3,82)	88	7,50 (3,66)		
Monat 12	90	8,54 (3,81)	85	7,90 (4,25)	-0,64 [-1,85; 0,56] η^2 = 0,006 P = 0,293	-0,36 [-1,38; 0,66] η^2 = 0,003 P = 0,485

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 9: Ergebnisse der Analyse des COMPASS-31 Scores (Orthostase) für alle Auswertepopulationen.

COMPASS-31 Orthostase Score [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	90	21,87 (9,97)	86	21,49 (11,45)		
Monat 12	90	21,11 (9,80)	86	19,77 (10,81)	-1,34 [-4,41; 1,72] η^2 = 0,004 P = 0,388	-1,90 [-4,87; 1,08] η^2 = 0,009 P = 0,210
Auswertepopulation B						
Baseline	99	22,46 (9,88)	95	21,73 (11,43)		
Monat 12	93	20,99 (9,67)	94	20,26 (10,81)	-0,73 [-3,69; 2,23] η^2 = 0,001 P = 0,625	-1,01 [-3,81; 1,80] η^2 = 0,003 P = 0,479
Auswertepopulation A						
Baseline	99	22,46 (9,88)	135	22,13 (11,40)		
Monat 12	93	20,99 (9,67)	112	20,75 (10,43)	-0,24 [-3,03; 2,55] η^2 < 0,001 P = 0,866	-0,69 [-3,17; 1,79] η^2 = 0,001 P = 0,586
Auswertepopulation P						
Baseline	98	22,53 (9,90)	93	21,63 (11,44)		
Monat 12	93	20,99 (9,67)	94	20,26 (10,81)	-0,73 [-3,69; 2,23] η^2 = 0,001 P = 0,625	-1,09 [-3,91; 1,73] η^2 = 0,003 P = 0,448
Auswertepopulation C						
Baseline	92	22,04 (9,91)	89	21,39 (11,60)		
Monat 12	91	21,14 (9,49)	89	19,78 (10,69)	-1,37 [-4,34; 1,60] η^2 = 0,005 P = 0,365	-1,95 [-4,85; 0,95] η^2 = 0,010 P = 0,187

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 10: Ergebnisse der Analyse des COMPASS-31 Scores (Pupillomotorik) für alle Auswertepopulationen.

COMPASS-31 Pupillomotorik Score [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	89	2,35 (1,20)	82	2,43 (1,15)		
Monat 12	89	2,71 (1,11)	82	2,74 (1,03)	0,03 [-0,30; 0,35] η^2 < 0,001 P = 0,865	-0,10 [-0,37; 0,17] η^2 = 0,003 P = 0,475
Auswertepopulation B						
Baseline	98	2,40 (1,18)	90	2,44 (1,15)		
Monat 12	93	2,71 (1,10)	94	2,73 (1,05)	0,02 [-0,29; 0,33] η^2 < 0,001 P = 0,913	-0,09 [-0,35; 0,17] η^2 = 0,003 P = 0,485
Auswertepopulation A						
Baseline	98	2,40 (1,18)	127	2,41 (1,09)		
Monat 12	93	2,71 (1,10)	113	2,72 (1,05)	0,01 [-0,28; 0,31] η^2 < 0,001 P = 0,931	-0,09 [-0,32; 0,15] η^2 = 0,002 P = 0,470
Auswertepopulation P						
Baseline	97	2,39 (1,18)	88	2,45 (1,16)		
Monat 12	93	2,71 (1,10)	94	2,73 (1,05)	0,02 [-0,29; 0,33] η^2 < 0,001 P = 0,913	-0,09 [-0,35; 0,17] η^2 = 0,003 P = 0,476
Auswertepopulation C						
Baseline	91	2,34 (1,19)	83	2,43 (1,16)		
Monat 12	91	2,72 (1,11)	89	2,68 (1,03)	-0,04 [-0,36; 0,28] η^2 < 0,001 P = 0,802	-0,16 [-0,43; 0,11] η^2 = 0,009 P = 0,233

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 11: Ergebnisse der Analyse des COMPASS-31 Scores (Sekretomotorik) für alle Auswertepopulationen.

COMPASS-31 Sekretomotorik Score [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	88	5,11 (3,67)	82	5,17 (3,64)		
Monat 12	88	5,45 (3,56)	82	5,67 (3,83)	0,22 [-0,90; 1,33] η^2 < 0,001 P = 0,703	0,05 [-0,93; 1,02] η^2 < 0,001 P = 0,922
Auswertepopulation B						
Baseline	97	5,17 (3,68)	92	5,29 (3,61)		
Monat 12	93	5,39 (3,54)	93	5,55 (3,81)	0,16 [-0,90; 1,23] η^2 < 0,001 P = 0,765	-0,24 [-1,15; 0,67] η^2 = 0,001 P = 0,608
Auswertepopulation A						
Baseline	97	5,17 (3,68)	132	5,10 (3,71)		
Monat 12	93	5,39 (3,54)	110	5,26 (3,98)	-0,13 [-1,18; 0,92] η^2 < 0,001 P = 0,805	-0,40 [-1,25; 0,46] η^2 = 0,004 P = 0,363
Auswertepopulation P						
Baseline	96	5,13 (3,69)	90	5,21 (3,62)		
Monat 12	93	5,39 (3,54)	93	5,55 (3,81)	0,16 [-0,90; 1,23] η^2 < 0,001 P = 0,765	-0,23 [-1,14; 0,68] η^2 = 0,001 P = 0,621
Auswertepopulation C						
Baseline	90	5,12 (3,66)	86	5,36 (3,60)		
Monat 12	91	5,46 (3,53)	88	5,53 (3,76)	0,06 [-1,01; 1,14] η^2 < 0,001 P = 0,906	-0,13 [-1,08; 0,83] η^2 < 0,001 P = 0,793

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 12: Ergebnisse der Analyse des COMPASS-31 Scores (Vasomotorik) für alle Auswertepopulationen.

COMPASS-31 Vasomotorik Score [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	92	0,85 (1,36)	82	0,93 (1,47)		
Monat 12	92	1,24 (1,53)	82	1,48 (1,64)	0,24 [-0,23; 0,72] η^2 = 0,006 P = 0,313	0,12 [-0,31; 0,55] η^2 = 0,002 P = 0,590
Auswertepopulation B						
Baseline	100	0,95 (1,44)	94	0,88 (1,43)		
Monat 12	94	1,23 (1,52)	91	1,44 (1,61)	0,21 [-0,25; 0,66] η^2 = 0,004 P = 0,372	0,15 [-0,25; 0,55] η^2 = 0,003 P = 0,470
Auswertepopulation A						
Baseline	100	0,95 (1,44)	134	1,03 (1,50)		
Monat 12	94	1,23 (1,52)	110	1,39 (1,59)	0,15 [-0,28; 0,58] η^2 = 0,002 P = 0,481	0,10 [-0,26; 0,47] η^2 = 0,001 P = 0,573
Auswertepopulation P						
Baseline	99	0,94 (1,44)	92	0,86 (1,42)		
Monat 12	94	1,23 (1,52)	91	1,44 (1,61)	0,21 [-0,25; 0,66] η^2 = 0,004 P = 0,372	0,14 [-0,26; 0,55] η^2 = 0,003 P = 0,481
Auswertepopulation C						
Baseline	93	0,91 (1,43)	88	0,85 (1,40)		
Monat 12	92	1,22 (1,51)	86	1,44 (1,60)	0,22 [-0,24; 0,68] η^2 = 0,005 P = 0,345	0,13 [-0,30; 0,56] η^2 = 0,002 P = 0,552

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 13: Ergebnisse der Analyse des COMPASS-31 Scores (Gesamtscore) für alle Auswertepopulationen.

COMPASS-31 Gesamtscore Score [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	78	39,32 (13,77)	69	37,71 (15,19)		
Monat 12	78	39,06 (14,76)	69	38,98 (13,98)	-0,08 [-4,79; 4,62] η^2 < 0,001 P = 0,972	-0,32 [-4,53; 3,89] η^2 < 0,001 P = 0,882
Auswertepopulation B						
Baseline	93	40,31 (13,96)	83	38,61 (15,25)		
Monat 12	87	39,36 (14,37)	87	39,72 (14,84)	0,36 [-4,01; 4,74] η^2 < 0,001 P = 0,870	-1,38 [-5,14; 2,38] η^2 = 0,003 P = 0,470
Auswertepopulation A						
Baseline	93	40,31 (13,96)	119	39,49 (15,43)		
Monat 12	87	39,36 (14,37)	103	40,09 (14,77)	0,72 [-3,47; 4,91] η^2 < 0,001 P = 0,733	-1,27 [-4,63; 2,10] η^2 = 0,003 P = 0,459
Auswertepopulation P						
Baseline	92	40,25 (14,02)	81	38,35 (15,27)		
Monat 12	87	39,36 (14,37)	87	39,72 (14,84)	0,36 [-4,01; 4,74] η^2 < 0,001 P = 0,870	-1,50 [-5,28; 2,28] η^2 = 0,003 P = 0,434
Auswertepopulation C						
Baseline	86	39,33 (13,53)	77	38,14 (15,52)		
Monat 12	85	39,67 (13,89)	82	38,93 (14,54)	-0,74 [-5,08; 3,61] η^2 < 0,001 P = 0,737	-0,92 [-4,99; 3,15] η^2 = 0,001 P = 0,656

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 14: Ergebnisse der Analyse des SF-36 (körperliche Funktionsfähigkeit) für alle Auswertepopulationen.

SF-36: Körperliche Funktionsfähigkeit [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	91	41,87 (18,81)	83	40,72 (18,66)		
Monat 12	91	42,09 (26,00)	83	37,41 (20,52)	-4,68 [-11,73; 2,38] η^2 = 0,010 P = 0,192	-1,98 [-8,53; 4,57] η^2 = 0,002 P = 0,552
Auswertepopulation B						
Baseline	99	40,35 (19,21)	94	40,43 (18,54)		
Monat 12	94	41,81 (25,86)	92	36,03 (20,64)	-5,78 [-12,55; 1,00] η^2 = 0,015 P = 0,094	-3,83 [-9,86; 2,19] η^2 = 0,008 P = 0,211
Auswertepopulation A						
Baseline	99	40,35 (19,21)	134	38,47 (18,54)		
Monat 12	94	41,81 (25,86)	111	35,23 (21,96)	-6,58 [-13,17; 0,00] η^2 = 0,019 P = 0,050	-4,09 [-9,45; 1,26] η^2 = 0,010 P = 0,133
Auswertepopulation P						
Baseline	98	40,56 (19,20)	92	40,76 (18,59)		
Monat 12	94	41,81 (25,86)	92	36,03 (20,64)	-5,78 [-12,55; 1,00] η^2 = 0,015 P = 0,094	-3,22 [-9,29; 2,85] η^2 = 0,006 P = 0,297
Auswertepopulation C						
Baseline	93	41,61 (18,68)	89	40,00 (18,28)		
Monat 12	93	41,61 (25,93)	89	36,85 (20,34)	-4,76 [-11,60; 2,08] η^2 = 0,010 P = 0,171	-1,87 [-8,24; 4,51] η^2 = 0,002 P = 0,564

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 15: Ergebnisse der Analyse des SF-36 (körperliche Schmerzen) für alle Auswertepopulationen.

SF-36: Körperliche Schmerzen [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	89	31,46 (20,80)	85	33,94 (22,69)		
Monat 12	89	37,67 (26,96)	85	33,24 (19,83)	-4,43 [-11,54; 2,68] η^2 = 0,009 P = 0,220	-3,17 [-9,69; 3,34] η^2 = 0,006 P = 0,338
Auswertepopulation B						
Baseline	98	30,71 (20,56)	94	33,22 (23,26)		
Monat 12	93	38,12 (26,93)	94	33,35 (19,56)	-4,77 [-11,55; 2,02] η^2 = 0,010 P = 0,167	-3,13 [-9,32; 3,05] η^2 = 0,005 P = 0,319
Auswertepopulation A						
Baseline	98	30,71 (20,56)	134	33,99 (23,18)		
Monat 12	93	38,12 (26,93)	113	33,25 (19,90)	-4,87 [-11,31; 1,58] η^2 = 0,011 P = 0,138	-3,87 [-9,58; 1,85] η^2 = 0,009 P = 0,184
Auswertepopulation P						
Baseline	97	30,70 (20,67)	92	33,72 (23,26)		
Monat 12	93	38,12 (26,93)	94	33,35 (19,56)	-4,77 [-11,55; 2,02] η^2 = 0,010 P = 0,167	-3,13 [-9,35; 3,08] η^2 = 0,005 P = 0,321
Auswertepopulation C						
Baseline	91	31,26 (20,88)	87	32,59 (21,94)		
Monat 12	92	37,80 (26,90)	89	33,60 (19,72)	-4,20 [-11,14; 2,73] η^2 = 0,008 P = 0,233	-2,21 [-8,67; 4,26] η^2 = 0,003 P = 0,502

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 16: Ergebnisse der Analyse des SF-36 (emotionales Wohlbefinden) für alle Auswertepopulationen.

SF-36: Emotionales Wohlbefinden [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	87	60,28 (17,08)	86	58,09 (17,05)		
Monat 12	87	63,82 (18,14)	86	63,67 (18,32)	-0,14 [-5,61; 5,33] η^2 < 0,001 P = 0,959	2,45 [-1,95; 6,86] η^2 = 0,007 P = 0,273
Auswertepopulation B						
Baseline	96	59,21 (18,28)	96	58,21 (17,76)		
Monat 12	93	63,14 (19,10)	92	63,57 (18,34)	0,43 [-5,01; 5,86] η^2 < 0,001 P = 0,877	2,76 [-1,47; 6,99] η^2 = 0,009 P = 0,200
Auswertepopulation A						
Baseline	96	59,21 (18,28)	135	58,55 (18,56)		
Monat 12	93	63,14 (19,10)	111	63,68 (18,69)	0,54 [-4,70; 5,77] η^2 < 0,001 P = 0,840	2,12 [-1,74; 5,98] η^2 = 0,005 P = 0,280
Auswertepopulation P						
Baseline	95	59,41 (18,26)	94	58,13 (17,93)		
Monat 12	93	63,14 (19,10)	92	63,57 (18,34)	0,43 [-5,01; 5,86] η^2 < 0,001 P = 0,877	2,74 [-1,49; 6,97] η^2 = 0,009 P = 0,202
Auswertepopulation C						
Baseline	89	60,90 (16,98)	89	57,57 (18,11)		
Monat 12	92	63,22 (19,19)	88	64,09 (17,87)	0,87 [-4,59; 6,33] η^2 < 0,001 P = 0,753	3,02 [-1,13; 7,17] η^2 = 0,012 P = 0,152

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 17: Ergebnisse der Analyse des SF-36 (allgemeine Gesundheitswahrnehmung) für alle Auswertepopulationen.

SF-36: Allgemeine Gesundheitswahrnehmung [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	90	27,61 (12,92)	86	27,62 (13,05)		
Monat 12	90	28,11 (14,19)	86	26,57 (11,41)	-1,54 [-5,38; 2,30] η^2 = 0,004 P = 0,430	-1,24 [-5,12; 2,63] η^2 = 0,002 P = 0,526
Auswertepopulation B						
Baseline	98	27,19 (12,98)	95	27,37 (13,31)		
Monat 12	93	27,96 (14,07)	93	26,18 (11,71)	-1,77 [-5,52; 1,97] η^2 = 0,005 P = 0,351	-1,39 [-5,07; 2,30] η^2 = 0,003 P = 0,458
Auswertepopulation A						
Baseline	98	27,19 (12,98)	135	28,44 (13,65)		
Monat 12	93	27,96 (14,07)	112	27,28 (12,48)	-0,68 [-4,34; 2,98] η^2 < 0,001 P = 0,714	-0,22 [-3,68; 3,25] η^2 < 0,001 P = 0,902
Auswertepopulation P						
Baseline	97	27,16 (13,05)	93	27,53 (13,40)		
Monat 12	93	27,96 (14,07)	93	26,18 (11,71)	-1,77 [-5,52; 1,97] η^2 = 0,005 P = 0,351	-1,42 [-5,11; 2,27] η^2 = 0,003 P = 0,448
Auswertepopulation C						
Baseline	92	27,50 (12,81)	88	27,16 (13,41)		
Monat 12	93	27,96 (14,07)	88	26,19 (11,35)	-1,76 [-5,53; 2,00] η^2 = 0,005 P = 0,356	-1,47 [-5,30; 2,36] η^2 = 0,003 P = 0,451

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 18: Ergebnisse der Analyse des SF-36 (Änderung des Gesundheitszustands) für alle Auswertepopulationen.

SF-36: Änderung des Gesundheitszustands [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	92	26,63 (30,74)	86	20,93 (28,92)		
Monat 12	92	52,99 (28,91)	86	41,86 (25,56)	-11,13 [-19,22; -3,03] η^2 = 0,040 P = 0,007	-8,75 [-17,06; -0,45] η^2 = 0,025 P = 0,039
Auswertepopulation B						
Baseline	100	25,50 (29,94)	96	20,57 (28,33)		
Monat 12	94	53,19 (28,69)	93	42,47 (25,75)	-10,72 [-18,59; -2,85] η^2 = 0,038 P = 0,008	-9,06 [-17,27; -0,86] η^2 = 0,026 P = 0,031
Auswertepopulation A						
Baseline	100	25,50 (29,94)	136	22,24 (28,49)		
Monat 12	94	53,19 (28,69)	112	42,63 (26,35)	-10,56 [-18,13; -2,99] η^2 = 0,036 P = 0,006	-8,52 [-16,53; -0,51] η^2 = 0,024 P = 0,037
Auswertepopulation P						
Baseline	99	25,76 (29,98)	94	20,48 (28,39)		
Monat 12	94	53,19 (28,69)	93	42,47 (25,75)	-10,72 [-18,59; -2,85] η^2 = 0,038 P = 0,008	-8,97 [-17,22; -0,72] η^2 = 0,026 P = 0,033
Auswertepopulation C						
Baseline	93	26,61 (30,58)	89	19,10 (27,96)		
Monat 12	93	52,96 (28,75)	89	43,26 (25,77)	-9,70 [-17,70; -1,70] η^2 = 0,031 P = 0,018	-7,27 [-15,57; 1,03] η^2 = 0,017 P = 0,086

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 19: Ergebnisse der Analyse des SF-36 (emotionale Rollenfunktion) für alle Auswertepopulationen.

SF-36: Emotionale Rollenfunktion [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	89	68,54 (42,45)	86	60,85 (43,74)		
Monat 12	89	68,16 (44,07)	86	76,36 (39,56)	8,19 [-4,32; 20,70] η^2 = 0,010 P = 0,198	13,67 [2,84; 24,50] η^2 = 0,036 P = 0,014
Auswertepopulation B						
Baseline	99	65,99 (43,64)	95	62,46 (43,82)		
Monat 12	92	69,20 (43,71)	94	75,18 (40,59)	5,97 [-6,22; 18,17] η^2 = 0,005 P = 0,335	12,14 [1,60; 22,68] η^2 = 0,028 P = 0,024
Auswertepopulation A						
Baseline	99	65,99 (43,64)	135	63,46 (43,84)		
Monat 12	92	69,20 (43,71)	113	72,27 (42,70)	3,07 [-8,88; 15,02] η^2 = 0,001 P = 0,613	8,96 [-1,30; 19,21] η^2 = 0,015 P = 0,086
Auswertepopulation P						
Baseline	98	65,65 (43,73)	93	61,65 (43,94)		
Monat 12	92	69,20 (43,71)	94	75,18 (40,59)	5,97 [-6,22; 18,17] η^2 = 0,005 P = 0,335	12,53 [2,01; 23,06] η^2 = 0,030 P = 0,020
Auswertepopulation C						
Baseline	92	68,84 (42,16)	88	60,98 (44,12)		
Monat 12	91	68,86 (43,83)	89	76,40 (39,95)	7,54 [-4,80; 19,88] η^2 = 0,008 P = 0,230	14,55 [3,76; 25,35] η^2 = 0,040 P = 0,009

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 20: Ergebnisse der Analyse des SF-36 (körperliche Rollenfunktion) für alle Auswertepopulationen.

SF-36: Körperliche Rollenfunktion [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	90	1,67 (8,21)	86	1,45 (7,02)		
Monat 12	90	6,94 (20,54)	86	1,74 (6,41)	-5,20 [-9,77; -0,63] η^2 = 0,028 P = 0,026	-4,94 [-9,76; -0,12] η^2 = 0,024 P = 0,045
Auswertepopulation B						
Baseline	99	1,52 (7,84)	96	1,30 (6,66)		
Monat 12	93	6,72 (20,24)	93	2,15 (7,95)	-4,57 [-9,02; -0,12] η^2 = 0,022 P = 0,044	-4,44 [-9,06; 0,17] η^2 = 0,020 P = 0,059
Auswertepopulation A						
Baseline	99	1,52 (7,84)	135	1,11 (6,01)		
Monat 12	93	6,72 (20,24)	112	3,35 (12,33)	-3,37 [-7,91; 1,16] η^2 = 0,010 P = 0,144	-3,55 [-8,34; 1,24] η^2 = 0,012 P = 0,145
Auswertepopulation P						
Baseline	98	1,53 (7,88)	94	1,33 (6,73)		
Monat 12	93	6,72 (20,24)	93	2,15 (7,95)	-4,57 [-9,02; -0,12] η^2 = 0,022 P = 0,044	-4,36 [-8,97; 0,25] η^2 = 0,019 P = 0,063
Auswertepopulation C						
Baseline	92	1,36 (7,74)	89	1,40 (6,91)		
Monat 12	92	6,79 (20,33)	88	2,27 (8,16)	-4,52 [-9,12; 0,08] η^2 = 0,021 P = 0,054	-4,33 [-9,20; 0,55] η^2 = 0,018 P = 0,082

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 21: Ergebnisse der Analyse des SF-36 (soziale Funktion) für alle Auswertepopulationen.

SF-36: Soziale Funktion [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	84	27,23 (20,09)	86	23,40 (18,17)		
Monat 12	84	33,78 (26,94)	86	21,66 (20,24)	-12,12 [-19,33; -4,92] η^2 = 0,062 P = 0,001	-8,49 [-15,72; -1,25] η^2 = 0,032 P = 0,022
Auswertepopulation B						
Baseline	97	25,52 (19,84)	96	23,57 (18,66)		
Monat 12	89	32,58 (26,82)	93	22,58 (21,67)	-10,00 [-17,12; -2,89] η^2 = 0,041 P = 0,006	-6,92 [-13,72; -0,12] η^2 = 0,022 P = 0,046
Auswertepopulation A						
Baseline	97	25,52 (19,84)	135	23,06 (18,57)		
Monat 12	89	32,58 (26,82)	112	23,77 (21,75)	-8,81 [-15,57; -2,06] η^2 = 0,032 P = 0,011	-5,79 [-12,01; 0,43] η^2 = 0,017 P = 0,068
Auswertepopulation P						
Baseline	96	25,65 (19,90)	94	23,80 (18,79)		
Monat 12	89	32,58 (26,82)	93	22,58 (21,67)	-10,00 [-17,12; -2,89] η^2 = 0,041 P = 0,006	-7,13 [-14,08; -0,18] η^2 = 0,023 P = 0,044
Auswertepopulation C						
Baseline	90	26,53 (19,73)	89	23,60 (18,70)		
Monat 12	89	32,58 (26,82)	88	21,88 (20,61)	-10,71 [-17,81; -3,61] η^2 = 0,048 P = 0,003	-7,94 [-15,05; -0,82] η^2 = 0,029 P = 0,029

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 22: Ergebnisse der Analyse des SF-36 (Vitalität) für alle Auswertepopulationen.

SF-36: Vitalität [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	90	16,39 (13,35)	85	15,82 (12,84)		
Monat 12	90	21,78 (18,62)	85	17,41 (15,31)	-4,37 [-9,47; 0,74] η^2 = 0,016 P = 0,093	-2,79 [-7,77; 2,18] η^2 = 0,007 P = 0,269
Auswertepopulation B						
Baseline	98	16,38 (13,67)	96	15,99 (12,77)		
Monat 12	94	21,38 (18,41)	91	18,30 (16,45)	-3,09 [-8,16; 1,98] η^2 = 0,008 P = 0,231	-1,58 [-6,35; 3,20] η^2 = 0,002 P = 0,515
Auswertepopulation A						
Baseline	98	16,38 (13,67)	136	17,83 (13,60)		
Monat 12	94	21,38 (18,41)	110	18,68 (16,43)	-2,70 [-7,51; 2,11] η^2 = 0,006 P = 0,270	-1,71 [-5,95; 2,53] η^2 = 0,003 P = 0,428
Auswertepopulation P						
Baseline	97	16,08 (13,43)	94	15,74 (12,74)		
Monat 12	94	21,38 (18,41)	91	18,30 (16,45)	-3,09 [-8,16; 1,98] η^2 = 0,008 P = 0,231	-1,60 [-6,40; 3,19] η^2 = 0,002 P = 0,510
Auswertepopulation C						
Baseline	91	16,37 (13,40)	89	15,73 (12,78)		
Monat 12	93	21,45 (18,50)	87	18,68 (16,61)	-2,77 [-7,96; 2,41] η^2 = 0,006 P = 0,292	-1,11 [-6,21; 3,98] η^2 = 0,001 P = 0,667

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 23: Ergebnisse der Analyse der Handgriffkraft Messungen (Fmean in Session 1) für alle Auswertepopulationen.

Handgriffkraft: Fmean in Session 1 [kg]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	88	16,24 (8,39)	82	15,92 (7,71)		
Monat 12	88	16,16 (8,24)	82	13,51 (7,71)	-2,66 [-5,08; -0,24] η^2 = 0,027 P = 0,032	-1,43 [-2,99; 0,13] η^2 = 0,020 P = 0,071
Auswertepopulation B						
Baseline	100	16,74 (9,07)	97	16,61 (7,71)		
Monat 12	89	16,13 (8,20)	89	13,61 (7,91)	-2,52 [-4,90; -0,13] η^2 = 0,024 P = 0,039	-1,62 [-3,08; -0,16] η^2 = 0,025 P = 0,029
Auswertepopulation A						
Baseline	100	16,74 (9,07)	137	17,56 (9,13)		
Monat 12	89	16,13 (8,20)	103	14,01 (7,97)	-2,11 [-4,42; 0,19] η^2 = 0,017 P = 0,072	-1,78 [-3,08; -0,47] η^2 = 0,031 P = 0,008
Auswertepopulation P						
Baseline	99	16,83 (9,08)	95	16,44 (7,70)		
Monat 12	89	16,13 (8,20)	89	13,61 (7,91)	-2,52 [-4,90; -0,13] η^2 = 0,024 P = 0,039	-1,56 [-3,04; -0,08] η^2 = 0,023 P = 0,038
Auswertepopulation C						
Baseline	93	17,02 (9,17)	89	16,34 (7,66)		
Monat 12	85	16,30 (8,25)	83	13,74 (7,85)	-2,55 [-5,01; -0,10] η^2 = 0,025 P = 0,042	-1,19 [-2,78; 0,41] η^2 = 0,013 P = 0,143

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 24: Ergebnisse der Analyse der Handgriffkraft Messungen (Fmean in Session 2) für alle Auswertepopulationen.

Handgriffkraft: Fmean in Session 2 [kg]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	87	14,61 (8,78)	81	14,38 (7,20)		
Monat 12	87	14,57 (8,30)	81	12,08 (7,57)	-2,48 [-4,91; -0,06] η^2 = 0,024 P = 0,045	-1,35 [-2,91; 0,22] η^2 = 0,018 P = 0,091
Auswertepopulation B						
Baseline	99	15,09 (9,22)	97	15,11 (7,27)		
Monat 12	89	14,58 (8,21)	88	12,19 (7,96)	-2,38 [-4,78; 0,02] η^2 = 0,021 P = 0,052	-1,60 [-3,04; -0,15] η^2 = 0,025 P = 0,031
Auswertepopulation A						
Baseline	99	15,09 (9,22)	137	16,14 (9,02)		
Monat 12	89	14,58 (8,21)	101	12,77 (7,98)	-1,81 [-4,13; 0,51] η^2 = 0,012 P = 0,125	-1,46 [-2,73; -0,19] η^2 = 0,023 P = 0,025
Auswertepopulation P						
Baseline	98	15,20 (9,21)	95	14,97 (7,29)		
Monat 12	89	14,58 (8,21)	88	12,19 (7,96)	-2,38 [-4,78; 0,02] η^2 = 0,021 P = 0,052	-1,60 [-3,06; -0,14] η^2 = 0,025 P = 0,032
Auswertepopulation C						
Baseline	92	15,35 (9,31)	89	14,86 (7,26)		
Monat 12	85	14,80 (8,25)	83	12,24 (7,80)	-2,56 [-5,00; -0,11] η^2 = 0,025 P = 0,041	-1,27 [-2,87; 0,33] η^2 = 0,015 P = 0,119

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 25: Ergebnisse der Analyse der Handgriffkraft Messungen (Fmax in Session 1) für alle Auswertepopulationen.

Handgriffkraft: Fmax in Session 1 [kg]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	88	20,03 (8,96)	82	19,78 (8,82)		
Monat 12	88	20,19 (9,68)	82	16,55 (9,32)	-3,63 [-6,52; -0,75] η^2 = 0,036 P = 0,014	-2,20 [-4,18; -0,22] η^2 = 0,029 P = 0,030
Auswertepopulation B						
Baseline	100	20,66 (9,56)	97	20,44 (8,91)		
Monat 12	89	20,13 (9,65)	89	16,71 (9,47)	-3,41 [-6,24; -0,59] η^2 = 0,031 P = 0,018	-2,54 [-4,55; -0,52] η^2 = 0,033 P = 0,014
Auswertepopulation A						
Baseline	100	20,66 (9,56)	137	21,61 (10,62)		
Monat 12	89	20,13 (9,65)	103	17,21 (9,49)	-2,92 [-5,65; -0,19] η^2 = 0,023 P = 0,036	-1,74 [-3,60; 0,11] η^2 = 0,016 P = 0,065
Auswertepopulation P						
Baseline	99	20,71 (9,60)	95	20,31 (8,95)		
Monat 12	89	20,13 (9,65)	89	16,71 (9,47)	-3,41 [-6,24; -0,59] η^2 = 0,031 P = 0,018	-2,61 [-4,61; -0,60] η^2 = 0,035 P = 0,011
Auswertepopulation C						
Baseline	93	20,86 (9,75)	89	20,23 (8,97)		
Monat 12	85	20,27 (9,68)	83	16,84 (9,44)	-3,43 [-6,35; -0,52] η^2 = 0,032 P = 0,021	-1,89 [-3,90; 0,12] η^2 = 0,021 P = 0,065

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 26: Ergebnisse der Analyse der Handgriffkraft Messungen (Fmax in Session 2) für alle Auswertepopulationen.

Handgriffkraft: Fmax in Session 2 [kg]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	87	17,79 (9,61)	81	17,84 (8,03)		
Monat 12	87	17,77 (9,41)	81	15,00 (9,28)	-2,76 [-5,61; 0,09] η^2 = 0,022 P = 0,057	-1,64 [-3,56; 0,28] η^2 = 0,018 P = 0,093
Auswertepopulation B						
Baseline	99	18,44 (10,21)	97	18,45 (8,17)		
Monat 12	89	17,77 (9,31)	88	15,13 (9,76)	-2,64 [-5,47; 0,19] η^2 = 0,019 P = 0,067	-1,55 [-3,42; 0,33] η^2 = 0,014 P = 0,106
Auswertepopulation A						
Baseline	99	18,44 (10,21)	137	19,64 (9,95)		
Monat 12	89	17,77 (9,31)	101	15,72 (9,67)	-2,05 [-4,78; 0,67] η^2 = 0,012 P = 0,139	-0,62 [-2,34; 1,10] η^2 = 0,002 P = 0,478
Auswertepopulation P						
Baseline	98	18,55 (10,19)	95	18,35 (8,22)		
Monat 12	89	17,77 (9,31)	88	15,13 (9,76)	-2,64 [-5,47; 0,19] η^2 = 0,019 P = 0,067	-1,77 [-3,64; 0,11] η^2 = 0,019 P = 0,065
Auswertepopulation C						
Baseline	92	18,68 (10,41)	89	18,23 (8,22)		
Monat 12	85	17,93 (9,27)	83	15,20 (9,63)	-2,73 [-5,61; 0,15] η^2 = 0,021 P = 0,063	-1,47 [-3,40; 0,46] η^2 = 0,014 P = 0,133

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 27: Ergebnisse der Analyse der Handgriffkraft Messungen (Fatigue in Ratio in Session 1) für alle Auswertepopulationen.

Handgriffkraft: Fatigue Ratio in Session 1	Kontrollgruppe		Interventionsgruppe		Differenz (%) [95-% Konfidenzintervall]	
	n	MW (SD) Median (IQR)	n	MW (SD) Median (IQR)	Unadjustiert[c]	Adjustiert[a,b,c]
Complete-Case						
Baseline	88	1,30 (0,24) 1,23 (1,16 — 1,38)	82	1,29 (0,19) 1,25 (1,15 — 1,39)		
Monat 12	88	1,30 (0,28) 1,24 (1,14 — 1,39)	82	1,25 (0,19) 1,21 (1,14 — 1,29)	-3,17 % [-7,72; 1,60] η^2 = 0,010 P = 0,187	-3,06 % [-7,36; 1,44] η^2 = 0,011 P = 0,179
Auswertepopulation B						
Baseline	100	1,32 (0,28) 1,23 (1,15 — 1,40)	97	1,27 (0,19) 1,24 (1,12 — 1,38)		
Monat 12	89	1,30 (0,28) 1,24 (1,14 — 1,39)	89	1,28 (0,27) 1,21 (1,14 — 1,29)	-1,82 % [-6,74; 3,37] η^2 = 0,003 P = 0,483	-1,75 % [-7,07; 3,87] η^2 = 0,002 P = 0,531
Auswertepopulation A						
Baseline	100	1,32 (0,28) 1,23 (1,15 — 1,40)	137	1,27 (0,19) 1,24 (1,12 — 1,38)		
Monat 12	89	1,30 (0,28) 1,24 (1,14 — 1,39)	103	1,28 (0,28) 1,20 (1,14 — 1,30)	-1,68 % [-6,49; 3,38] η^2 = 0,002 P = 0,507	3,40 % [-2,22; 9,34] η^2 = 0,007 P = 0,240
Auswertepopulation P						
Baseline	99	1,31 (0,28) 1,23 (1,15 — 1,39)	95	1,27 (0,19) 1,24 (1,13 — 1,38)		
Monat 12	89	1,30 (0,28) 1,24 (1,14 — 1,39)	89	1,28 (0,27) 1,21 (1,14 — 1,29)	-1,82 % [-6,74; 3,37] η^2 = 0,003 P = 0,483	-2,11 % [-7,34; 3,41] η^2 = 0,003 P = 0,443
Auswertepopulation C						
Baseline	93	1,31 (0,28) 1,23 (1,15 — 1,37)	89	1,28 (0,19) 1,24 (1,13 — 1,38)		
Monat 12	85	1,30 (0,28) 1,23 (1,12 — 1,37)	83	1,27 (0,26) 1,19 (1,14 — 1,28)	-2,12 % [-7,09; 3,11] η^2 = 0,004 P = 0,417	-2,60 % [-7,41; 2,47] η^2 = 0,007 P = 0,307

MW: Mittelwert (arithmetisch), SD: Standardabweichung, IQR: Interquartilsspannweite, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

[c] Die Differenzen stellen prozentuale Veränderungen dar; abgeleitet aus einem Modell mit logarithmisch transformierter abhängiger Variable.

Tabelle 28: Ergebnisse der Analyse der Handgriffkraft Messungen (Fatigue in Ratio in Session 2) für alle Auswertepopulationen.

Handgriffkraft: Fatigue Ratio in Session 2	Kontrollgruppe		Interventionsgruppe		Differenz (%) [95-% Konfidenzintervall]	
	n	MW (SD) Median (IQR)	n	MW (SD) Median (IQR)	Unadjustiert[c]	Adjustiert[a,b,c]
Complete-Case						
Baseline	87	1,31 (0,33) 1,21 (1,13 — 1,36)	81	1,30 (0,21) 1,25 (1,15 — 1,36)		
Monat 12	87	1,27 (0,20) 1,23 (1,15 — 1,33)	81	1,31 (0,50) 1,23 (1,16 — 1,35)	1,41 % [-3,78; 6,87] η^2 = 0,002 P = 0,600	-0,15 % [-5,28; 5,25] η^2 < 0,001 P = 0,955
Auswertepopulation B						
Baseline	99	1,35 (0,48) 1,21 (1,13 — 1,37)	97	1,27 (0,20) 1,23 (1,13 — 1,35)		
Monat 12	89	1,27 (0,20) 1,23 (1,15 — 1,33)	88	1,33 (0,52) 1,23 (1,16 — 1,36)	2,40 % [-2,98; 8,08] η^2 = 0,004 P = 0,387	2,02 % [-3,99; 8,40] η^2 = 0,002 P = 0,517
Auswertepopulation A						
Baseline	99	1,35 (0,48) 1,21 (1,13 — 1,37)	137	1,27 (0,20) 1,21 (1,14 — 1,35)		
Monat 12	89	1,27 (0,20) 1,23 (1,15 — 1,33)	101	1,31 (0,48) 1,23 (1,16 — 1,34)	1,34 % [-3,70; 6,66] η^2 = 0,001 P = 0,607	6,46 % [0,12; 13,21] η^2 = 0,019 P = 0,046
Auswertepopulation P						
Baseline	98	1,34 (0,49) 1,21 (1,13 — 1,36)	95	1,28 (0,20) 1,23 (1,14 — 1,35)		
Monat 12	89	1,27 (0,20) 1,23 (1,15 — 1,33)	88	1,33 (0,52) 1,23 (1,16 — 1,36)	2,40 % [-2,98; 8,08] η^2 = 0,004 P = 0,387	1,72 % [-4,17; 7,98] η^2 = 0,002 P = 0,572
Auswertepopulation C						
Baseline	92	1,34 (0,49) 1,21 (1,13 — 1,33)	89	1,28 (0,21) 1,23 (1,14 — 1,35)		
Monat 12	85	1,26 (0,19) 1,23 (1,15 — 1,32)	83	1,31 (0,49) 1,24 (1,16 — 1,36)	2,02 % [-3,12; 7,43] η^2 = 0,003 P = 0,447	-0,00 % [-5,11; 5,38] η^2 < 0,001 P = 1,000

MW: Mittelwert (arithmetisch), SD: Standardabweichung, IQR: Interquartilsspannweite, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

[c] Die Differenzen stellen prozentuale Veränderungen dar; abgeleitet aus einem Modell mit logarithmisch transformierter abhängiger Variable.

Tabelle 29: Ergebnisse der Analyse der Handgriffkraft Messungen (Recovery Ratio) für alle Auswertepopulationen.

Handgriffkraft: Recovery Ratio	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	87	0,86 (0,20)	81	0,91 (0,18)		
Monat 12	87	0,88 (0,18)	81	0,88 (0,19)	0,00 [-0,05; 0,06] η^2 < 0,001 P = 0,977	-0,00 [-0,06; 0,05] η^2 < 0,001 P = 0,926
Auswertepopulation B						
Baseline	99	0,86 (0,21)	97	0,91 (0,17)		
Monat 12	89	0,88 (0,18)	88	0,87 (0,20)	-0,02 [-0,07; 0,04] η^2 = 0,002 P = 0,575	-0,01 [-0,07; 0,05] η^2 = 0,001 P = 0,676
Auswertepopulation A						
Baseline	99	0,86 (0,21)	137	0,91 (0,17)		
Monat 12	89	0,88 (0,18)	101	0,88 (0,19)	-0,01 [-0,06; 0,05] η^2 < 0,001 P = 0,844	0,02 [-0,04; 0,08] η^2 = 0,002 P = 0,564
Auswertepopulation P						
Baseline	98	0,86 (0,21)	95	0,91 (0,17)		
Monat 12	89	0,88 (0,18)	88	0,87 (0,20)	-0,02 [-0,07; 0,04] η^2 = 0,002 P = 0,575	-0,02 [-0,08; 0,04] η^2 = 0,002 P = 0,595
Auswertepopulation C						
Baseline	92	0,86 (0,21)	89	0,91 (0,18)		
Monat 12	85	0,89 (0,18)	83	0,87 (0,18)	-0,02 [-0,07; 0,04] η^2 = 0,002 P = 0,536	-0,01 [-0,07; 0,04] η^2 = 0,002 P = 0,618

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 30: Ergebnisse der Analyse der Positive Schellong-Tests (nach 0 Minuten Stehen) für alle Auswertepopulationen.

Schellong-Test nach 0 Minuten Stehen	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	Positiv (%)	n	Positiv (%)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	86	49 (56,98)	81	24 (29,63)		
Monat 12	86	32 (37,21)	81	24 (29,63)	0,71 [0,37; 1,36] P = 0,302	1,04 [0,49; 2,22] P = 0,911
Auswertepopulation B						
Baseline	99	51 (51,52)	96	27 (28,12)		
Monat 12	87	33 (37,93)	88	26 (29,55)	0,69 [0,36; 1,29] P = 0,243	1,00 [0,49; 2,02] P = 0,998
Auswertepopulation A						
Baseline	99	51 (51,52)	135	36 (26,67)		
Monat 12	87	33 (37,93)	102	31 (30,39)	0,71 [0,39; 1,31] P = 0,277	1,15 [0,59; 2,25] P = 0,679
Auswertepopulation P						
Baseline	98	51 (52,04)	94	26 (27,66)		
Monat 12	87	33 (37,93)	88	26 (29,55)	0,69 [0,36; 1,29] P = 0,243	0,96 [0,47; 1,96] P = 0,907
Auswertepopulation C						
Baseline	92	50 (54,35)	88	26 (29,55)		
Monat 12	85	31 (36,47)	83	25 (30,12)	0,75 [0,39; 1,44] P = 0,385	1,09 [0,51; 2,33] P = 0,815

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 31: Ergebnisse der Analyse der Positive Schellong-Tests (nach 2 Minuten Stehen) für alle Auswertepopulationen.

Schellong-Test nach 2 Minuten Stehen	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	Positiv (%)	n	Positiv (%)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	83	42 (50,60)	80	23 (28,75)		
Monat 12	83	40 (48,19)	80	27 (33,75)	0,55 [0,29; 1,04] P = 0,064	0,81 [0,38; 1,73] P = 0,585
Auswertepopulation B						
Baseline	98	47 (47,96)	95	25 (26,32)		
Monat 12	85	41 (48,24)	88	30 (34,09)	0,56 [0,30; 1,03] P = 0,061	0,84 [0,40; 1,75] P = 0,639
Auswertepopulation A						
Baseline	98	47 (47,96)	134	40 (29,85)		
Monat 12	85	41 (48,24)	102	35 (34,31)	0,56 [0,31; 1,02] P = 0,056	0,88 [0,44; 1,73] P = 0,700
Auswertepopulation P						
Baseline	97	46 (47,42)	93	24 (25,81)		
Monat 12	85	41 (48,24)	88	30 (34,09)	0,56 [0,30; 1,03] P = 0,061	0,82 [0,40; 1,68] P = 0,580
Auswertepopulation C						
Baseline	91	43 (47,25)	87	23 (26,44)		
Monat 12	83	39 (46,99)	83	29 (34,94)	0,61 [0,32; 1,14] P = 0,117	0,89 [0,42; 1,89] P = 0,757

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 32: Ergebnisse der Analyse der Positive Schellong-Tests (nach 5 Minuten Stehen) für alle Auswertepopulationen.

Schellong-Test nach 5 Minuten Stehen	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	Positiv (%)	n	Positiv (%)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	80	46 (57,50)	75	26 (34,67)		
Monat 12	80	45 (56,25)	75	25 (33,33)	0,39 [0,20; 0,75] P = 0,005	0,54 [0,26; 1,13] P = 0,101
Auswertepopulation B						
Baseline	97	53 (54,64)	94	30 (31,91)		
Monat 12	82	47 (57,32)	81	27 (33,33)	0,37 [0,20; 0,71] P = 0,003	0,56 [0,28; 1,12] P = 0,099
Auswertepopulation A						
Baseline	97	53 (54,64)	133	45 (33,83)		
Monat 12	82	47 (57,32)	95	33 (34,74)	0,40 [0,21; 0,73] P = 0,003	0,69 [0,35; 1,36] P = 0,279
Auswertepopulation P						
Baseline	96	53 (55,21)	92	29 (31,52)		
Monat 12	82	47 (57,32)	81	27 (33,33)	0,37 [0,20; 0,71] P = 0,003	0,57 [0,28; 1,16] P = 0,120
Auswertepopulation C						
Baseline	90	51 (56,67)	86	28 (32,56)		
Monat 12	80	45 (56,25)	76	27 (35,53)	0,43 [0,22; 0,82] P = 0,011	0,62 [0,30; 1,30] P = 0,204

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 33: Ergebnisse der Analyse der Positive Schellong-Tests (nach 10 Minuten Stehen) für alle Auswertepopulationen.

Schellong-Test nach 10 Minuten Stehen	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	Positiv (%)	n	Positiv (%)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	78	49 (62,82)	70	23 (32,86)		
Monat 12	78	41 (52,56)	70	26 (37,14)	0,53 [0,27; 1,04] P = 0,063	0,89 [0,40; 2,00] P = 0,776
Auswertepopulation B						
Baseline	95	56 (58,95)	92	33 (35,87)		
Monat 12	80	42 (52,50)	76	28 (36,84)	0,53 [0,28; 1,01] P = 0,052	0,89 [0,43; 1,86] P = 0,755
Auswertepopulation A						
Baseline	95	56 (58,95)	131	49 (37,40)		
Monat 12	80	42 (52,50)	88	33 (37,50)	0,54 [0,29; 1,01] P = 0,053	0,99 [0,51; 1,92] P = 0,978
Auswertepopulation P						
Baseline	94	56 (59,57)	90	32 (35,56)		
Monat 12	80	42 (52,50)	76	28 (36,84)	0,53 [0,28; 1,01] P = 0,052	0,90 [0,44; 1,84] P = 0,764
Auswertepopulation C						
Baseline	88	55 (62,50)	85	31 (36,47)		
Monat 12	78	40 (51,28)	71	26 (36,62)	0,55 [0,28; 1,06] P = 0,075	0,93 [0,42; 2,06] P = 0,854

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 34: Ergebnisse der Analyse der Ddurchschnittlichen Schrittzahl pro Tag nach Zeitpunkten (mittels Troiano-Algorithmus ermittelt) für alle Auswertepopulationen.

Schrittzahl: Akzelerometer [Schritte/Tag] (nach Troiano)	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	56	4894,26 (2828,90)	45	4983,83 (2509,99)		
Monat 12	56	4296,44 (2681,43)	45	4660,16 (2894,30)	363,73 [-739,82; 1467,27] η^2 = 0,004 P = 0,515	198,87 [-580,51; 978,25] η^2 = 0,003 P = 0,614
Auswertepopulation B						
Baseline	88	5052,41 (2647,52)	91	4679,40 (2559,23)		
Monat 12	62	4303,14 (2684,13)	50	4654,60 (2875,80)	351,46 [-692,39; 1395,32] η^2 = 0,004 P = 0,506	-458,59 [-3185,08; 2267,91] η^2 = 0,005 P = 0,735
Auswertepopulation A						
Baseline	88	5052,41 (2647,52)	118	4498,32 (2551,62)		
Monat 12	62	4303,14 (2684,13)	60	4453,58 (2769,45)	150,44 [-827,14; 1128,01] η^2 < 0,001 P = 0,761	-648,94 [-3377,94; 2080,05] η^2 = 0,003 P = 0,636
Auswertepopulation P						
Baseline	88	5052,41 (2647,52)	89	4713,91 (2559,30)		
Monat 12	62	4303,14 (2684,13)	50	4654,60 (2875,80)	351,46 [-692,39; 1395,32] η^2 = 0,004 P = 0,506	-498,07 [-2714,48; 1718,34] η^2 = 0,004 P = 0,653
Auswertepopulation C						
Baseline	84	5112,48 (2684,97)	84	4642,61 (2462,40)		
Monat 12	58	4420,48 (2696,39)	47	4557,48 (2879,06)	137,00 [-944,86; 1218,86] η^2 < 0,001 P = 0,802	148,97 [-647,94; 945,88] η^2 = 0,002 P = 0,711

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 35: Ergebnisse der Analyse des IPAQ (gehen/leichte Aktivität) für alle Auswertepopulationen.

IPAQ: Gehen/leichte Aktivität [MET-Minuten/ Woche]	Kontrollgruppe		Interventionsgruppe		Differenz (%) [95-% Konfidenzintervall]	
	n	MW (SD) Median (IQR)	n	MW (SD) Median (IQR)	Unadjustiert[c]	Adjustiert[a,b,c]
Complete-Case						
Baseline	4	1831,50 (2537,57) 792,00 (198,00 — 2425,50)	9	575,67 (698,69) 396,00 (198,00 — 594,00)		
Monat 12	4	1361,25 (1868,58) 495,00 (371,25 — 1485,00)	9	876,33 (1176,04) 594,00 (297,00 — 693,00)	-23,56 % [-78,92; 177,16] η^2 = 0,019 P = 0,655	-1,54 % [-91,23; 1005,75] η^2 < 0,001 P = 0,987
Auswertepopulation B						
Baseline	11	865,50 (1598,77) 297,00 (173,25 — 594,00)	17	561,97 (623,45) 247,50 (198,00 — 594,00)		
Monat 12	17	670,68 (954,24) 396,00 (198,00 — 594,00)	13	680,31 (1010,28) 396,00 (264,00 — 693,00)	0,79 % [-51,29; 108,56] η^2 < 0,001 P = 0,983	-5,95 % [-88,38; 661,01] η^2 = 0,001 P = 0,938
Auswertepopulation A						
Baseline	11	865,50 (1598,77) 297,00 (173,25 — 594,00)	19	638,29 (671,65) 396,00 (198,00 — 693,00)		
Monat 12	17	670,68 (954,24) 396,00 (198,00 — 594,00)	14	695,36 (972,28) 396,00 (272,25 — 693,00)	6,57 % [-47,42; 115,99] η^2 = 0,001 P = 0,855	-5,23 % [-88,34; 670,35] η^2 = 0,001 P = 0,945
Auswertepopulation P						
Baseline	11	865,50 (1598,77) 297,00 (173,25 — 594,00)	17	561,97 (623,45) 247,50 (198,00 — 594,00)		
Monat 12	17	670,68 (954,24) 396,00 (198,00 — 594,00)	13	680,31 (1010,28) 396,00 (264,00 — 693,00)	0,79 % [-51,29; 108,56] η^2 < 0,001 P = 0,983	-5,59 % [-88,36; 665,71] η^2 = 0,001 P = 0,941
Auswertepopulation C						
Baseline	10	942,15 (1663,81) 346,50 (198,00 — 693,00)	15	597,30 (656,85) 247,50 (181,50 — 693,00)		
Monat 12	16	708,47 (972,31) 396,00 (272,25 — 594,00)	12	720,50 (1044,29) 396,00 (288,75 — 693,00)	-4,37 % [-53,54; 96,84] η^2 < 0,001 P = 0,900	29,00 % [-77,39; 636,15] η^2 = 0,056 P = 0,666

MW: Mittelwert (arithmetisch), SD: Standardabweichung, IQR: Interquartilsspannweite, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

[c] Die Differenzen stellen prozentuale Veränderungen dar; abgeleitet aus einem Modell mit logarithmisch transformierter abhängiger Variable.

Tabelle 36: Ergebnisse der Analyse des IPAQ (moderate Aktivität) für alle Auswertepopulationen.

IPAQ: Moderate Aktivität [MET-Minuten/ Woche]	Kontrollgruppe		Interventionsgruppe		Differenz (%) [95-% Konfidenzintervall]	
	n	MW (SD) Median (IQR)	n	MW (SD) Median (IQR)	Unadjustiert[c]	Adjustiert[a,b,c]
Complete-Case						
Baseline	34	1852,94 (3220,54) 640,00 (260,00 — 1170,00)	27	1406,07 (1365,18) 1064,00 (540,00 — 1680,00)		
Monat 12	34	1254,71 (1774,08) 480,00 (330,00 — 960,00)	27	1257,93 (1090,10) 960,00 (480,00 — 1680,00)	26,92 % [-25,59; 116,50] η^2 = 0,013 P = 0,375	4,71 % [-43,01; 92,36] η^2 < 0,001 P = 0,880
Auswertepopulation B						
Baseline	55	1354,69 (2638,54) 480,00 (190,00 — 960,00)	52	1264,62 (1171,24) 880,00 (360,00 — 1680,00)		
Monat 12	54	968,15 (1477,53) 480,00 (240,00 — 960,00)	36	1185,67 (1074,93) 750,00 (450,00 — 1680,00)	61,82 % [-0,79; 163,94] η^2 = 0,042 P = 0,054	45,02 % [-11,76; 138,34] η^2 = 0,022 P = 0,140
Auswertepopulation A						
Baseline	55	1354,69 (2638,54) 480,00 (190,00 — 960,00)	70	1358,11 (2160,84) 720,00 (360,00 — 1680,00)		
Monat 12	54	968,15 (1477,53) 480,00 (240,00 — 960,00)	42	1131,05 (1041,62) 720,00 (375,00 — 1640,00)	55,79 % [-1,60; 146,66] η^2 = 0,038 P = 0,058	38,37 % [-12,47; 118,72] η^2 = 0,018 P = 0,162
Auswertepopulation P						
Baseline	55	1354,69 (2638,54) 480,00 (190,00 — 960,00)	52	1264,62 (1171,24) 880,00 (360,00 — 1680,00)		
Monat 12	54	968,15 (1477,53) 480,00 (240,00 — 960,00)	36	1185,67 (1074,93) 750,00 (450,00 — 1680,00)	61,82 % [-0,79; 163,94] η^2 = 0,042 P = 0,054	50,69 % [-9,17; 149,98] η^2 = 0,026 P = 0,111
Auswertepopulation C						
Baseline	53	1402,79 (2676,71) 480,00 (240,00 — 960,00)	48	1229,17 (1187,67) 760,00 (344,00 — 1680,00)		
Monat 12	52	993,08 (1499,51) 480,00 (240,00 — 960,00)	35	1144,11 (1060,89) 720,00 (420,00 — 1600,00)	49,71 % [-8,18; 144,10] η^2 = 0,031 P = 0,104	1,52 % [-44,69; 86,34] η^2 < 0,001 P = 0,961

MW: Mittelwert (arithmetisch), SD: Standardabweichung, IQR: Interquartilsspannweite, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

[c] Die Differenzen stellen prozentuale Veränderungen dar; abgeleitet aus einem Modell mit logarithmisch transformierter abhängiger Variable.

Tabelle 37: Ergebnisse der Analyse des IPAQ (anstregende Aktivität) für alle Auswertepopulationen.

IPAQ: Anstregende Aktivität [MET-Minuten/ Woche]	Kontrollgruppe		Interventionsgruppe		Differenz (%) [95-% Konfidenzintervall]	
	n	MW (SD) Median (IQR)	n	MW (SD) Median (IQR)	Unadjustiert[c]	Adjustiert[a,b,c]
Complete-Case						
Baseline	68	3185,29 (3845,26) 1860,00 (960,00 — 3780,00)	66	2572,61 (2176,82) 1800,00 (960,00 — 3360,00)		
Monat 12	68	3048,94 (3771,34) 1440,00 (720,00 — 3660,00)	66	2666,67 (3442,50) 1420,00 (776,00 — 3360,00)	-19,82 % [-68,47; 103,89] η^2 = 0,001 P = 0,641	-46,49 % [-76,20; 20,34] η^2 = 0,013 P = 0,130
Auswertepopulation B						
Baseline	83	3165,78 (4096,89) 1680,00 (680,00 — 3684,00)	80	2610,90 (2947,00) 1680,00 (812,00 — 3360,00)		
Monat 12	77	2746,60 (3639,97) 1200,00 (480,00 — 3360,00)	73	2544,66 (3308,09) 1200,00 (760,00 — 3360,00)	2,24 % [-33,16; 56,40] η^2 < 0,001 P = 0,918	-12,88 % [-50,74; 54,07] η^2 = 0,004 P = 0,629
Auswertepopulation A						
Baseline	83	3165,78 (4096,89) 1680,00 (680,00 — 3684,00)	114	2404,49 (2757,59) 1440,00 (720,00 — 3020,00)		
Monat 12	77	2746,60 (3639,97) 1200,00 (480,00 — 3360,00)	86	2294,88 (3111,31) 1200,00 (720,00 — 2460,00)	-6,88 % [-37,50; 38,73] η^2 < 0,001 P = 0,724	-9,99 % [-46,95; 52,72] η^2 = 0,004 P = 0,691
Auswertepopulation P						
Baseline	82	3202,44 (4108,38) 1680,00 (720,00 — 3722,00)	80	2610,90 (2947,00) 1680,00 (812,00 — 3360,00)		
Monat 12	77	2746,60 (3639,97) 1200,00 (480,00 — 3360,00)	73	2544,66 (3308,09) 1200,00 (760,00 — 3360,00)	2,24 % [-33,16; 56,40] η^2 < 0,001 P = 0,918	-13,47 % [-48,65; 45,79] η^2 = 0,004 P = 0,581
Auswertepopulation C						
Baseline	77	3286,75 (4171,59) 1680,00 (720,00 — 3760,00)	75	2349,23 (2128,76) 1680,00 (784,00 — 3360,00)		
Monat 12	75	2760,64 (3678,16) 1200,00 (480,00 — 3360,00)	71	2589,30 (3340,07) 1200,00 (764,00 — 3360,00)	13,03 % [-23,31; 66,58] η^2 = 0,003 P = 0,533	-6,85 % [-37,80; 39,50] η^2 < 0,001 P = 0,729

MW: Mittelwert (arithmetisch), SD: Standardabweichung, IQR: Interquartilsspannweite, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

[c] Die Differenzen stellen prozentuale Veränderungen dar; abgeleitet aus einem Modell mit logarithmisch transformierter abhängiger Variable.

Tabelle 38: Ergebnisse der Analyse der Blutparameter (Albumin) für alle Auswertepopulationen.

Albumin [g / l]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	56	45,29 (2,67)	46	45,08 (2,58)		
Monat 12	56	45,04 (2,51)	46	44,98 (2,66)	-0,06 [-1,08; 0,96] η^2 < 0,001 P = 0,904	0,04 [-0,82; 0,91] η^2 < 0,001 P = 0,923
Auswertepopulation B						
Baseline	60	45,48 (2,76)	51	45,14 (2,54)		
Monat 12	67	45,17 (2,49)	61	44,86 (2,63)	-0,31 [-1,21; 0,58] η^2 = 0,004 P = 0,493	-0,02 [-0,77; 0,73] η^2 = 0,002 P = 0,959
Auswertepopulation A						
Baseline	60	45,48 (2,76)	64	45,11 (2,52)		
Monat 12	67	45,17 (2,49)	69	44,89 (2,53)	-0,28 [-1,13; 0,57] η^2 = 0,003 P = 0,514	-0,11 [-0,97; 0,75] η^2 = 0,003 P = 0,796
Auswertepopulation P						
Baseline	59	45,43 (2,75)	50	45,13 (2,57)		
Monat 12	67	45,17 (2,49)	61	44,86 (2,63)	-0,31 [-1,21; 0,58] η^2 = 0,004 P = 0,493	-0,12 [-0,97; 0,72] η^2 = 0,004 P = 0,768
Auswertepopulation C						
Baseline	58	45,42 (2,78)	47	44,94 (2,52)		
Monat 12	66	45,18 (2,51)	58	44,88 (2,68)	-0,30 [-1,22; 0,62] η^2 = 0,003 P = 0,524	0,05 [-0,83; 0,93] η^2 < 0,001 P = 0,917

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 39: Ergebnisse der Analyse der Blutparameter (anorganisches Phosphat) für alle Auswertepopulationen.

Anorganisches Phosphat [mmol / l]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	56	0,90 (0,16)	46	0,94 (0,17)		
Monat 12	56	0,91 (0,19)	46	0,95 (0,16)	0,03 [-0,03; 0,10] η^2 = 0,010 P = 0,326	0,02 [-0,04; 0,09] η^2 = 0,006 P = 0,467
Auswertepopulation B						
Baseline	60	0,90 (0,16)	51	0,94 (0,18)		
Monat 12	67	0,91 (0,18)	61	0,96 (0,18)	0,05 [-0,01; 0,12] η^2 = 0,022 P = 0,094	0,04 [-0,02; 0,10] η^2 = 0,016 P = 0,222
Auswertepopulation A						
Baseline	60	0,90 (0,16)	65	0,93 (0,18)		
Monat 12	67	0,91 (0,18)	69	0,96 (0,17)	0,05 [-0,01; 0,11] η^2 = 0,023 P = 0,080	0,04 [-0,03; 0,11] η^2 = 0,015 P = 0,235
Auswertepopulation P						
Baseline	59	0,90 (0,16)	50	0,94 (0,18)		
Monat 12	67	0,91 (0,18)	61	0,96 (0,18)	0,05 [-0,01; 0,12] η^2 = 0,022 P = 0,094	0,04 [-0,02; 0,09] η^2 = 0,014 P = 0,208
Auswertepopulation C						
Baseline	58	0,90 (0,16)	47	0,93 (0,18)		
Monat 12	66	0,91 (0,18)	58	0,96 (0,17)	0,05 [-0,01; 0,11] η^2 = 0,020 P = 0,117	0,02 [-0,04; 0,09] η^2 = 0,005 P = 0,486

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 40: Ergebnisse der Analyse der Blutparameter (Calcium) für alle Auswertepopulationen.

Calcium [mmol / l]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	56	2,35 (0,10)	46	2,36 (0,09)		
Monat 12	56	2,32 (0,09)	46	2,36 (0,09)	0,04 [0,00; 0,08] η^2 = 0,046 P = 0,031	0,04 [0,00; 0,07] η^2 = 0,052 P = 0,026
Auswertepopulation B						
Baseline	60	2,35 (0,09)	51	2,35 (0,09)		
Monat 12	67	2,33 (0,09)	61	2,36 (0,09)	0,03 [-0,00; 0,06] η^2 = 0,026 P = 0,069	0,03 [0,00; 0,06] η^2 = 0,038 P = 0,026
Auswertepopulation A						
Baseline	60	2,35 (0,09)	64	2,36 (0,08)		
Monat 12	67	2,33 (0,09)	69	2,36 (0,09)	0,03 [-0,00; 0,06] η^2 = 0,025 P = 0,064	0,03 [-0,00; 0,06] η^2 = 0,035 P = 0,051
Auswertepopulation P						
Baseline	59	2,35 (0,09)	50	2,35 (0,09)		
Monat 12	67	2,33 (0,09)	61	2,36 (0,09)	0,03 [-0,00; 0,06] η^2 = 0,026 P = 0,069	0,03 [0,00; 0,06] η^2 = 0,039 P = 0,029
Auswertepopulation C						
Baseline	58	2,35 (0,09)	47	2,35 (0,09)		
Monat 12	66	2,33 (0,09)	58	2,36 (0,09)	0,03 [-0,00; 0,06] η^2 = 0,029 P = 0,060	0,04 [0,01; 0,08] η^2 = 0,059 P = 0,020

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 41: Ergebnisse der Analyse der Blutparameter (Creatinkinase) für alle Auswertepopulationen.

Creatinkinase [U / l]	Kontrollgruppe		Interventionsgruppe		Differenz (%) [95-% Konfidenzintervall]	
	n	MW (SD) Median (IQR)	n	MW (SD) Median (IQR)	Unadjustiert[c]	Adjustiert[a,b,c]
Complete-Case						
Baseline	56	80,11 (49,24) 69,50 (53,00 — 92,25)	46	71,74 (26,32) 69,00 (52,25 — 83,75)		
Monat 12	56	72,36 (26,77) 68,00 (57,00 — 80,25)	46	82,78 (55,65) 73,50 (57,00 — 92,00)	7,14 % [-8,23; 25,08] η^2 = 0,008 P = 0,379	9,80 % [-2,77; 23,99] η^2 = 0,024 P = 0,130
Auswertepopulation B						
Baseline	60	81,35 (48,61) 72,00 (53,75 — 92,25)	51	73,86 (27,50) 70,00 (54,00 — 89,00)		
Monat 12	67	75,76 (34,89) 69,00 (57,00 — 81,50)	61	80,97 (50,26) 73,00 (57,00 — 89,00)	3,25 % [-10,25; 18,78] η^2 = 0,002 P = 0,652	6,75 % [-7,73; 23,52] η^2 = 0,012 P = 0,371
Auswertepopulation A						
Baseline	60	81,35 (48,61) 72,00 (53,75 — 92,25)	64	72,31 (26,16) 70,00 (53,00 — 85,00)		
Monat 12	67	75,76 (34,89) 69,00 (57,00 — 81,50)	69	78,29 (48,12) 70,00 (57,00 — 85,00)	-0,01 % [-12,75; 14,58] η^2 < 0,001 P = 0,998	5,38 % [-8,82; 21,79] η^2 = 0,009 P = 0,470
Auswertepopulation P						
Baseline	59	80,05 (47,97) 71,00 (53,50 — 88,50)	50	74,16 (27,69) 72,50 (53,50 — 91,00)		
Monat 12	67	75,76 (34,89) 69,00 (57,00 — 81,50)	61	80,97 (50,26) 73,00 (57,00 — 89,00)	3,25 % [-10,25; 18,78] η^2 = 0,002 P = 0,652	7,85 % [-6,45; 24,32] η^2 = 0,015 P = 0,291
Auswertepopulation C						
Baseline	58	79,84 (48,36) 71,00 (53,25 — 84,75)	47	75,02 (27,82) 75,00 (54,00 — 93,00)		
Monat 12	66	75,73 (35,15) 68,50 (57,00 — 82,75)	58	82,09 (51,08) 73,00 (59,00 — 92,00)	4,83 % [-9,12; 20,92] η^2 = 0,003 P = 0,515	10,95 % [-1,99; 25,59] η^2 = 0,030 P = 0,099

MW: Mittelwert (arithmetisch), SD: Standardabweichung, IQR: Interquartilsspannweite, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

[c] Die Differenzen stellen prozentuale Veränderungen dar; abgeleitet aus einem Modell mit logarithmisch transformierter abhängiger Variable.

Tabelle 42: Ergebnisse der Analyse der Blutparameter (Ferritin) für alle Auswertepopulationen.

Ferritin [µg / l]	Kontrollgruppe		Interventionsgruppe		Differenz (%) [95-% Konfidenzintervall]	
	n	MW (SD) Median (IQR)	n	MW (SD) Median (IQR)	Unadjustiert[c]	Adjustiert[a,b,c]
Complete-Case						
Baseline	56	84,30 (78,61) 57,75 (37,77 — 100,78)	46	79,66 (61,07) 56,65 (42,23 — 110,58)		
Monat 12	56	76,41 (67,83) 55,60 (33,17 — 103,60)	46	79,86 (57,56) 62,15 (40,40 — 101,47)	13,96 % [-15,59; 53,87] η^2 = 0,007 P = 0,390	7,18 % [-10,67; 28,61] η^2 = 0,006 P = 0,452
Auswertepopulation B						
Baseline	60	87,59 (80,14) 58,55 (37,77 — 112,45)	51	79,89 (59,62) 57,00 (40,70 — 117,05)		
Monat 12	67	80,36 (67,16) 59,40 (34,20 — 114,05)	61	73,02 (54,15) 55,00 (34,40 — 98,00)	-3,29 % [-26,18; 26,71] η^2 < 0,001 P = 0,807	0,04 % [-15,32; 18,18] η^2 = 0,002 P = 0,996
Auswertepopulation A						
Baseline	60	87,59 (80,14) 58,55 (37,77 — 112,45)	64	80,32 (56,42) 64,40 (42,03 — 108,97)		
Monat 12	67	80,36 (67,16) 59,40 (34,20 — 114,05)	69	76,62 (57,31) 59,60 (34,70 — 101,10)	1,32 % [-21,95; 31,52] η^2 < 0,001 P = 0,921	6,43 % [-12,42; 29,34] η^2 = 0,005 P = 0,526
Auswertepopulation P						
Baseline	59	87,46 (80,82) 58,40 (37,55 — 119,40)	50	80,90 (59,78) 58,30 (42,23 — 119,33)		
Monat 12	67	80,36 (67,16) 59,40 (34,20 — 114,05)	61	73,02 (54,15) 55,00 (34,40 — 98,00)	-3,29 % [-26,18; 26,71] η^2 < 0,001 P = 0,807	0,80 % [-14,56; 18,91] η^2 = 0,002 P = 0,924
Auswertepopulation C						
Baseline	58	85,58 (80,22) 57,75 (37,33 — 103,92)	47	79,22 (58,53) 57,00 (42,25 — 117,05)		
Monat 12	66	79,37 (67,18) 59,35 (34,15 — 108,60)	58	71,29 (50,95) 54,30 (34,47 — 97,97)	-3,23 % [-26,37; 27,18] η^2 < 0,001 P = 0,812	6,49 % [-11,59; 28,28] η^2 = 0,005 P = 0,503

MW: Mittelwert (arithmetisch), SD: Standardabweichung, IQR: Interquartilsspannweite, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

[c] Die Differenzen stellen prozentuale Veränderungen dar; abgeleitet aus einem Modell mit logarithmisch transformierter abhängiger Variable.

Tabelle 43: Ergebnisse der Analyse der Blutparameter (Interleukin-8 nach Erylyse) für alle Auswertepopulationen.

Interleukin-8 nach Erylyse [pg / ml]	Kontrollgruppe		Interventionsgruppe		Differenz (%) [95-% Konfidenzintervall]	
	n	MW (SD) Median (IQR)	n	MW (SD) Median (IQR)	Unadjustiert[c]	Adjustiert[a,b,c]
Complete-Case						
Baseline	55	205,18 (321,12) 140,00 (109,10 — 190,00)	42	185,36 (119,73) 153,40 (123,90 — 192,15)		
Monat 12	55	174,05 (66,92) 158,00 (133,20 — 205,00)	42	243,62 (336,05) 186,20 (141,05 — 220,00)	17,78 % [-2,97; 42,98] η^2 = 0,029 P = 0,097	13,14 % [-4,34; 33,80] η^2 = 0,024 P = 0,147
Auswertepopulation B						
Baseline	58	203,39 (312,99) 139,90 (110,90 — 190,50)	49	178,56 (112,73) 149,80 (120,00 — 190,80)		
Monat 12	68	167,46 (63,80) 155,80 (125,50 — 198,05)	60	214,84 (284,39) 170,10 (134,80 — 203,00)	11,92 % [-4,31; 30,89] η^2 = 0,016 P = 0,157	7,43 % [-18,36; 41,37] η^2 = 0,016 P = 0,593
Auswertepopulation A						
Baseline	58	203,39 (312,99) 139,90 (110,90 — 190,50)	62	172,31 (103,23) 150,80 (119,70 — 189,80)		
Monat 12	68	167,46 (63,80) 155,80 (125,50 — 198,05)	68	213,96 (272,82) 170,10 (124,40 — 203,00)	11,12 % [-4,97; 29,94] η^2 = 0,013 P = 0,185	14,70 % [-7,02; 41,51] η^2 = 0,024 P = 0,195
Auswertepopulation P						
Baseline	57	202,40 (315,68) 139,80 (110,20 — 189,00)	48	179,20 (113,84) 150,80 (119,90 — 191,25)		
Monat 12	68	167,46 (63,80) 155,80 (125,50 — 198,05)	60	214,84 (284,39) 170,10 (134,80 — 203,00)	11,92 % [-4,31; 30,89] η^2 = 0,016 P = 0,157	5,74 % [-18,98; 38,02] η^2 = 0,013 P = 0,668
Auswertepopulation C						
Baseline	56	199,66 (317,85) 139,60 (109,65 — 186,45)	45	159,95 (67,93) 148,60 (119,60 — 189,20)		
Monat 12	67	166,95 (64,14) 154,40 (124,80 — 195,30)	57	208,94 (289,08) 163,60 (134,20 — 199,00)	9,21 % [-6,64; 27,77] η^2 = 0,010 P = 0,268	9,94 % [-7,53; 30,70] η^2 = 0,014 P = 0,279

MW: Mittelwert (arithmetisch), SD: Standardabweichung, IQR: Interquartilsspannweite, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

[c] Die Differenzen stellen prozentuale Veränderungen dar; abgeleitet aus einem Modell mit logarithmisch transformierter abhängiger Variable.

Tabelle 44: Ergebnisse der Analyse der Blutparameter (Laktatdehydrogenase) für alle Auswertepopulationen.

Laktatdehydrogenase [U / l]	Kontrollgruppe		Interventionsgruppe		Differenz (%) [95-% Konfidenzintervall]	
	n	MW (SD) Median (IQR)	n	MW (SD) Median (IQR)	Unadjustiert[c]	Adjustiert[a,b,c]
Complete-Case						
Baseline	56	204,61 (38,72) 199,00 (182,75 — 236,00)	46	195,65 (30,74) 194,00 (176,75 — 213,25)		
Monat 12	56	213,25 (36,62) 213,50 (187,50 — 231,50)	46	230,83 (94,58) 221,50 (189,25 — 244,75)	4,97 % [-3,63; 14,33] η^2 = 0,013 P = 0,263	5,97 % [-2,48; 15,16] η^2 = 0,020 P = 0,169
Auswertepopulation B						
Baseline	60	204,13 (38,04) 199,00 (182,75 — 230,00)	51	196,16 (30,92) 194,00 (173,50 — 221,50)		
Monat 12	67	215,33 (36,59) 217,00 (188,50 — 233,50)	61	229,64 (86,08) 222,00 (193,00 — 249,00)	3,78 % [-3,79; 11,94] η^2 = 0,007 P = 0,334	5,73 % [-2,05; 14,12] η^2 = 0,019 P = 0,150
Auswertepopulation A						
Baseline	60	204,13 (38,04) 199,00 (182,75 — 230,00)	64	192,23 (33,10) 193,50 (171,00 — 214,75)		
Monat 12	67	215,33 (36,59) 217,00 (188,50 — 233,50)	69	225,67 (81,94) 217,00 (186,00 — 245,00)	2,26 % [-4,88; 9,94] η^2 = 0,003 P = 0,542	4,84 % [-2,62; 12,87] η^2 = 0,015 P = 0,206
Auswertepopulation P						
Baseline	59	203,86 (38,31) 199,00 (181,50 — 232,00)	50	196,78 (30,91) 194,50 (176,75 — 221,75)		
Monat 12	67	215,33 (36,59) 217,00 (188,50 — 233,50)	61	229,64 (86,08) 222,00 (193,00 — 249,00)	3,78 % [-3,79; 11,94] η^2 = 0,007 P = 0,334	6,30 % [-1,99; 15,28] η^2 = 0,021 P = 0,137
Auswertepopulation C						
Baseline	58	203,64 (38,60) 196,00 (180,25 — 234,00)	47	198,11 (31,28) 195,00 (179,00 — 223,50)		
Monat 12	66	214,71 (36,52) 215,50 (188,25 — 232,75)	58	229,19 (86,41) 223,00 (193,50 — 248,75)	3,92 % [-3,69; 12,14] η^2 = 0,008 P = 0,319	6,48 % [-1,90; 15,58] η^2 = 0,025 P = 0,131

MW: Mittelwert (arithmetisch), SD: Standardabweichung, IQR: Interquartilsspannweite, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

[c] Die Differenzen stellen prozentuale Veränderungen dar; abgeleitet aus einem Modell mit logarithmisch transformierter abhängiger Variable.

Tabelle 45: Ergebnisse der Analyse der Blutparameter (NT-pro-BNP) für alle Auswertepopulationen.

NT-pro-BNP [ng / l]	Kontrollgruppe		Interventionsgruppe		Differenz (%) [95-% Konfidenzintervall]	
	n	MW (SD) Median (IQR)	n	MW (SD) Median (IQR)	Unadjustiert[c]	Adjustiert[a,b,c]
Complete-Case						
Baseline	56	58,52 (50,18) 45,00 (22,25 — 81,50)	46	76,80 (65,01) 69,50 (28,25 — 94,00)		
Monat 12	56	58,52 (57,24) 52,00 (27,75 — 67,25)	46	92,78 (84,19) 68,00 (34,25 — 121,00)	59,04 % [10,75; 128,38] η^2 = 0,061 P = 0,012	19,71 % [-7,18; 54,40] η^2 = 0,021 P = 0,164
Auswertepopulation B						
Baseline	60	56,17 (49,62) 42,50 (19,00 — 79,50)	51	72,12 (63,66) 63,00 (27,50 — 88,00)		
Monat 12	67	57,42 (55,38) 50,00 (22,00 — 67,50)	61	82,38 (77,82) 59,00 (28,00 — 114,00)	41,57 % [1,40; 97,66] η^2 = 0,033 P = 0,041	10,36 % [-19,80; 51,87] η^2 = 0,007 P = 0,537
Auswertepopulation A						
Baseline	60	56,17 (49,62) 42,50 (19,00 — 79,50)	64	76,52 (72,73) 63,00 (28,75 — 88,00)		
Monat 12	67	57,42 (55,38) 50,00 (22,00 — 67,50)	69	80,30 (74,23) 59,00 (29,00 — 109,00)	41,55 % [3,09; 94,35] η^2 = 0,034 P = 0,032	25,01 % [-11,73; 77,05] η^2 = 0,023 P = 0,203
Auswertepopulation P						
Baseline	59	57,00 (49,62) 43,00 (19,50 — 80,00)	50	72,70 (64,16) 64,50 (27,25 — 88,00)		
Monat 12	67	57,42 (55,38) 50,00 (22,00 — 67,50)	61	82,38 (77,82) 59,00 (28,00 — 114,00)	41,57 % [1,40; 97,66] η^2 = 0,033 P = 0,041	12,25 % [-20,10; 57,70] η^2 = 0,009 P = 0,497
Auswertepopulation C						
Baseline	58	56,62 (49,97) 42,50 (19,25 — 80,25)	47	69,34 (61,54) 63,00 (27,50 — 87,50)		
Monat 12	66	56,82 (55,58) 49,50 (21,00 — 66,25)	58	79,64 (77,08) 54,50 (28,00 — 112,75)	37,98 % [-1,78; 93,84] η^2 = 0,028 P = 0,063	22,86 % [-5,68; 60,03] η^2 = 0,026 P = 0,125

MW: Mittelwert (arithmetisch), SD: Standardabweichung, IQR: Interquartilsspannweite, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

[c] Die Differenzen stellen prozentuale Veränderungen dar; abgeleitet aus einem Modell mit logarithmisch transformierter abhängiger Variable.

Tabelle 46: Ergebnisse der Analyse des PHQ-9 für alle Auswertepopulationen.

PHQ-9 [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	90	12,82 (4,27)	84	12,92 (4,00)		
Monat 12	90	11,50 (5,02)	84	12,13 (4,13)	0,63 [-0,75; 2,01] η^2 = 0,005 P = 0,368	0,10 [-1,03; 1,23] η^2 < 0,001 P = 0,862
Auswertepopulation B						
Baseline	98	13,18 (4,66)	92	12,86 (4,11)		
Monat 12	94	11,44 (4,94)	94	11,90 (4,33)	0,47 [-0,87; 1,81] η^2 = 0,003 P = 0,491	-0,01 [-1,09; 1,07] η^2 < 0,001 P = 0,987
Auswertepopulation A						
Baseline	98	13,18 (4,66)	131	12,85 (4,41)		
Monat 12	94	11,44 (4,94)	113	12,05 (4,43)	0,62 [-0,67; 1,90] η^2 = 0,004 P = 0,345	0,18 [-0,80; 1,17] η^2 < 0,001 P = 0,712
Auswertepopulation P						
Baseline	97	13,20 (4,69)	91	12,85 (4,13)		
Monat 12	94	11,44 (4,94)	94	11,90 (4,33)	0,47 [-0,87; 1,81] η^2 = 0,003 P = 0,491	-0,06 [-1,15; 1,03] η^2 < 0,001 P = 0,913
Auswertepopulation C						
Baseline	91	12,80 (4,27)	86	12,90 (4,20)		
Monat 12	92	11,40 (4,93)	89	11,85 (4,31)	0,45 [-0,91; 1,81] η^2 = 0,002 P = 0,513	-0,14 [-1,28; 0,99] η^2 < 0,001 P = 0,806

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 47: Ergebnisse der Analyse der ESS für alle Auswertepopulationen.

ESS [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	84	10,08 (5,03)	84	10,33 (5,30)		
Monat 12	84	9,76 (5,62)	84	9,86 (5,70)	0,10 [-1,63; 1,82] η^2 < 0,001 P = 0,913	-0,41 [-1,58; 0,76] η^2 = 0,003 P = 0,488
Auswertepopulation B						
Baseline	92	10,16 (5,14)	92	10,21 (5,35)		
Monat 12	93	9,56 (5,49)	94	9,68 (5,70)	0,12 [-1,49; 1,74] η^2 < 0,001 P = 0,882	-0,48 [-1,55; 0,60] η^2 = 0,004 P = 0,383
Auswertepopulation A						
Baseline	92	10,16 (5,14)	129	9,91 (5,31)		
Monat 12	93	9,56 (5,49)	113	9,81 (5,71)	0,26 [-1,29; 1,80] η^2 < 0,001 P = 0,746	-0,23 [-1,22; 0,77] η^2 < 0,001 P = 0,652
Auswertepopulation P						
Baseline	91	10,11 (5,15)	91	10,26 (5,35)		
Monat 12	93	9,56 (5,49)	94	9,68 (5,70)	0,12 [-1,49; 1,74] η^2 < 0,001 P = 0,882	-0,54 [-1,62; 0,55] η^2 = 0,005 P = 0,330
Auswertepopulation C						
Baseline	85	9,94 (4,97)	86	10,20 (5,35)		
Monat 12	91	9,60 (5,42)	89	9,46 (5,55)	-0,14 [-1,76; 1,47] η^2 < 0,001 P = 0,861	-0,74 [-1,93; 0,45] η^2 = 0,009 P = 0,221

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 48: Ergebnisse der Analyse des EQ-5D-5L Indexwerts für alle Auswertepopulationen.

EQ-5D-5L Indexwert [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	92	0,48 (0,25)	81	0,48 (0,24)		
Monat 12	92	0,51 (0,29)	81	0,49 (0,26)	-0,02 [-0,10; 0,06] η^2 = 0,002 P = 0,611	-0,00 [-0,09; 0,08] η^2 < 0,001 P = 0,921
Auswertepopulation B						
Baseline	99	0,46 (0,25)	89	0,48 (0,25)		
Monat 12	95	0,51 (0,29)	94	0,49 (0,26)	-0,02 [-0,10; 0,05] η^2 = 0,002 P = 0,550	-0,01 [-0,09; 0,07] η^2 < 0,001 P = 0,758
Auswertepopulation A						
Baseline	99	0,46 (0,25)	129	0,48 (0,24)		
Monat 12	95	0,51 (0,29)	113	0,48 (0,30)	-0,03 [-0,11; 0,05] η^2 = 0,003 P = 0,450	-0,03 [-0,11; 0,06] η^2 = 0,002 P = 0,526
Auswertepopulation P						
Baseline	98	0,46 (0,25)	87	0,48 (0,24)		
Monat 12	95	0,51 (0,29)	94	0,49 (0,26)	-0,02 [-0,10; 0,05] η^2 = 0,002 P = 0,550	-0,01 [-0,08; 0,07] η^2 < 0,001 P = 0,879
Auswertepopulation C						
Baseline	92	0,47 (0,25)	82	0,47 (0,24)		
Monat 12	93	0,51 (0,29)	89	0,49 (0,26)	-0,01 [-0,09; 0,07] η^2 < 0,001 P = 0,742	0,02 [-0,07; 0,10] η^2 = 0,001 P = 0,684

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 49: Ergebnisse der Analyse des EQ-5D-5L Indexwerts für alle Auswertepopulationen (nur BKK).

EQ-5D-5L Indexwert [Punkte]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Baseline	22	0,49 (0,22)	19	0,54 (0,20)		
Monat 12	22	0,53 (0,31)	19	0,49 (0,25)	-0,04 [-0,22; 0,13] η^2 = 0,006 P = 0,620	-0,04 [-0,26; 0,18] η^2 = 0,005 P = 0,701
Auswertepopulation B						
Baseline	25	0,43 (0,26)	20	0,56 (0,21)		
Monat 12	22	0,53 (0,31)	20	0,48 (0,25)	-0,05 [-0,23; 0,12] η^2 = 0,010 P = 0,538	-0,03 [-0,24; 0,18] η^2 = 0,004 P = 0,747
Auswertepopulation A						
Baseline	25	0,43 (0,26)	26	0,56 (0,21)		
Monat 12	22	0,53 (0,31)	24	0,52 (0,25)	-0,01 [-0,18; 0,15] η^2 < 0,001 P = 0,884	0,00 [-0,20; 0,20] η^2 = 0,002 P = 0,975
Auswertepopulation P						
Baseline	24	0,45 (0,25)	20	0,56 (0,21)		
Monat 12	22	0,53 (0,31)	20	0,48 (0,25)	-0,05 [-0,23; 0,12] η^2 = 0,010 P = 0,538	-0,03 [-0,24; 0,17] η^2 = 0,003 P = 0,741
Auswertepopulation C						
Baseline	21	0,48 (0,22)	19	0,54 (0,20)		
Monat 12	21	0,52 (0,31)	20	0,48 (0,25)	-0,04 [-0,22; 0,14] η^2 = 0,005 P = 0,661	-0,03 [-0,25; 0,20] η^2 = 0,002 P = 0,810

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 50: Ergebnisse der Analyse der QALYs geschätzt über Baseline und Monat 12 in beiden Gruppen für alle Auswertepopulationen (nur BKK).

QALYs (Variante 1)	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Monat 12	22	0,51 (0,20)	19	0,52 (0,16)	0,00 [-0,11; 0,12] η^2 < 0,001 P = 0,934	-0,02 [-0,15; 0,11] η^2 = 0,004 P = 0,726
Auswertepopulation B						
Monat 12	22	0,51 (0,20)	19	0,52 (0,16)	0,00 [-0,11; 0,12] η^2 < 0,001 P = 0,934	0,04 [-0,10; 0,17] η^2 = 0,008 P = 0,592
Auswertepopulation A						
Monat 12	22	0,51 (0,20)	23	0,54 (0,17)	0,03 [-0,08; 0,14] η^2 = 0,008 P = 0,555	0,06 [-0,07; 0,19] η^2 = 0,020 P = 0,370
Auswertepopulation P						
Monat 12	22	0,51 (0,20)	19	0,52 (0,16)	0,00 [-0,11; 0,12] η^2 < 0,001 P = 0,934	0,01 [-0,12; 0,14] η^2 = 0,002 P = 0,848
Auswertepopulation C						
Monat 12	21	0,50 (0,19)	19	0,52 (0,16)	0,02 [-0,10; 0,13] η^2 = 0,002 P = 0,762	-0,01 [-0,14; 0,12] η^2 < 0,001 P = 0,860

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 51: Ergebnisse der Analyse der QALYs geschätzt über Baseline und Monat 12 in der Kontrollgruppe und über Baseline, Monat 3, 6, 9 und 12 in der Interventionsgruppe für alle Auswertepopulationen (nur BKK).

QALYs (Variante 2)	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Monat 12	25	0,45 (0,25)	20	0,49 (0,17)	0,04 [-0,09; 0,17] η^2 = 0,008 P = 0,561	0,01 [-0,13; 0,15] η^2 < 0,001 P = 0,858
Auswertepopulation B						
Monat 12	25	0,45 (0,25)	21	0,49 (0,17)	0,02 [-0,11; 0,16] η^2 = 0,003 P = 0,703	-0,01 [-0,14; 0,12] η^2 = 0,002 P = 0,888
Auswertepopulation A						
Monat 12	25	0,45 (0,25)	27	0,51 (0,18)	0,03 [-0,10; 0,16] η^2 = 0,005 P = 0,635	0,02 [-0,11; 0,16] η^2 = 0,004 P = 0,703
Auswertepopulation P						
Monat 12	24	0,47 (0,24)	21	0,49 (0,17)	0,01 [-0,12; 0,13] η^2 < 0,001 P = 0,922	-0,04 [-0,16; 0,09] η^2 = 0,010 P = 0,556
Auswertepopulation C						
Monat 12	21	0,50 (0,19)	20	0,50 (0,17)	-0,02 [-0,13; 0,10] η^2 = 0,002 P = 0,773	-0,05 [-0,17; 0,08] η^2 = 0,018 P = 0,442

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

Tabelle 52: Ergebnisse der Analyse der monatlichen Gesundheitskosten im Zeitraum Monat 3 bis Monat 12 für alle Auswertepopulationen (nur BKK).

Gesundheitskosten [€ / Monat]	Kontrollgruppe		Interventionsgruppe		Differenz (%) [95-% Konfidenzintervall]	
	n	MW (SD) Median (IQR)	n	MW (SD) Median (IQR)	Unadjustiert[c]	Adjustiert[a,b,c]
Complete-Case						
Monat 12	25	431,79 (727,24) 218,21 (122,09 — 360,58)	20	644,63 (940,99) 376,98 (221,53 — 593,51)	43,12 % [-21,62; 161,33] η^2 = 0,032 P = 0,236	29,56 % [-37,49; 168,53] η^2 = 0,014 P = 0,476
Auswertepopulation B						
Monat 12	25	431,79 (727,24) 218,21 (122,09 — 360,58)	21	663,62 (921,29) 399,69 (225,19 — 648,96)	50,64 % [-16,87; 172,96] η^2 = 0,042 P = 0,172	35,11 % [-33,33; 173,83] η^2 = 0,019 P = 0,393
Auswertepopulation A						
Monat 12	25	431,79 (727,24) 218,21 (122,09 — 360,58)	27	604,82 (832,11) 399,69 (216,48 — 611,99)	34,47 % [-24,91; 140,82] η^2 = 0,020 P = 0,312	17,20 % [-40,84; 132,15] η^2 = 0,005 P = 0,642
Auswertepopulation P						
Monat 12	24	446,60 (739,02) 224,72 (144,18 — 386,65)	21	663,62 (921,29) 399,69 (225,19 — 648,96)	43,48 % [-21,05; 160,75] η^2 = 0,033 P = 0,230	16,31 % [-42,23; 134,17] η^2 = 0,005 P = 0,664
Auswertepopulation C						
Monat 12	21	470,81 (788,89) 218,21 (122,09 — 464,88)	20	661,62 (945,17) 376,98 (221,53 — 593,51)	39,49 % [-27,38; 167,92] η^2 = 0,027 P = 0,309	3,50 % [-52,25; 124,36] η^2 < 0,001 P = 0,928

MW: Mittelwert (arithmetisch), SD: Standardabweichung, IQR: Interquartilsspannweite, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

[c] Die Differenzen stellen prozentuale Veränderungen dar; abgeleitet aus einem Modell mit logarithmisch transformierter abhängiger Variable.

Tabelle 53: Ergebnisse der Analyse der Krankheitstage im Zeitraum Monat 3 bis Monat 12 für alle Auswertepopulationen (nur BKK).

Krankentage [Tage]	Kontrollgruppe		Interventionsgruppe		Differenz [95-% Konfidenzintervall]	
	n	MW (SD)	n	MW (SD)	Unadjustiert	Adjustiert[a,b]
Complete-Case						
Monat 12	25	131,24 (123,91)	20	172,35 (105,13)	41,11 [-29,06; 111,28] η^2 = 0,031 P = 0,244	11,75 [-65,57; 89,07] η^2 = 0,003 P = 0,760
Auswertepopulation B						
Monat 12	25	131,24 (123,91)	21	174,62 (102,99)	43,38 [-25,15; 111,91] η^2 = 0,036 P = 0,209	10,80 [-63,83; 85,43] η^2 = 0,002 P = 0,771
Auswertepopulation A						
Monat 12	25	131,24 (123,91)	27	146,04 (116,49)	14,80 [-52,16; 81,76] η^2 = 0,004 P = 0,659	3,02 [-69,29; 75,34] η^2 < 0,001 P = 0,933
Auswertepopulation P						
Monat 12	24	125,29 (122,87)	21	174,62 (102,99)	49,33 [-19,40; 118,06] η^2 = 0,046 P = 0,155	19,97 [-56,43; 96,38] η^2 = 0,008 P = 0,599
Auswertepopulation C						
Monat 12	21	124,29 (124,44)	20	183,35 (97,37)	59,06 [-11,76; 129,89] η^2 = 0,068 P = 0,100	41,74 [-39,25; 122,73] η^2 = 0,032 P = 0,301

MW: Mittelwert (arithmetisch), SD: Standardabweichung, n: Fallzahl.

[a] Adjustiert für Baseline, Alter, Geschlecht, BMI, Erkrankungsdauer, Schwere der Erkrankung und Bildungsgrad.

[b] Die berichteten Werte stellen die gepoolten Ergebnisse aus m = 30 multiple imputierten Datensätzen dar.

CFS_CARE

Care Concept for Patients with Myalgic Encephalomyelitis/Chronic Fatigue Syndrome

Statistical Analysis Plan

Version 1.0

Apr 29, 2025

Approved by:

Carmen Scheibenbogen	_____
Principal Investigator	Place, date, signature
Ulrike Grittner	_____
Statistician	Place, date, signature
Lukas Mödl	_____
Statistician	Place, date, signature

Revision History

Version	Date	Revision
1.0	Apr 29, 2025	Initial version

Contents

1 Background **1**

2 Objective **1**

2.1 Primary Hypothesis 1

2.2 Secondary Hypotheses 1

3 Study Design **3**

3.1 Recruitment 3

3.2 Inclusion and Exclusion Criteria 4

3.3 Study Arms 5

3.3.1 Treatment Arm 5

3.3.2 Control Arm 5

3.4 Allocation 5

3.5 Sample Size Calculation 6

4 Outcomes **6**

5 Analysis Sets **11**

5.1 Full Analysis Set 11

5.2 Per Protocol Analysis Set 11

6 Data Handling **11**

6.1 Data Collection and Privacy 11

6.2 Missing Data 11

6.3 Data Anomalies 11

6.4 Data Transformations 12

7 Statistical Analyses **12**

7.1 Summary Statistics 12

7.2 Statistical Modelling 12

7.2.1 Primary Analysis 12

7.2.2 Secondary Analyses 13

7.2.3 Subgroup Analyses 13

8 Reporting of Results **13**

9 Software **14**

10 References **14**

1 Background

Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS) is a complex condition characterized by profound and persistent fatigue that is not improved by rest. This fatigue is often accompanied by a range of symptoms, including a significant worsening of symptoms following even minor physical or mental exertion (post-exertional malaise). Other common symptoms include cognitive dysfunction (e.g., memory and concentration problems), sleep disorders, widespread pain, and orthostatic intolerance.

The exact cause of ME/CFS is still unclear. It frequently follows infections, particularly the Epstein-Barr virus, and exhibits familial patterns, though no specific causative genes have been identified to this date.

Although ME/CFS is not a rare condition, it lacks widespread medical understanding. Diagnosis relies on symptom evaluation, as there is currently no definitive diagnostic test. This often leads to inaccurate diagnosis and a failure to provide the necessary individualized medical and rehabilitative care.

2 Objective

The Fatigue Center at Charité is launching CFS_CARE, a novel, interdisciplinary care model for ME/CFS patients that aims to improve physical health, enhance workplace participation, and minimize long-term care needs. This program provides comprehensive diagnostic evaluations and personalized therapy recommendations, including activity management (pacing), mind-body relaxation techniques, symptom-specific treatments like sleep and pain management, nutritional guidance, social and emotional support, and workplace reintegration assistance. The objective of CFS_CARE is to evaluate the effectiveness of this program.

2.1 Primary Hypothesis

Physical Health: At the 12-month follow-up, patients in the intervention group report better physical health on the physical functioning sub-scale of the 36-Item Short Form Survey (SF-36) compared to patients in the control group.

2.2 Secondary Hypotheses

Disability: At the 12-month follow-up, patients in the intervention group report higher scores on the Bell disability scale compared to patients in the control group.

Disability Longitudinal: Over time, patients in the intervention group report higher scores on the Bell disability scale.[a]

Fatigue: At the 12-month follow-up, patients in the intervention group report lower scores on the Chalder fatigue scale compared to patients in the control group.

Fatigue Longitudinal: Over time, patients in the intervention group report lower scores on the Chalder fatigue scale.[a]

Symptom Severity: At the 12-month follow-up, patients in the intervention group report reduced symptom severity according to the Canadian Consensus Criteria (CCC) compared to the control group.

Symptom Severity Longitudinal: Over time, patients in the intervention group report reduced symptom severity according to the CCC.[a]

Autonomic Function: At the 12-month follow-up, patients in the intervention group report lower scores on the Composite Autonomic Symptom Score-31 (COMPASS31) scale compared to the control group.

Autonomic Function Longitudinal: Over time, patients in the intervention group report lower scores on the COMPASS31 scale.[a]

Health Perception: At the 12-month follow-up, patients in the intervention group report higher scores on the SF-36 scale compared to the control group.

Health Perception Longitudinal: Over time, patients in the intervention group report higher scores on SF-36 scale.[a]

Hand Grip Strength: At the 12-month follow-up, patients in the intervention group show better hand grip strength compared to the control group.

Cardiovascular Function: At the 12-month follow-up, patients in the intervention group exhibit fewer positive Schellong tests for postural tachycardia syndrome compared to the control group.

Step Count: At the 6-month and the 12-month follow-up, patients in the intervention group show a higher average number of steps per day over time (measured with a pedometer over a ten-day period) compared to the control group.

Physical Activity: At the 12-month follow-up, patients in the intervention group report higher physical activity over time on the International Physical Activity Questionnaire (IPAQ) scale compared to the control group.

Physical Activity Longitudinal: Over time, patients in the intervention group report higher physical activity on the IPAQ scale.[a]

Blood Parameters: At the 12-month follow-up, patients in the intervention group show improved blood parameters compared to the control group.

Depressive Symptoms: At the 12-month follow-up, patients in the intervention group report fewer depressive symptoms according to the PHQ-9 scale compared to the control group.

Daytime Sleepiness: At the 12-month follow-up, patients in the intervention group report lower scores on the Epworth Sleepiness Scale (ESS) compared to the control group.

Life Quality: At the 12-month follow-up, patients in the intervention group report higher scores on the European Quality of Life 5 Dimensions 5 Level (EQ-5D-5L) scale compared to the control group.

Life Quality Longitudinal: Over time, patients in the intervention group report lower scores on the EQ-5D-5L scale.[a]

Sick Leave: At the 12-month follow-up, patients in the intervention group have fewer total sick days from three months post-enrollment onwards compared to the control group.[b]

Work Productivity: At the 12-month follow-up, patients in the intervention group have an increase in daily working hours compared to the control group.[b]

Health Care Costs: At the 12-month follow-up, patients in the intervention group have reduced health care costs during the study period compared to the control group.[b]

Adjusted Life Quality: At the 12-month follow-up, patients in the intervention group more Quality Adjusted Life Years (QALYs) compared to the control group.

[a]: Only for the intervention subgroup.

[b]: Only for the subgroup of patients insured by the Verkehrsbau Union, the Bahn, or the Siemens Betriebskrankenkasse (BKK).

3 Study Design

CFS_CARE is a prospective, single-center, non-randomized controlled two-arm intervention study evaluating effectiveness of the developed care model. The study population is visualized in Figure 1.

3.1 Recruitment

Participants are recruited from the Charité outpatient clinic. The study is promoted through the websites of the Institute for Medical Immunology and the Fatigue Center at Charité, as well as through patient organizations, including Fatigatio e.V., Lost-Voices-Stiftung, and the German Society for ME/CFS.

Eligibility requires a confirmed ME/CFS diagnosis, defined by fulfillment of the CCC from 2003, a disease duration of more than 6 months (per definition of the CCC) but less than 5 years, and a Bell Disability score between 30 and 70. Despite these criteria, only approximately half of all adults presenting to Charité with suspected ME/CFS received a confirmed diagnosis within the past two years. Therefore, to enhance diagnostic accuracy, a pre-consultation questionnaire

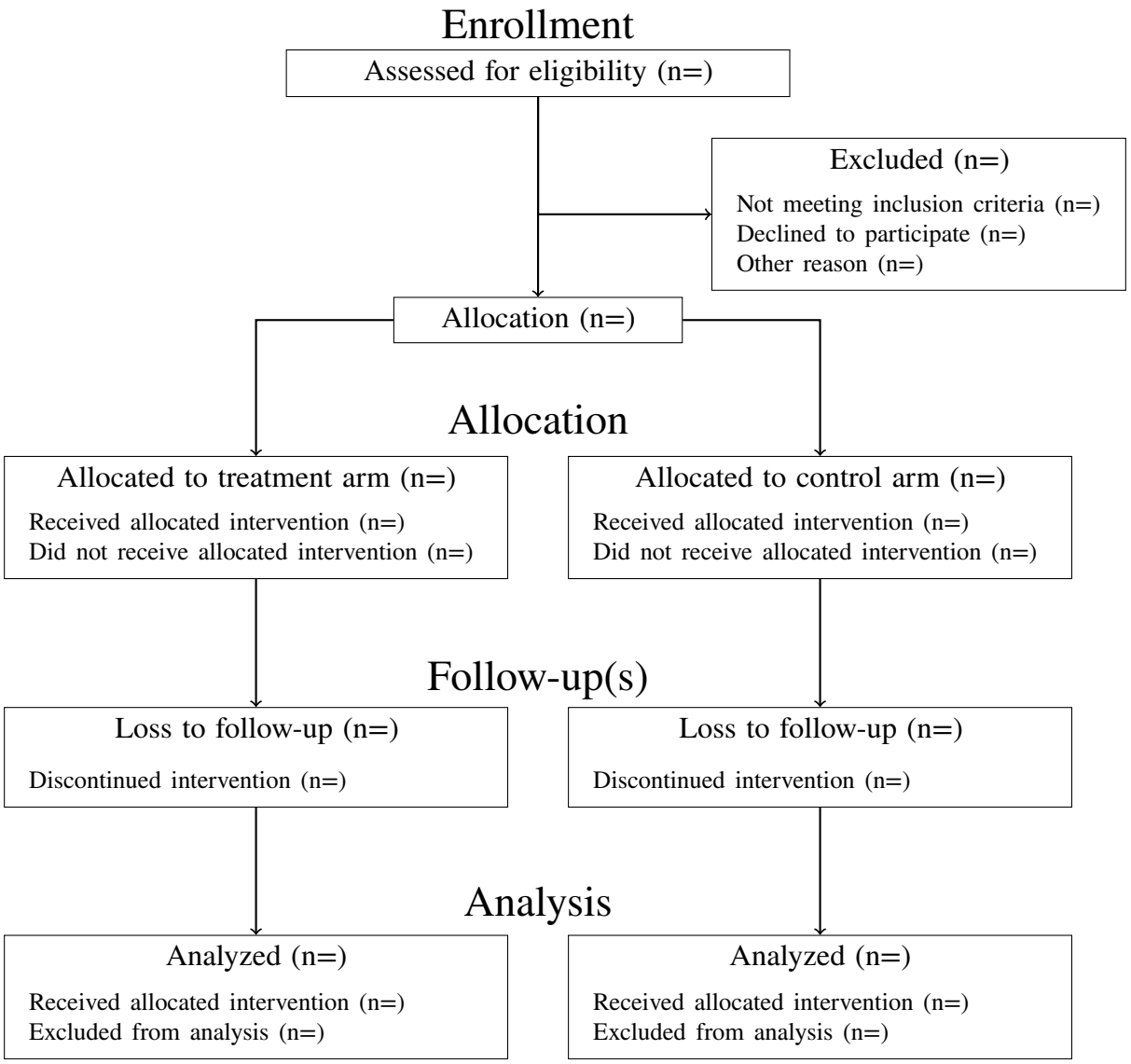


Figure 1: CONSORT flow diagram of the study population.

is administered, and telephone history and physician consultation are conducted to clarify ambiguous cases. This process is expected to confirm ME/CFS in approximately 80 % of eligible participants. A recruitment target of 300 individuals is set to achieve the planned sample size (see Section 3.5).

3.2 Inclusion and Exclusion Criteria

Inclusion criteria:

- Capacity to provide informed consent.
- Current statutory health insurance coverage.
- Age between 18 and 65 years (inclusive).

- Fulfillment of the 2003 CCC.
- Disease duration of more than 6 months but less than 5 years.
- Bell Disability Scale score between 30 and 70 (inclusive).

Exclusion criteria:

- Other illnesses associated with chronic fatigue, e.g., fatigue after cancer.

3.3 Study Arms

3.3.1 Treatment Arm

Patients undergo a comprehensive evaluation at Charité, including sleep medicine examinations, sports medicine assessments of physical performance, and cardiological evaluations of endothelial function and heart rate variability. Patients also receive counseling on activity limits and stress management and are provided with an activity tracker, which monitors various physiological data. Subsequently, patients are admitted to a specialized rehabilitation clinic in Kreischa, Germany. The inpatient rehabilitation program focuses on several key areas: teaching activity management (pacing) and relaxation techniques (mind-body therapy); providing symptomatic treatment for sleep and pain, along with physical therapy, lymphatic drainage, circulatory support, and mobility aids; offering nutritional counseling; addressing secondary social and emotional challenges; and educating family members. A crucial part of the rehabilitation process involves analyzing work capacity and the workplace, and planning reintegration measures to potentially restore the patient's (partial) ability to work. If occupational disability is determined, social workers provide counseling and application support during and after the rehabilitation stay. Following the rehabilitation stay, follow-up care is provided by the Charité. Data assessments occur every 3 months. A final assessment conducted by Charité concludes the 12-month study period.

3.3.2 Control Arm

The care provided in the control arm of the project reflects the current standard of care for ME/CFS patients in Germany. Following diagnosis, patients receive continued care from their referring physician. Both physicians and patients will receive a detailed report with therapy recommendations offered by the Charité, as well as offers for further education and telephone consultation. A final assessment conducted by Charité concludes the 12-month study period.

3.4 Allocation

Patients are not be randomly assigned to intervention and control arm. Because the German Federal Pension Insurance (DRV) funds the rehabilitation, eligibility for the inpatient rehabilita-

tion intervention is restricted to patients who meet all inclusion criteria and are insured by the DRV. Therefore, randomization is not feasible. Patients who met the inclusion criteria but are not insured by the DRV are eligible for inclusion in the control group.

3.5 Sample Size Calculation

Limited published data exist evaluating structured specialist care for patients with ME/CFS. A study by Brown, Khorana, and Jason [1] used the Energy Envelope Theory, which posits that patients who remain active within their individual activity limits (pacing) achieve greater functional improvements than patients who exceed their individual limits. Their study, which included 17 patients who did not exceed their activity limits at baseline and 27 patients who did, observed differences in physical health after 12 months. The subgroup with lower activity at baseline showed a Cohen's d of 0.49 (mean improvement: 16.7 vs. 3.6 points, pooled standard deviation: 26.6), while the subgroup with higher activity at baseline showed a Cohen's d of 0.68 (mean improvement: 19.3 vs. 1.2 points, pooled standard deviation: 26.6).

CFS_CARE includes patients with various underlying conditions. Therefore, somewhat smaller effects are anticipated. With 106 patients per group, a t-test for independent samples with 80 % power can detect a difference with an effect size of 0.387 or greater at a two-sided significance level of 5 %. A dropout rate of approximately 10 % is assumed, therefore 120 patients per group are enrolled. The risk of dropouts in the control group is considered relatively low, as these patients continue to receive care from their referring physicians and undergo a comprehensive examination at Charité after 12 months.

4 Outcomes

Table 1 summarizes all measured outcomes.

Table 1: Overview over all outcomes assessed in the treatment (T) and control (C) arms.

Variable	Baseline	3-month Follow-up	6-month Follow-up	9-month Follow-up	12-month Follow-up
SF-36	T + C	T	T	T	T + C
Physical functioning[‡]					
Physical role functioning					
Bodily pain					
General health perception					
Vitality					
Social role functioning					
Emotional role functioning					
Mental health					
Bell disability scale	T + C	T	T	T	T + C
Chalder fatigue scale	T + C	T	T	T	T + C
Physical					
Mental					
Total					
Symptom severity	T + C	T	T	T	T + C
COMPASS31 scale	T + C	T	T	T	T + C
Orthostatic intolerance					
Vasomotor					
Secretomotor					
Gastrointestinal					

Table 1: Overview over all outcomes assessed in the treatment (T) and control (C) arms.

Variable	Baseline	3-month Follow-up	6-month Follow-up	9-month Follow-up	12-month Follow-up
Bladder					
Pupillomotor changes					
Hand grip strength	T + C	T	T	T	T + C
Average Session 1					
Average Session 2					
Maximal Session 1					
Maximal Session 2					
Fatigue ratio Session 1					
Fatigue ratio Session 2					
Recovery ratio					
Schellong test	T + C				T + C
Step count	T + C		T + C		T + C
IPAQ	T + C	T	T	T	T + C
Sitting					
Walking					
Moderate activities					
Vigorous activities					
Blood parameters	T + C				T + C
Creatine kinase					
Lactate dehydrogenase					

Table 1: Overview over all outcomes assessed in the treatment (T) and control (C) arms.

Variable	Baseline	3-month Follow-up	6-month Follow-up	9-month Follow-up	12-month Follow-up
N-terminal fragment B-type natriuretic peptide					
Albumin					
Ferritin					
Transferrin saturation					
Calcium					
Phosphate					
Interleukin-8					
PHQ-9 scale	T + C				T + C
ESS	T + C				T + C
EQ-5D-5L	T + C	T	T	T	T + C
Sick leave					T + C
Working hours					T + C
Health care costs[1]					T + C
QALY[2]					T + C

‡: Primary outcome.

[1]: All inpatient and outpatient BKK-related claims data from the study period are included, covering hospital, pharmaceutical, therapy, medical aid, transportation, home nursing, domestic assistance, ambulatory physician, and SGB XI social care services. Due to reporting delays, outpatient costs are directly considered only up to month 18. Total costs are standardized to monthly averages per patient. These averages are then aggregated per patient to calculate a total mean monthly cost.

[2]: QALYs are calculated from the EQ-5D-5L, assuming linear changes between baseline and follow-up. Cost-effectiveness is determined by comparing quality-of-life gains and costs. the incremental cost-effectiveness ratio is calculated by estimating the mean additional cost per extra QALY. A threshold of 50,000 € per extra QALY is often used in studies as a cost-effectiveness limit. If the ratio is below this threshold, the intervention is considered cost-effective.

5 Analysis Sets

5.1 Full Analysis Set

The full analysis set comprises all patients assigned to a treatment arm, regardless of adherence. Section 7 details its application.

5.2 Per Protocol Analysis Set

The per protocol analysis set comprises all patients who completed the study as assigned, excluding those lost to follow-up, withdrew, or changed treatment arms. Section 7 details its application.

6 Data Handling

6.1 Data Collection and Privacy

All data are pseudonymized and collected via an electronic Case Report Form within REDCap. All entered data is monitored and undergoes quality checks. Health care cost data (provided by Verkehrsbau Union, Bahn, and Siemens BKK) are linked to the data base using patient pseudonyms.

6.2 Missing Data

Reasons for missing values are reported and documented. If feasible, missing data in the full analysis set (defined in Section 5.1) are imputed using Multiple Imputation by Chained Equations. The imputation model uses predictive mean matching based on all outcomes defined in Section 1 as well as the covariates considered in Section 7.2. A total of 30 complete datasets are generated, with 50 iterations per imputation. Imputation accuracy and model convergence is thoroughly assessed using descriptive measures, including kernel density estimates and trace plots.

6.3 Data Anomalies

Comprehensive data quality checks, including descriptive statistics (see Section 7.1), are conducted to identify potential data anomalies. When anomalies are detected (e.g., outliers), an investigation is performed to determine if they result from data entry errors or other similar issues. If a data entry error is confirmed, the data are corrected or removed. In cases where the source of the anomaly cannot be definitively determined, the original values are retained.

6.4 Data Transformations

In cases of skewed distributions, transformations such as the Box-Cox transform may be applied to improve symmetry. The rationale for, and the application of, any transformations are carefully discussed and justified.

7 Statistical Analyses

7.1 Summary Statistics

Summary statistics for ordinal, interval, and ratio outcomes include the number of missing values (if any), number of non-missing values, mean $\pm$ standard deviation, and Tukey's five-point summary (minimum, 25th percentile, median, 75th percentile, maximum). In addition to tabular presentation, the distribution of each outcome is visualized by treatment arm and follow-up using a combination of box and violin plots, where appropriate. Summary statistics for nominal outcomes include the number of missing values (if any), number of non-missing values, and absolute and relative frequencies.

Summary statistics are calculated for both the full analysis set and the per-protocol analysis set (defined in Section 5.1 and Section 5.2, respectively).

7.2 Statistical Modelling

To evaluate primary and secondary outcomes, separate Analysis of Covariance (ANCOVA) models are employed to estimate treatment effects. The model equation is:

$$Y_t = \theta + \beta \times T + \gamma \times Y_0 + \boldsymbol{\delta} \bullet \boldsymbol{X} + \varepsilon,$$

where Y_t denotes the outcome value at month t post enrollment, T is a binary indicator for the intervention group, Y_0 denotes the outcome value at baseline, and $\boldsymbol{X}$ denotes a set of confounding variables (sex, age at baseline, disease duration in months at baseline, disease severity on the Bell disability scale at baseline or maximal hand grip strength at baseline, disease trigger, vocational training status at baseline, and employment status at baseline). The remainder term ε captures residual deviations in the postulated relationship.

Model fits are primarily assessed visually using quantile-quantile and fitted versus residuals plots. If there is evidence of non-linearity, heteroscedasticity, or strong deviations from normality, data transformations are considered (see Section 6.4).

7.2.1 Primary Analysis

The primary analysis utilized the ANCOVA model, as specified, for $t = 12$.

A two-sided significance level of $\alpha = 5$ % is used.

The primary analysis uses the full analysis set (defined in Section 5.1), followed by a sensitivity analysis using the per-protocol analysis set (defined in Section 5.2).

7.2.2 Secondary Analyses

The secondary analyses also employ the ANCOVA model, as specified above. For longitudinal analyses, the ANCOVA model is modified to account for repeated measurements within the intervention group. Specifically, a set $\boldsymbol{C}$ of time index indicators for the 3-, 6-, 9-, and 12-month follow-ups, and an individual-specific random effect τ are included in the model. Since longitudinal analyses, which include time points $t = 0, 3, 6, 9, 12$, are limited to the intervention group, both coefficients β and γ are constrained to zero to ensure identifiability.

All secondary analyses are exploratory.

All secondary analyses use the full analysis set (defined in Section 5.1).

7.2.3 Subgroup Analyses

Subgroup analyses extend the primary and secondary analyses by modifying the ANCOVA model to examine subgroup-specific effects. Specifically, a subgroup indicator G and the interaction $T \times G$ are included. If feasible, the following subgroups are investigated:

- Women vs. men: With ME/CFS occurring twice as often in women, their response to treatment effectiveness may vary compared to men.
- Fatigue onset due to COVID-19 infection vs. other causes: Because different causes of fatigue onset likely involve distinct pathophysiological mechanisms, treatment effectiveness may vary.
- With breathing-related sleep disorders vs. insomnia: Insomnia is a common comorbidity in ME/CFS patients, and it may affect treatment effectiveness.
- With vs. without postural tachycardia syndrome: Postural tachycardia syndrome is a common comorbidity in ME/CFS patients, and it may affect treatment effectiveness.

All subgroup analyses are exploratory.

All subgroup analyses use the full analysis set (defined in Section 5.1).

8 Reporting of Results

Summary statistics are reported by treatment arm and follow-up. Estimated adjusted (marginal) mean differences are presented with p-values, 95 % confidence intervals, and explained variance.

Other metrics such as R^2, or the random effects variance may also be reported.

P-values and confidence intervals are derived from corresponding t-distributions. Explained variance is estimated using η^2.

All reported non-integer values are rounded to three decimal places.

9 Software

All statistical analyses are conducted using R (version ≥4.4+) with the `tidyverse` (version ≥1.3+), `lme4` (version ≥1.1+), and `mice` (version ≥3.17+) packages.

10 References

[1] Maureen Brown, Neetu Khorana, and Leonard A. Jason. "The role of changes in activity as a function of perceived available and expended energy in nonpharmacological treatment outcomes for ME/CFS." *Journal of Clinical Psychology* 67.3 (2011), pp. 253–260.

CFS_CARE

Care Concept for Patients with Myalgic Encephalomyelitis/Chronic Fatigue Syndrome

Statistical Analysis Plan

Version 2.0

Dec 31, 2025

Revision History

Version	Date	Revision
1.0	Apr 29, 2025	Initial version
2.0	Dec 20, 2025	Revised version Changes: • Clarified symptom severity hypotheses. • Removed productivity hypothesis. • Clarified some secondary hypotheses (health perception over time, symptoms, step count). • Clarified outcome definitions and assessment time points (Table 1). • Added baseline characteristics to Section 4. • Added precise variable definitions for health care costs, sick leave, Schellong test, and step count. • Added Section 6: Covariates. • Reworked Figure 1: CONSORT diagram; and moved it to Section 7. • Reworked former Section 5: Analysis Sets (now Section 6: Study Population). • Revised the description of the imputation model. • Added a section for the cost effectiveness analysis. • Revised Statistical Analyses section. • Fixed language as well as several grammar and spelling errors.

Major changes from version 1.0 to version 2.0 are marked in red.

Contents

1 Background **1**

2 Objective **1**

2.1 Primary Hypothesis 1

2.2 Secondary Hypotheses 1

3 Study Design **3**

3.1 Recruitment 3

3.2 Inclusion and Exclusion Criteria 4

3.3 Study Arms 4

3.3.1 Treatment Arm 4

3.3.2 Control Arm 5

3.4 Allocation 5

3.5 Sample Size Calculation 5

4 Outcomes and Baseline Characteristica **5**

5 Covariates **10**

6 Study Population **10**

6.1 Sub-Population A 10

6.2 Sub-Population B 10

6.3 Sub-Population P 10

6.4 Sub-Population C 10

7 Data Handling **12**

7.1 Data Collection and Privacy 12

7.2 Missing Data 12

7.3 Data Anomalies 12

7.4 Data Transformations 12

8 Statistical Analyses **13**

8.1 Summary Statistics 13

8.2 Statistical Modelling 13

8.2.1 Primary Analysis 13

8.2.2 Secondary Analyses 14

8.2.3 Subgroup Analyses 14

8.2.4 Cost Effectiveness Analysis 15

9 Reporting of Results **15**

10 Software **15**

11 References **16**

1 Background

Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (ME/CFS) is a complex condition characterized by profound and persistent fatigue that is not improved by rest. This fatigue is often accompanied by a range of symptoms, including a significant worsening of symptoms following even minor physical or mental exertion (post-exertional malaise). Other common symptoms include cognitive dysfunction (e.g., memory and concentration problems), sleep disorders, widespread pain, and orthostatic intolerance.

The exact cause of ME/CFS is still unclear. It frequently follows infections, particularly the Epstein-Barr virus, and exhibits familial patterns, though no specific causative genes have been identified to this date.

Although ME/CFS is not a rare condition, it lacks widespread medical understanding. Diagnosis relies on symptom evaluation, as there is currently no definitive diagnostic test. This often leads to inaccurate diagnosis and a failure to provide the necessary individualized medical and rehabilitative care.

2 Objective

The Fatigue Center at Charité is launching CFS_CARE, a novel, interdisciplinary care model for ME/CFS patients that aims to improve physical health, enhance workplace participation, and minimize long-term care needs. This program provides comprehensive diagnostic evaluations and personalized therapy recommendations, including activity management (pacing), mind-body relaxation techniques, symptom-specific treatments like sleep and pain management, nutritional guidance, social and emotional support, and workplace reintegration assistance. The objective of CFS_CARE is to evaluate the effectiveness of this program compared to the standard of care.

2.1 Primary Hypothesis

Physical functioning: After 12 months, patients in the intervention group report better physical health on the physical functioning scale of the 36-Item Short Form Survey (SF-36) compared to patients in the control group.

2.2 Secondary Hypotheses

Physical functioning over time: Over time, patients in the intervention group report higher scores on the physical functioning scale of the SF-36.[a]

Disability: After 12 months, patients in the intervention group report higher scores on the Bell disability scale compared to patients in the control group.

Disability over time: Over time, patients in the intervention group report higher scores on the Bell disability scale.[a]

Fatigue: After 12 months, patients in the intervention group report lower scores on the Chalder fatigue scale compared to patients in the control group.

Fatigue over time: Over time, patients in the intervention group report lower scores on the Chalder fatigue scale.[a]

Symptoms: After 12 months, patients in the intervention group report less severe symptoms according to the Canadian Consensus Criteria (CCC) compared to the control group.

Symptoms over time: Over time, patients in the intervention group report less severe symptoms according to the CCC.[a]

Autonomic function: After 12 months, patients in the intervention group report lower scores on the Composite Autonomic Symptom Score 31 (COMPASS 31) scale compared to the control group.

Autonomic function over time: Over time, patients in the intervention group report lower scores on the COMPASS 31 scale.[a]

Health perception: After 12 months, patients in the intervention group report higher scores on the scales of the SF-36 compared to the control group.

~~**Health Perception over time:** Over time, patients in the intervention group report higher scores on the scales of the SF-36.[a]~~

Hand grip strength: After 12 months, patients in the intervention group show better hand grip strength compared to the control group.

Cardiovascular function: After 12 months, patients in the intervention group exhibit fewer positive Schellong tests compared to the control group.

~~**Step count:** At the 6-month and the 12-month follow-up, patients in the intervention group show a higher average number of steps per day over time (measured with a pedometer over a ten-day period) compared to the control group.~~

Step count: At the 12-month follow-up, patients in the intervention group show a higher average number of steps per day (measured with an acelerometer) compared to the control group.

Physical activity: After 12 months, patients in the intervention group report higher physical activity over time on the International Physical Activity Questionnaire (IPAQ) compared to the control group.

Physical activity over time: Over time, patients in the intervention group report higher physical activity on the IPAQ.[a]

Blood Parameters: After 12 months, patients in the intervention group show improved blood parameters compared to the control group.

Depressive symptoms: After 12 months, patients in the intervention group report fewer depressive symptoms according to the Patient Health Questionnaire 9 (PHQ-9) compared to the control group.

Daytime sleepiness: After 12 months, patients in the intervention group report lower scores on the Epworth Sleepiness Scale (ESS) compared to the control group.

Quality of life: After 12 months, patients in the intervention group report higher European Quality of Life 5 Dimensions 5 Level (EQ-5D-5L) scores compared to the control group.

Quality of life over time: Over time, patients in the intervention group report lower EQ-5D-5L scores.[a]

Sick leave: After 12 months, patients in the intervention group have fewer total sick days from three months post-enrollment onwards compared to the control group.[b]

~~**Work productivity:** After 12 months, patients in the intervention group have an increase in daily working hours compared to the control group.[b]~~

Health care costs: After 12 months, patients in the intervention group have reduced health care costs during the study period compared to the control group.[b]

Adjusted quality of life: After 12 months, patients in the intervention group more Quality Adjusted Life Years (QALYs) compared to the control group.[b]

[a]: Only for the intervention group.

[b]: Only for the subgroup of patients insured by the Verkehrsbau Union, the Bahn, or the Siemens Betriebskrankenkasse (BKK).

3 Study Design

CFS_CARE is a prospective, single-center, non-randomized controlled two-arm intervention study evaluating effectiveness of the new care model.

3.1 Recruitment

Participants are recruited from the Charité outpatient clinic. The study is promoted through the websites of the Institute for Medical Immunology and the Fatigue Center at Charité, as well as through patient organizations, including Fatigatio e.V., Lost-Voices-Stiftung, and the German Society for ME/CFS.

Eligibility requires a confirmed ME/CFS diagnosis, defined by fulfillment of the CCC from 2003, a disease duration of more than 6 months (per definition of the CCC) but less than 5 years, and a Bell Disability score between 30 and 70. Despite these criteria, only approximately half of all adults presenting to Charité with suspected ME/CFS received a confirmed diagnosis within

the past two years. Therefore, to enhance diagnostic accuracy, a pre-consultation questionnaire is administered, and telephone history and physician consultation are conducted to clarify ambiguous cases. This process is expected to confirm ME/CFS in approximately 80 % of eligible participants. A recruitment target of 300 individuals is set to achieve the planned sample size (see Section 3.5).

3.2 Inclusion and Exclusion Criteria

Inclusion criteria are:

- Capacity to provide informed consent.
- Current statutory health insurance coverage.
- Age between 18 and 65 years (inclusive).
- Fulfillment of the 2003 CCC.
- Disease duration of more than 6 months but less than 5 years.
- Bell Disability Scale score between 30 and 70 (inclusive).

Exclusion criteria are:

- Other illnesses associated with chronic fatigue, e.g., fatigue after cancer.

3.3 Study Arms

3.3.1 Treatment Arm

Patients undergo a comprehensive evaluation at Charité, including sleep medicine examinations, sports medicine assessments of physical performance, and cardiological evaluations of endothelial function and heart rate variability. Patients also receive counseling on activity limits and stress management and are provided with an activity tracker, which monitors various physiological data. Subsequently, patients are admitted to a specialized rehabilitation clinic in Kreischa, Germany. The inpatient rehabilitation program focuses on several key areas: teaching activity management (pacing) and relaxation techniques (mind-body therapy); providing symptomatic treatment for sleep and pain, along with physical therapy, lymphatic drainage, circulatory support, and mobility aids; offering nutritional counseling; addressing secondary social and emotional challenges; and educating family members. A crucial part of the rehabilitation process involves analyzing work capacity and the workplace, and planning reintegration measures to potentially restore the patient's (partial) ability to work. If occupational disability is determined, social workers provide counseling and application support during and after the rehabilitation stay. Following the rehabilitation stay, follow-up care is provided by the Charité. Data assessments occur every 3 months. A final assessment conducted by Charité concludes the 12-month study period.

3.3.2 Control Arm

The care provided in the control arm of the project reflects the current standard of care for ME/CFS patients in Germany. Following diagnosis, patients receive continued care from their referring physician. Both physicians and patients will receive a detailed report with therapy recommendations offered by the Charité, as well as offers for further education and telephone consultation. A final assessment conducted by Charité concludes the 12-month study period.

3.4 Allocation

Patients are not be randomly assigned to intervention and control arm. Because the German Federal Pension Insurance (DRV) funds the rehabilitation, eligibility for the inpatient rehabilitation intervention is restricted to patients who meet all inclusion criteria and are insured by the DRV. Therefore, randomization is not feasible. Patients who met the inclusion criteria but are not insured by the DRV are eligible for inclusion in the control group.

3.5 Sample Size Calculation

Limited published data exist evaluating structured specialist care for patients with ME/CFS. A study by Brown, Khorana, and Jason [4] used the Energy Envelope Theory, which posits that patients who remain active within their individual activity limits (pacing) achieve greater functional improvements than patients who exceed their individual limits. Their study, which included 17 patients who did not exceed their activity limits at baseline and 27 patients who did, observed differences in physical health after 12 months. The subgroup with lower activity at baseline showed a Cohen's d of 0.49 (mean improvement: 16.7 vs. 3.6 points, pooled standard deviation: 26.6), while the subgroup with higher activity at baseline showed a Cohen's d of 0.68 (mean improvement: 19.3 vs. 1.2 points, pooled standard deviation: 26.6).

CFS_CARE includes patients with various underlying conditions. Therefore, somewhat smaller effects are anticipated. With 106 patients per group, a t-test for independent samples with 80 % power can detect a difference with an effect size of 0.387 or greater at a two-sided significance level of 5 %. A dropout rate of approximately 10 % is assumed, therefore 120 patients per group are enrolled. The risk of dropouts in the control group is considered relatively low, as these patients continue to receive care from their referring physicians and undergo a comprehensive examination at Charité after 12 months.

4 Outcomes and Baseline Characteristica

Table 1 presents all considered outcomes and their corresponding assessment time points. Table 2 summarizes key characteristics assessed at baseline.

Table 1: Overview over all outcomes assessed in the treatment (T) and control (C) arms.

Variable	Unit	Baseline	Month 3	Month 6	Month 9	Month 12
SF-36						
Physical functioning[‡]		T + C	T + C	T + C	T	T + C
Physical role functioning		T + C				T + C
Bodily pain		T + C				T + C
General health perception		T + C				T + C
Vitality		T + C				T + C
Social role functioning		T + C				T + C
Emotional role functioning		T + C				T + C
Mental health		T + C				T + C
Bell disability scale		T + C	T	T	T	T + C
Chalder fatigue scale						
Physical		T + C	T	T	T	T + C
Mental		T + C	T	T	T	T + C
Total		T + C	T	T	T	T + C
Symptoms						
Headache		T + C	T	T	T	T + C
Muscle		T + C	T	T	T	T + C
COMPASS 31 scale						
Orthostatic intolerance		T + C	T	T	T	T + C
Vasomotor		T + C	T	T	T	T + C

Table 1: Overview over all outcomes assessed in the treatment (T) and control (C) arms.

Variable	Unit	Baseline	Month 3	Month 6	Month 9	Month 12
Secretomotor		T + C	T	T	T	T + C
Gastrointestinal		T + C	T	T	T	T + C
Bladder		T + C	T	T	T	T + C
Pupillomotor		T + C	T	T	T	T + C
Hand grip strength						
Average Session 1	kg	T + C				T + C
Average Session 2	kg	T + C				T + C
Maximal Session 1	kg	T + C				T + C
Maximal Session 2	kg	T + C				T + C
Fatigue ratio Session 1		T + C				T + C
Fatigue ratio Session 2		T + C				T + C
Recovery ratio		T + C				T + C
Positive Schellong test[1]						
After 0 minutes standing		T + C	T	T	T	T + C
After 2 minutes standing		T + C	T	T	T	T + C
After 5 minutes standing		T + C	T	T	T	T + C
After 10 minutes standing		T + C	T	T	T	T + C
Step count[1]	steps/day	T + C		T + C		T + C
IPAQ						
Walking	MET-min	T + C	T	T	T	T + C

Table 1: Overview over all outcomes assessed in the treatment (T) and control (C) arms.

Variable	Unit	Baseline	Month 3	Month 6	Month 9	Month 12
Moderate activities	MET-min	T + C	T	T	T	T + C
Vigorous activities	MET-min	T + C	T	T	T	T + C
Blood parameters						
Creatine kinase	U/l	T + C				T + C
Lactate dehydrogenase	U/l	T + C				T + C
N-terminal fragment B-type natriuretic peptide	ng/l	T + C				T + C
Albumin	g/l	T + C				T + C
Ferritin	µg/l	T + C				T + C
Calcium	mmol/l	T + C				T + C
Inorganic phosphate	mmol/l	T + C				T + C
Interleukin-8 after erylysis	pg/ml	T + C				T + C
PHQ-9 scale		T + C				T + C
ESS		T + C				T + C
EQ-5D-5L		T + C	T	T	T	T + C
Sick leave[3]	days					T + C
Health care costs[4]	Euro					T + C
QALY[5]					T + C	

‡: Primary outcome.

[1]: A positive Schellong test is defined by the presence of at least one of the following criteria: a decrease in systolic blood pressure of $\geq$ 20 mmHg, a decrease in diastolic blood pressure of $\geq$ 10 mmHg, a heart rate increase of $\geq$ 30 bpm, or an absolute standing heart rate of $\geq$ 120 bpm.

[2]: Average daily steps were calculated by dividing the total step count by the wear duration (days), as estimated by the Troiano algorithm [3].

[3]: Sick days are calculated pro rata for the observation period of 90 to 365 days after the inclusion date.

[4]: All inpatient and outpatient BKK-related claims data from the study period are included, covering hospital, pharmaceutical, therapy, medical aid, transportation, home nursing, domestic assistance, ambulatory physician, and SGB XI social care services. Specifically, eligible costs must be incurred between 90 and 365 days after the inclusion date. Due to reporting delays, outpatient costs are directly considered only up to month 18. Total costs are standardized to monthly averages per patient. These averages are then aggregated per patient to calculate a total mean monthly cost.

[5]: QALYs are calculated from the EQ-5D-5L, assuming linear changes between baseline and follow-up (trapezodial rule).

Table 2 summarizes key characteristics assessed at baseline.

5 Covariates

Among the baseline characteristics presented in Table 2, the following variables were identified as potential confounding factors for subsequent analyses:

- Age (in years at) baseline.
- Body mass index (in kg/m^2) at baseline.
- Sex (male, female).
- Disease severity (on the Bell disability scale) at baseline.
- Disease duration (in months) at baseline.
- Education (according to the International Standard Classification of Education; [5]) at baseline.

6 Study Population

Figure 1 illustrates the schematic flow of the study population from recruitment to final analysis.

6.1 Sub-Population A

Sub-population A is based on the intention-to-treat (ITT) principle and comprises all patients assigned to intervention or control group, regardless of adherence.

6.2 Sub-Population B

Sub-population C comprises all patients from the intervention group who completed rehab and all patients initially allocated to the control group.

6.3 Sub-Population P

Sub-population P comprises all patients who completed completed the study as assigned.

6.4 Sub-Population C

Sub-population C comprises all patients who completed the study as assigned, excluding those lost to follow-up, withdrew, changed treatment arms, or have missing data in the primary end outcome.

Table 2: Baseline characteristics.

Category	Variable
Patient	- Age (in years) - Body mass index (in kg/m^2) - Sex (male, female)
Disease	- Disease severity (on the Bell disability scale) - Disease duration (in months) - Disease trigger (COVID-19, Epstein–Barr virus infection, other respiratory infections, other infections, no documented infection).
Health-related lifestyle factors	- Alcohol consumption (frequency per week, no alcohol consumption) - Smoker (active, former, no)
Co-morbidities	- Allergies - Cardiovascular diseases - Gastrointestinal diseases - Nephrological and urogenital disorders - Endocrine and metabolic disorders - Musculoskeletal disorders - Neurological disorders - Pulmonary diseases - Dermatological diseases - Psychiatric disorders - Oncological diseases - Other diseases.
Family and children	- Family status (married, separated, single) - Children (yes, no)
Education, training, and employment	- Education (lower secondary school, intermediate secondary school, higher education entrance qualification) - Training (no formal qualification; currently in school or vocational training, vocational training, school-based vocational education or technical college, university of applied sciences or university, other) - Employment (employed, unemployed).

7 Data Handling

7.1 Data Collection and Privacy

All data are pseudonymized and collected via an electronic Case Report Form within REDCap. All entered data is monitored and undergoes quality checks. Health care cost data (provided by Verkehrsbau Union, Bahn, and Siemens BKK) are linked to the data base using patient pseudonyms.

7.2 Missing Data

Reasons for missing values are reported and documented. ~~If feasible, missing data in the full analysis set (defined in Section **??**) are imputed using Multiple Imputation by Chained Equations.~~ Where feasible and provided that missingness does not exceed 50 %, missing values are addressed using Multiple Imputation by Chained Equations (MICE), assuming that data are Missing at Random (MAR) or Missing Completely at Random (MCAR). ~~The imputation model uses predictive mean matching based on all outcomes defined in Section 1 as well as the covariates considered in Section 8.2~~. All outcomes outlined in Section 4, except for BKK-related data (which are considered complete by design), are imputed for all populations using predictive mean matching. For nominal outcomes, multinomial logistic regression may be employed. The imputation models incorporate all covariates specified in Section 5, along with longitudinal measurements of the respective outcome from all available time points. For the sub-population C only missing data in the covariates are imputed. A total of 30 complete datasets are generated, with 50 iterations per imputation. Estimates from the imputed data sets are pooled using Rubin's rules [2]. Imputation accuracy and model convergence is thoroughly assessed using descriptive measures, including kernel density estimates and trace plots.

7.3 Data Anomalies

Comprehensive data quality checks, including descriptive statistics (see Section 8.1), are conducted to identify potential data anomalies. When anomalies are detected (e.g., outliers), an investigation is performed to determine if they result from data entry errors or other similar issues. If a data entry error is confirmed, the data are corrected or removed. In cases where the source of the anomaly cannot be definitively determined, the original values are retained.

7.4 Data Transformations

In cases of skewed distributions, transformations such as the Box-Cox transform may be applied to improve symmetry. The rationale for, and the application of, any transformations are carefully discussed and justified.

8 Statistical Analyses

8.1 Summary Statistics

Summary statistics for ordinal, interval, and ratio outcomes include the number of missing values (if any), number of non-missing values, mean with standard deviation, and Tukey's five-point summary (minimum, 25th percentile, median, 75th percentile, maximum). In addition to tabular presentation, the distribution of each outcome is visualized by treatment arm and follow-up using a combination of box and violin plots, where appropriate. Summary statistics for nominal outcomes include the number of missing values (if any), number of non-missing values, and absolute and relative frequencies.

~~Summary statistics are calculated for both the full analysis set and the per-protocol analysis set (defined in Section ?? and Section ??, respectively).~~

8.2 Statistical Modelling

~~To evaluate primary and secondary outcomes, separate Analysis of Covariance (ANCOVA) models are employed to estimate treatment effects. The model equation is:~~

$$~~Y_t = \theta + \beta \times T + \gamma \times Y_0 + \boldsymbol{\delta} \bullet \boldsymbol{X} + \varepsilon,~~$$

~~where Y_t denotes the outcome value at month t post enrollment, T is a binary indicator for the intervention group, Y_0 denotes the outcome value at baseline, and $\boldsymbol{X}$ denotes a set of confounding variables (sex, age at baseline, disease duration in months at baseline, disease severity on the Bell disability scale at baseline or maximal hand grip strength at baseline, disease trigger, vocational training status at baseline, and employment status at baseline). The remainder term ε captures residual deviations in the postulated relationship.~~

Generalized linear (mixed) models are used for all statistical analyses, with the model choice depending on the scale of the respective outcome. The model fits are primarily assessed visually using quantile-quantile and fitted versus residuals plots. If there is evidence of non-linearity, heteroscedasticity, or strong deviations from normality (symmetry), data transformations are considered (see Section 7.4).

8.2.1 Primary Analysis

~~The primary analysis utilized the ANCOVA model, as specified, for $t = 12$.~~ The primary analysis employs an Analysis of Covariance (ANCOVA) model to assess the treatment effect on physical functioning (SF-36) at month 12. This model estimates the difference between the intervention and control groups, adjusted for baseline scores and the covariates detailed in Section 5.

The two-sided significance level is set to $\alpha = 5\ \%$.

~~The primary analysis uses the full analysis set (defined in Section **??**), followed by a sensitivity analysis using the per-protocol analysis set (defined in Section **??**).~~ The primary analysis is conducted on sub-population B. To assess the robustness of the results, subsequent sensitivity analyses are performed on sub-populations C, A, and P.

8.2.2 Secondary Analyses

~~The secondary analyses also employ the ANCOVA model, as specified above. For longitudinal analyses, the ANCOVA model is modified to account for repeated measurements within the intervention group. Specifically, a set $\boldsymbol{C}$ of time index indicators for the 3-, 6-, 9-, and 12-month follow-ups, and an individual-specific random effect τ are included in the model. Since longi-tudinal analyses, which include time points $t = 0, 3, 6, 9, 12$, are limited to the intervention group, both coefficients β and γ are constrained to zero to ensure identifiability.~~ Separate ANCOVA models are employed to compare the intervention and control groups at month 12. Each model includes the outcome's baseline value and the specific covariates listed in Section 5 as independent variables. In the case of a binary outcome, the ANCOVA framework is extended to a generalized linear model (GLM) using a logit link function. For longitudinal analyses in the intervention group, linear or logistic mixed models are utilized. Each model includes the timepoint (linear), the outcome's baseline, and the covariates specified in Section 5 as fixed effects. Patient-specific random intercepts are included to account for repeated measurements per patient.

All secondary analyses are exploratory.

All secondary analyses are conducted on sub-population B and C. Sensitivity analyses are performed on sub-populations A and P.

8.2.3 Subgroup Analyses

Differences in treatment effects across subgroups are assessed using the linear model framework with a subgroup-by-treatment interaction. If feasible, the following subgroups are investigated:

- Women vs. men: With ME/CFS occurring twice as often in women, their response to treatment effectiveness may vary compared to men.
- Severe vs. moderate-severe: the effectiveness may vary dependening on the disease severity at baseline (measured by the Bell disability scale).
- ~~Fatigue onset due to COVID-19 infection vs. other causes: Because different causes of fatigue onset likely involve distinct pathophysiological mechanisms, treatment effectiveness may vary.~~
- ~~With breathing-related sleep disorders vs. insomnia: Insomnia is a common comorbidity in ME/CFS patients, and it may affect treatment effectiveness.~~

- ~~With vs. without postural tachycardia syndrome: Postural tachycardia syndrome is a common comorbidity in ME/CFS patients, and it may affect treatment effectiveness.~~

All subgroup analyses are exploratory.

All subgroup analyses are conducted on sub-populations B and C, and only for the primary outcome.

8.2.4 Cost Effectiveness Analysis

Cost-effectiveness is assessed by evaluating the relationship between incremental costs and gains in quality of life. The incremental cost-effectiveness ratio (ICER) represents the average additional cost required to achieve one extra QALY. The ICER is defined as:

$$\text{ICER} = \frac{\overline{(\text{cost intervention group})} - \overline{(\text{cost control group})}}{\overline{(\text{QALY intervention group})} - \overline{(\text{QALY control group})}}.$$

The bars represent means, which are estimated using the previously described ANCOVA models. If log-transformations are employed, Duan's smearing factor [1] is used for the back transform. To account for statistical uncertainty, 95 % confidence intervals for incremental costs, incremental QALYs, and the ICER are generated using a non-parametric bootstrap procedure with 2,000 replications. The results are interpreted against a threshold of 50,000 € per QALY. If the calculated ratio falls below this limit, the intervention is considered cost-effective.

The cost-effectiveness analysis is exploratory.

The cost-effectiveness analysis is conducted on sub-populations B and C.

9 Reporting of Results

Summary statistics are reported by treatment arm and follow-up. Estimated adjusted (marginal) mean differences are presented with p-values, 95 % confidence intervals, and explained variance. Other metrics such as R^2, or the random effects variance may also be reported.

P-values and confidence intervals are derived from corresponding t-distributions. Explained variance is estimated using partial η^2.

All non-integer values are rounded to reasonable number of decimal places. P-values and partial η^2 are rounded always to three decimal places.

10 Software

All statistical analyses are conducted using R (version ≥4.4+) with the packages: `tidyverse` (version ≥1.3+), `lme4` (version ≥1.1+), and `mice` (version ≥3.17+).

11 References

[1] N. Duan. “Smearing estimate: a nonparametric retransformation method.” *Journal of the American Statistical Association* 78.383 (1983), pp. 605–610.

[2] D. B. Rubin. *Multiple Imputation for Nonresponse in Surveys*. John Wiley & Sons, 1987.

[3] R. P. Troiano et al. “Physical Activity in the United States Measured by Accelerometer.” *Medicine & Science in Sports & Exercise* 40.1 (2008), pp. 181–188.

[4] M. Brown, N. Khorana, and L. A. Jason. “The role of changes in activity as a function of perceived available and expended energy in nonpharmacological treatment outcomes for ME/CFS.” *Journal of Clinical Psychology* 67.3 (2011), pp. 253–260.

[5] UNESCO Institute for Statistics. *International Standard Classification of Education: ISCED 2011*. Tech. rep. UNESCO, 2011.

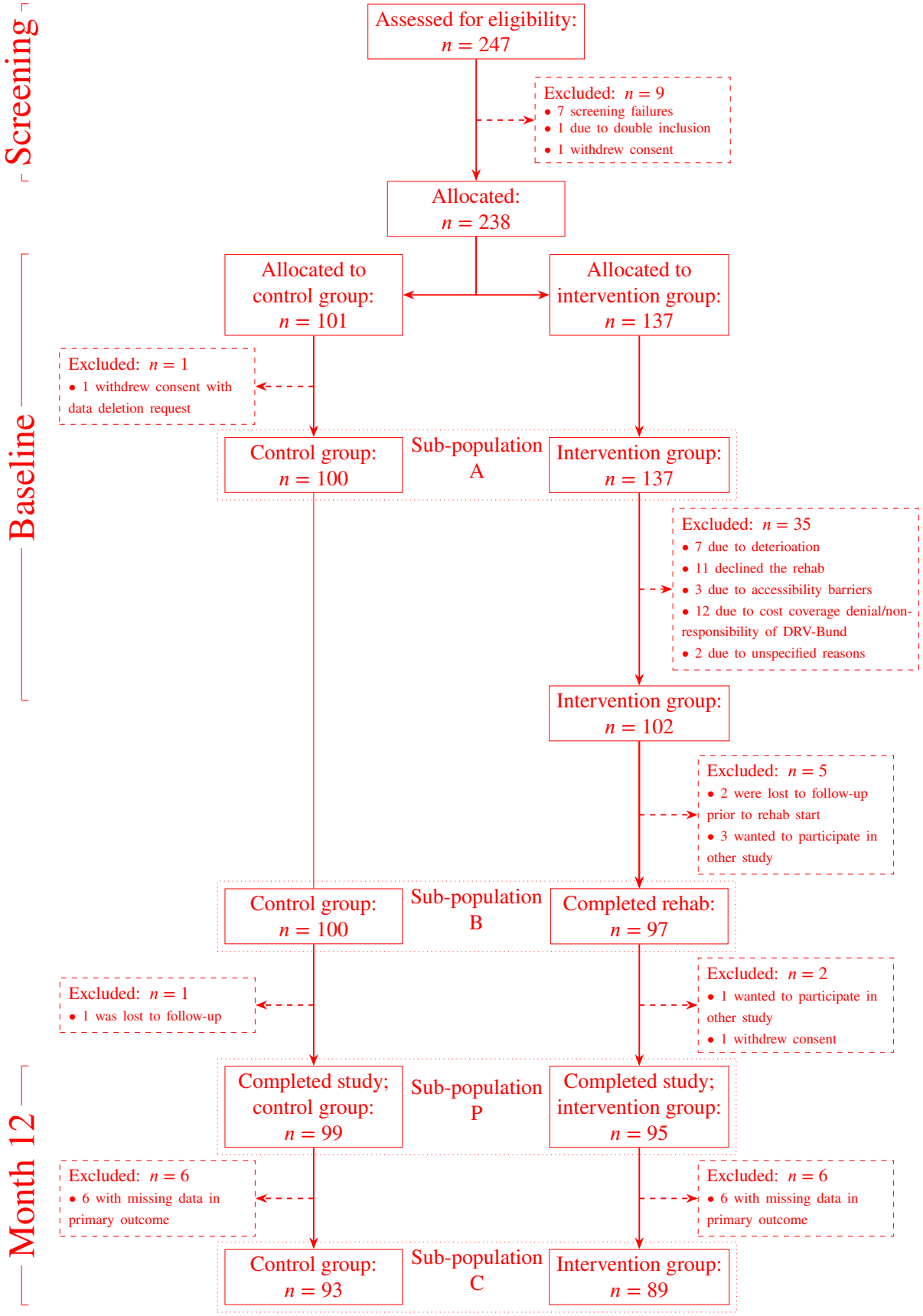


Figure 1: Schematic flow of the study population.

Baseline-Charakteristika der Patient:innen in Auswertepopulation A

Tabelle D1: Grundlegende Merkmale der Auswertepopulation A.

Merkmal	Gesamt N = 237	Kontrollgruppe N = 137	Interventionsgruppe N = 100	SMD [a]
Patient:innendaten				
Alter [Jahre], MW (SD)	41,4 (10,5)	39,3 (11,1)	42,8 (9,9)	0,3
Body Mass Index [kg/m^2], MW (SD)	25,0 (5,1)	25,0 (5,0)	25,0 (5,2)	0,0
Geschlecht, n (%)				0,2
weiblich	201 (15,2 %)	83 (83,0 %)	111 (82,2 %)	
männlich	36 (84,2 %)	17 (17,0 %)	24 (17,8 %)	
Alkohol und Rauchen				
Derzeitiger Alkoholkonsum, n (%)				0,2
Kein derzeitiger Alkoholkonsum	105 (46,1 %)	37 (39,4 %)	68 (50,7 %)	
1x pro Woche	70 (30,7 %)	36 (38,3 %)	34 (25,4 %)	
2x pro Woche	44 (19,3 %)	16 (17,0 %)	28 (20,9 %)	
3x pro Woche	9 (3,9 %)	5 (5,3 %)	4 (3,0 %)	
Keine Angabe	9	6	3	
Rauchen, n (%)				0,0
Nichtraucher	137 (59,1 %)	56 (57,1 %)	81 (60,4 %)	
Ehemalige:r Raucher:in	75 (32,3 %)	33 (33,7 %)	42 (31,3 %)	
Derzeitige:r Raucher:in	20 (8,6 %)	9 (9,2 %)	11 (8,2 %)	
Keine Angabe	5	2	3	
Komorbiditäten (Patient:innenangabe), n (%)				
Allergien	143 (60,3 %)	51 (51,0 %)	92 (67,2 %)	0,3
Kardiovaskuläre Erkrankungen [1]	33 (13,9 %)	17 (17,0 %)	16 (11,7 %)	0,1
Gastrointestinale Erkrankungen [2]	22 (9,3 %)	8 (8,0 %)	14 (10,2 %)	0,1

Merkmal	Gesamt N = 237	Kontrollgruppe N = 137	Interventionsgruppe N = 100	SMD [a]
Erkrankungen des Urogenitaltraktes [3]	16 (6,8 %)	10 (10,0 %)	6 (4,4 %)	0,2
Endokrine Erkrankungen [4]	46 (19,4 %)	22 (22,0 %)	24 (17,5 %)	0,1
Muskuloskelettale Erkrankungen [5]	14 (5,9 %)	7 (7,0 %)	7 (5,1 %)	0,1
Neurologische Störungen [6]	18 (7,6 %)	9 (9,0 %)	9 (6,6 %)	0,1
Pulmonale Erkrankungen [7]	55 (23,2 %)	23 (23,0 %)	32 (23,4 %)	0,0
Hauterkrankungen [8]	15 (6,3 %)	6 (6,0 %)	9 (6,6 %)	0,0
Psychische Störungen [9]	24 (10,1 %)	7 (7,0 %)	17 (12,4 %)	0,2
Andere [10]	27 (11,4 %)	14 (14,0 %)	13 (9,5 %)	0,1
Familie und Kinder				
Familienstand, n (%)				0,1
Verheiratet	96 (42,7 %)	42 (43,8 %)	54 (41,9 %)	
Getrennt	19 (8,4 %)	6 (6,2 %)	13 (10,1 %)	
Ledig	110 (48,9 %)	48 (50,0 %)	62 (48,1 %)	
Keine Angabe	12	4	8	
Kinder, n (%)				0,0
Ja	68 (30,6 %)	31 (33,0 %)	37 (28,9 %)	
Nein	154 (69,4 %)	63 (67,0 %)	91 (71,1 %)	
Keine Angabe	15	6	9	
Ausbildung und Beschäftigung				
Ausbildung, n (%)				0,1
Kein Abschluss, noch in Ausbildung	15 (6,9 %)	7 (7,4 %)	8 (6,6 %)	
Beruflich-betrieblich	59 (27,2 %)	27 (28,4 %)	32 (26,2 %)	
Beruflich-schulisch, Fachschule	36 (16,6 %)	12 (12,6 %)	24 (19,7 %)	

Merkmal	Gesamt N = 237	Kontrollgruppe N = 137	Interventionsgruppe N = 100	SMD [a]
Fachhochschule, Universität	104 (47,9 %)	48 (50,5 %)	56 (45,9 %)	
Andere	3 (1,4 %)	1 (1,1 %)	2 (1,6 %)	
Keine Angabe	20	5	15	
Internationale Standardklassifikation für das Bildungswesen (ISCED), n (%)				0,1
Niedrig	7 (3,1 %)	5 (5,1 %)	2 (1,6 %)	
Mittel	100 (44,2 %)	38 (38,8 %)	62 (48,4 %)	
Hoch	119 (52,7 %)	55 (56,1 %)	64 (50,0 %)	
Beschäftigungsstatus, n (%)				0,0
Beschäftigt	208 (91,6 %)	89 (90,8 %)	119 (92,2 %)	
Nicht beschäftigt	19 (8,4 %)	9 (9,2 %)	10 (7,8 %)	
Keine Angabe	10	2	8	
Krankheitsmerkmale				
Schweregrad [Bell], MW (SD)	36,5 (7,5)	36,5 (7,8)	36,6 (7,3)	0,0
Dauer [Monate], MW (SD)	19,0 (11,7)	21,1 (14,0)	17,5 (9,5)	0,3
Auslöser (Trigger), n (%)				0,1
Covid-19	186 (80,2 %)	77 (78,6 %)	109 (81,3 %)	
Epstein-Barr-Virus	7 (3,0 %)	5 (5,1 %)	2.0 (1,5 %)	
Andere Atemwegsinfektionen	25 (10,8 %)	9 (9,2 %)	16 (11,9 %)	
Andere Infektion	5 (2,2 %)	3 (3,1 %)	2 (1,5 %)	
Nicht infektiös	9 (3,9 %)	4 (4,1 %)	5 (3,7 %)	
Unbekannt, keine Angabe	5	2	3	

MW: Mittelwert (arithemtisch), SD: Standardabweichung, n: Absolute Häufigkeit.
[a] Standardisierte Mittelwertsdifferenz (berechnet auf den relativen Häufigkeiten für kategoriale Merkmale).
[1] Arterieller Hypertonus, Herzrhythmusstörungen, usw.
[2] Reizdarmsyndrom, Gastritis, usw.
[3] Endometriose, Blasenstörungen, usw.
[4] Schilddrüsenüber- oder -unterfunktion, Diabetes mellitus, usw.
[5] Arthrose, Fibromyalgie Syndrom, usw.
[6] Migräne, Polyneuropathie, usw.
[7] Asthma, chronische Sinusitis, usw.
[8] Neurodermitis, Ekzem, usw.
[9] Depression, Angststörungen, usw.

[10] Tinnitus, Lip-/Lymphödem, u. a.
[11] Inklusive von Ehepartner:in getrennt lebend.
[12] Definiert als: Kein Schulabschluss, noch keinen Schulabschluss, Haupt-, Volks-, oder Realschulabschluss oder Abschluss an einer polytechnischen Oberschule; keine beruflich-betrieblich oder beruflich-schulische Ausbildung.
[13] Definiert als: Beruflich-betriebliche Ausbildung, beruflich-schulische Ausbildung, Fachhochschulreife, Abitur, allgemeine oder fachgebundene Hochschulreife, oder erweiterte Oberschule; keinen Abschluss an einer Fachschule, Fachakademie, Fachhochschule, Ingenieurschule, Universität oder Hochschule.
[14] Definiert als: Abschluss an einer Fachschule, Fachakademie, Fachhochschule, Ingenieurschule, Universität oder Hochschule.
[15] Pneumonie, Bronchitis, Sinusitis, Otitis, Influenza, usw.
[16] Meningitis, Pyelonephritis, Borreliose, Herpes, Zytomegalovirus.
[17] Impfung, körperliche Belastung.

Baseline-Charakteristika der Auswertepopulation P

Tabelle D2: Grundlegende Merkmale der Auswertepopulation P.

Merkmal	Gesamt N = 197	Kontrollgruppe N = 100	Interventionsgruppe N = 97	SMD [a]
Patient:innendaten				
Alter [Jahre], MW (SD)	41,6 (10,5)	39,3 (11,1)	44,1 (9,3)	0,5
Body Mass Index [kg/m2], MW (SD)	25,1 (5,1)	24,9 (4,9)	25,3 (5,4)	0,1
Geschlecht, n (%)				0,1
weiblich	168 (86,6 %)	83 (83,8 %)	85 (89,5 %)	
männlich	26 (13,4 %)	16 (16,2 %)	10 (10,5 %)	
Alkohol und Rauchen				
Derzeitiger Alkoholkonsum, n (%)				0,1
kein derzeitiger Alkoholkonsum	77 (41,2 %)	36 (38,7 %)	41 (43,6 %)	
1x pro Woche	65 (34,8 %)	36 (38,7 %)	29 (30,9 %)	
2x pro Woche	36 (19,3 %)	16 (17,2 %)	20 (21,3 %)	
3x pro Woche	9 (4,8 %)	5 (5,4 %)	4 (4,3 %)	
Keine Angabe	7	6	1	
Rauchen, n (%)				0,1
Nichtraucher	112 (58,9 %)	55 (56,7 %)	57 (61,3 %)	
Ehemalige:r Raucher:in	61 (31,2 %)	33 (34,0 %)	28 (30,1 %)	
Derzeitige:r Raucher:in	17 (8,9 %)	9 (9,3 %)	8 (8,6 %)	
Keine Angabe	7	2	2	
Komorbiditäten (Patient:innenangabe), n (%)				
Allergien	109 (56,2 %)	50 (50,5 %)	59 (62,1 %)	0,2
Kardiovaskuläre Erkrankungen [1]	29 (14,9 %)	17 (17,2 %)	12 (12,6 %)	0,1
Gastrointestinale Erkrankungen [2]	17 (8,8 %)	8 (8,1 %)	9 (9,5 %)	0,0
Erkrankungen des Urogenitaltraktes [3]	14 (7,2 %)	10 (10,1 %)	4 (4,2 %)	0,2
Endokrine Erkrankungen [4]	41 (21,1 %)	22 (22,2 %)	19 (20,0 %)	0,1
Muskuloskelettale Erkrankungen [5]	11 (5,7 %)	7 (7,1 %)	4 (4,2 %)	0,1

Merkmal	Gesamt N = 197	Kontrollgruppe N = 100	Interventionsgruppe N = 97	SMD [a]
Neurologische Störungen [6]	15 (7,7 %)	9 (9,1 %)	6 (6,3 %)	0,1
Pulmonale Erkrankungen [7]	45 (23,2 %)	22 (22,2 %)	23 (24,2 %)	0,0
Hauterkrankungen [8]	11 (5,7 %)	6 (6,1 %)	5 (5,3 %)	0,0
Psychische Störungen [9]	20 (10,3 %)	7 (7,1 %)	13 (13,7 %)	0,2
Andere [10]	22 (11,3 %)	14 (14,1 %)	8 (8,4 %)	0,2
Familie und Kinder				
Familienstand, n (%)				0,1
Verheiratet	82 (44,3 %)	42 (44,2 %)	40 (44,4 %)	
Getrennt	16 (8,6 %)	6 (6,3 %)	10 (11,1 %)	
Ledig	87 (47,0 %)	47 (49,5 %)	40 (44,4 %)	
Keine Angabe	9	4	5	
Kinder, n (%)				0,0
Ja	59 (32,4 %)	31 (33,0 %)	28 (31,5 %)	
Nein	123 (67,6 %)	62 (66,7 %)	61 (68,5 %)	
Keine Angabe	12	6	6	
Ausbildung und Beschäftigung				
Ausbildung, n (%)				0,1
Kein Abschluss, noch in Ausbildung	12 (6,7 %)	7 (7,4 %)	5 (5,9 %)	
Beruflich-betrieblich	52 (29,1 %)	26 (27,7 %)	26 (30,6 %)	
Beruflich-schulisch, Fachschule	26 (14,5 %)	12 (12,8 %)	14 (16,5 %)	
Fachhochschule, Universität	86 (48,0 %)	48 (51,1 %)	38 (44,7 %)	
Andere	3 (1,7 %)	1 (1,1 %)	2 (2,4 %)	
Keine Angabe	9	2	7	
Internationale Standardklassifikation für das Bildungswesen (ISCED), n (%)				0,2
Niedrig	6 (3,2 %)	5 (5,2 %)	1 (1,1 %)	
Mittel	79 (42,7 %)	37 (38,1 %)	42 (47,7 %)	

Merkmal	Gesamt N = 197	Kontrollgruppe N = 100	Interventionsgruppe N = 97	SMD [a]
Hoch	100 (54,1 %)	55 (56,7 %)	45 (51,1 %)	
Keine Angabe	9	2	7	
Beschäftigungsstatus, n (%)				0,0
Beschäftigt	171 (91,4 %)	88 (90,7 %)	83 (92,2 %)	
Nicht beschäftigt	16 (8,6 %)	9 (9,3 %)	7 (7,8 %)	
Keine Angabe	7	2	5	
Krankheitsmerkmale				
Schweregrad [Bell], MW (SD)	36,6 (7,7)	36,5 (7,8)	36,8 (7,6)	0,0
Dauer [Monate], MW (SD)	19,1 (12,0)	21,2 (14,0)	16,9 (9,1)	0,4
Auslöser (Trigger), n (%)				0,1
Covid-19	152 (79,2 %)	76 (78,4 %)	76 (80,0 %)	
Epstein-Barr-Virus	7 (3,6 %)	5 (5,2 %)	2 (2,1 %)	
Andere Atemwegsinfektionen	19 (9,9 %)	9 (9,3 %)	10 (10,5 %)	
Andere Infektion	5 (2,6 %)	3 (3,1 %)	2 (2,1 %)	
Nicht infektiös	9 (4,7 %)	4 (4,1 %)	5 (5,3 %)	
Unbekannt, keine Angabe	2	2	0	

MW: Mittelwert (arithemtisch), SD: Standardabweichung, n: Absolute Häufigkeit.
[a] Standardisierte Mittelwertsdifferenz (berechnet auf den relativen Häufigkeiten für kategoriale Merkmale).
[1] Arterieller Hypertonus, Herzrhythmusstörungen, usw.
[2] Reizdarmsyndrom, Gastritis, usw.
[3] Endometriose, Blasenstörungen, usw.
[4] Schilddrüsenüber- oder -unterfunktion, Diabetes mellitus, usw.
[5] Arthrose, Fibromyalgie Syndrom, usw.
[6] Migräne, Polyneuropathie, usw.
[7] Asthma, chronische Sinusitis, usw.
[8] Neurodermitis, Ekzem, usw.
[9] Depression, Angststörungen, usw.
[10] Tinnitus, Lip-/Lymphödem, u. a.
[11] Inklusive von Ehepartner:in getrennt lebend.

[12] Definiert als: Kein Schulabschluss, noch keinen Schulabschluss, Haupt-, Volks-, oder Realschulabschluss oder Abschluss an einer polytechnischen Oberschule; keine beruflich-betrieblich oder beruflich-schulische Ausbildung.
[13] Definiert als: Beruflich-betriebliche Ausbildung, beruflich-schulische Ausbildung, Fachhochschulreife, Abitur, allgemeine oder fachgebundene Hochschulreife, oder erweiterte Oberschule; keinen Abschluss an einer Fachschule, Fachakademie, Fachhochschule, Ingenieurschule, Universität oder Hochschule.
[14] Definiert als: Abschluss an einer Fachschule, Fachakademie, Fachhochschule, Ingenieurschule, Universität oder Hochschule.
[15] Pneumonie, Bronchitis, Sinusitis, Otitis, Influenza, usw.
[16] Meningitis, Pyelonephritis, Borreliose, Herpes, Zytomegalovirus.
[17] Impfung, körperliche Belastung.

Baseline-Charakteristika der BKK-Versicherungsnehmer:innen in Auswertepopulation B

Tabelle D3: Grundlegende Merkmale der BKK-Versicherungsnehmer:innen in Auswertepopulation B.

Merkmal	Gesamt N = 46	Kontrollgruppe N = 25	Interventionsgruppe N = 21	SMD [a]
Patient:innendaten				
Alter [Jahre], MW (SD)	39,8 (11,6)	36,5 (13,0)	43,8 (8,5)	0,7
Body Mass Index [kg/m^2], MW (SD)	26,4 (5,8)	25,7 (5,3)	27,2 (6,3)	0,3
Geschlecht, n (%)				0,1
weiblich	38 (82,6 %)	20 (80,0 %)	18 (85,7 %)	
männlich	8 (17,4 %)	5 (20,0 %)	3 (14,3 %)	
Alkohol und Rauchen				
Derzeitiger Alkoholkonsum, n (%)				0,2
kein derzeitiger Alkoholkonsum	20 (45,5 %)	12 (50,0 %)	8 (40,0 %)	
1x pro Woche	15 (34,1 %)	9 (37,5 %)	6 (30,0 %)	
2x pro Woche	8 (18,2 %)	3 (12,5 %)	5 (25,0 %)	
3x pro Woche	1 (2,3 %)	0 (0,0 %)	1 (5,0 %)	
Keine Angabe	2	1	1	
Rauchen, n (%)				0,1
Nichtraucher	30 (66,7 %)	17 (70,8 %)	13 (61,9 %)	
Ehemalige:r Raucher:in	13 (28,9 %)	6 (25,0 %)	7 (33,3 %)	
Derzeitige:r Raucher:in	2 (5,0 %)	1 (4,2 %)	1 (4,8 %)	
Keine Angabe	1	1	0	
Komorbiditäten (Patient:innenangabe), n (%)				
Allergien	25 (54,3 %)	13 (52,0 %)	12 (57,1 %)	0,1
Kardiovaskuläre Erkrankungen [1]	5 (10,9 %)	3 (12,0 %)	2 (9,5 %)	0,1

Gastrointestinale Erkrankungen [2]	7 (15,2 %)	4 (16,0 %)	3 (14,3 %)	0,0
Erkrankungen des Urogenitaltraktes [3]	4 (8,7 %)	3 (12,0 %)	1 (5,0 %)	0,2
Endokrine Erkrankungen [4]	9 (19,6 %)	6 (24,0 %)	3 (14,3 %)	0,2
Muskuloskelettale Erkrankungen [5]	4 (8,7 %)	3 (12,0 %)	1 (5,0 %)	0,3
Neurologische Störungen [6]	6 (13,0 %)	4 (16,0 %)	2 (9,5 %)	0,2
Pulmonale Erkrankungen [7]	12 (26,1 %)	8 (32,0 %)	4 (19,0 %)	0,3
Hauterkrankungen [8]	3 (6,5 %)	2 (8,0 %)	1 (4,8 %)	0,1
Psychische Störungen [9]	5 (10,9 %)	1 (4,0 %)	4 (19,0 %)	0,5
Andere [10]	5 (10,9 %)	3 (12,0 %)	2 (9,5 %)	0,0
Familie und Kinder				
Familienstand, n (%)				0,2
Verheiratet [11]	19 (41,3 %)	7 (28,0 %)	12 (57,1 %)	
Getrennt	4 (8,7 %)	2 (8,0 %)	2 (9,5 %)	
Ledig	23 (50,0 %)	16 (57,1 %)	7 (33,3 %)	
Keine Angabe	0	0	0	
Kinder, n (%)				0,1
Nein	35 (76,1 %)	18 (72,0 %)	17 (81,0 %)	
Ja	11 (23,9 %)	7 (28,0 %)	4 (19,0 %)	
Keine Angabe	0	0	0	
Ausbildung und Beschäftigung				
Ausbildung, n (%)				0,4

Kein Abschluss, noch in Ausbildung	4 (9,1 %)	3 (12,5 %)	1 (5,0 %)	
Beruf-betrieblich	14 (31,8 %)	5 (20,8 %)	9 (45,0 %)	
Beruflich-schulisch, Fachschule	10 (22,7 %)	4 (16,7 %)	6 (30,0 %)	
Fachhochschule, Universität	15 (34,1 %)	12 (50,0 %)	3 (15,0 %)	
Anderer Abschluss	1 (2,3 %)	0 (0,0 %)	1 (5,0 %)	
Keine Angabe	2	1	1	
Internationale Standardklassifikation für das Bildungswesen (ISCED), n (%)				0,5
Niedrig [12]	4 (8,9 %)	3 (12,0 %)	1 (5,0 %)	
Mittel [13]	20 (44,4 %)	7 (28,0 %)	13 (65,0 %)	
Hoch [14]	21 (46,7 %)	15 (60,0 %)	6 (30,0 %)	
Keine Angabe	1	0	1	
Beschäftigungsstatus, n (%)				0,2
Beschäftigt	38 (82,6 %)	20 (80,0 %)	18 (85,7 %)	
Nicht beschäftigt	8 (17,4 %)	5 (20,0 %)	3 (14,3 %)	
Keine Angabe	0	0	0	
Krankheitsmerkmale				
Schweregrad [Bell], MW (SD)	35,7 (7,5)	34,4 (6,6)	37,1 (8,5)	0,4
Dauer [Monate], MW (SD)	20,9 (14,3)	25,2 (16,5)	15,9 (9,4)	0,7
Auslöser (Trigger), n (%)				0,2
Covid-19	30 (68,2 %)	15 (65,2 %)	15 (71,4 %)	
Epstein-Barr-Virus	5 (11,4 %)	3 (13,0 %)	2 (9,5 %)	
Andere Atemwegsinfektionen [15]	5 (11,4 %)	3 (13,0 %)	2 (9,5 %)	
Andere Infektionen [16]	1 (2,3 %)	1 (4,3 %)	0 (0,0 %)	
Keine Infektion [17]	3 (6,8 %)	1 (4,3 %)	0 (0,0 %)	

Unbekannt, keine Angabe	2	2	0

MW: Mittelwert (arithemtisch), SD: Standardabweichung, n: Absolute Häufigkeit.
[a] Standardisierte Mittelwertsdifferenz (berechnet auf den relativen Häufigkeiten für kategoriale Merkmale).
[1] Arterieller Hypertonus, Herzrhythmusstörungen, usw.
[2] Reizdarmsyndrom, Gastritis, usw.
[3] Endometriose, Blasenstörungen, usw.
[4] Schilddrüsenüber- oder -unterfunktion, Diabetes mellitus, usw.
[5] Arthrose, Fibromyalgie Syndrom, usw.
[6] Migräne, Polyneuropathie, usw.
[7] Asthma, chronische Sinusitis, usw.
[8] Neurodermitis, Ekzem, usw.
[9] Depression, Angststörungen, usw.
[10] Tinnitus, Lip-/Lymphödem, u. a.
[11] Inklusive von Ehepartner:in getrennt lebend.
[12] Definiert als: Kein Schulabschluss, noch keinen Schulabschluss, Haupt-, Volks-, oder Realschulabschluss oder Abschluss an einer polytechnischen Oberschule; keine beruflich-betrieblich oder beruflich-schulische Ausbildung.
[13] Definiert als: Beruflich-betriebliche Ausbildung, beruflich-schulische Ausbildung, Fachhochschulreife, Abitur, allgemeine oder fachgebundene Hochschulreife, oder erweiterte Oberschule; keinen Abschluss an einer Fachschule, Fachakademie, Fachhochschule, Ingenieurschule, Universität oder Hochschule.
[14] Definiert als: Abschluss an einer Fachschule, Fachakademie, Fachhochschule, Ingenieurschule, Universität oder Hochschule.
[15] Pneumonie, Bronchitis, Sinusitis, Otitis, Influenza, usw.
[16] Meningitis, Pyelonephritis, Borreliose, Herpes, Zytomegalovirus.
[17] Impfung, körperliche Belastung.

Baseline-Charakteristika der BKK-Versicherungsnehmer:innen in Auswertepopulation C

Tabelle D4: Grundlegende Merkmale der BKK-Versicherungsnehmer:innen in Auswertepopulation C.

Merkmal	Gesamt N = 41	Kontrollgruppe N = 21	Interventionsgruppe N = 20	SMD [a]
Patient:innendaten				
Alter [Jahre], MW (SD)	41,0 (11,6)	37,6 (13,7)	44,7 (7,6)	0,6
Body Mass Index [kg/m^2], MW (SD)	26,5 (5,8)	25,7 (5,2)	27,4 (6,4)	0,3
Geschlecht, n (%)				0,1
weiblich	35 (85,4 %)	17 (81,0 %)	18 (90,0 %)	
männlich	6 (14,6 %)	4 (19,0 %)	2 (10,0 %)	
Alkohol und Rauchen				
Derzeitiger Alkoholkonsum, n (%)				0,3
kein derzeitiger Alkoholkonsum	17 (43,6 %)	10 (50,0 %)	7 (36,8 %)	
1x pro Woche	14 (35,9 %)	8 (40,0 %)	6 (31,6 %)	
2x pro Woche	7 (17,9 %)	2 (10,0 %)	5 (26,3 %)	
3x pro Woche	1 (2,6 %)	0 (0,0 %)	1 (5,3 %)	
Keine Angabe	2	1	1	
Rauchen, n (%)				0,0
Nichtraucher	26 (65,0 %)	13 (65,0 %)	13 (65,0 %)	
Ehemalige:r Raucher:in	12 (30,0 %)	6 (30,0 %)	6 (30,0 %)	
Derzeitige:r Raucher:in	2 (5,0 %)	1 (5,0 %)	1 (5,0 %)	
Keine Angabe	1	1	0	
Komorbiditäten (Patient:innenangabe), n (%)				
Allergien	22 (53,7 %)	10 (47,6 %)	12 (60,0 %)	0,2
Kardiovaskuläre Erkrankungen [1]	5 (12,2 %)	3 (14,3 %)	2 (10,0 %)	0,1

Merkmal	Gesamt N = 41	Kontrollgruppe N = 21	Interventionsgruppe N = 20	SMD [a]
Gastrointestinale Erkrankungen [2]	6 (14,6 %)	3 (14,3 %)	3 (15,0 %)	0,0
Erkrankungen des Urogenitaltraktes [3]	3 (7,3 %)	2 (9,5 %)	1 (5,0 %)	0,2
Endokrine Erkrankungen [4]	9 (22,0 %)	6 (28,6 %)	3 (15,0 %)	0,3
Muskuloskelettale Erkrankungen [5]	4 (9,8 %)	3 (14,3 %)	1 (5,0 %)	0,3
Neurologische Störungen [6]	6 (14,6 %)	4 (19,0 %)	2 (10,0 %)	0,3
Pulmonale Erkrankungen [7]	11 (26,8 %)	7 (33,3 %)	4 (20,0 %)	0,3
Hauterkrankungen [8]	3 (7,3 %)	2 (9,5 %)	1 (5,0 %)	0,2
Psychische Störungen [9]	5 (12,2 %)	1 (4,8 %)	4 (20,0 %)	0,5
Andere [10]	4 (9,8 %)	2 (9,5 %)	2 (10,0 %)	0,0
Familie und Kinder				
Familienstand, n (%)				0,2
Verheiratet [11]	19 (46,3 %)	7 (33,3 %)	12 (60,0 %)	
Getrennt	4 (9,8 %)	2 (9,5 %)	2 (10,0 %)	
Ledig	18 (43,9 %)	12 (57,1 %)	6 (30,0 %)	
Keine Angabe	0	0	0	
Kinder, n (%)				0,1
Nein	31 (75,6 %)	15 (71,4 %)	16 (80,0 %)	
Ja	10 (24,4 %)	6 (28,6 %)	4 (20,0 %)	

Merkmal	Gesamt N = 41	Kontrollgruppe N = 21	Interventionsgruppe N = 20	SMD [a]
Keine Angabe	0	0	0	
Ausbildung und Beschäftigung				
Ausbildung, n (%)				0,5
Kein Abschluss, noch in Ausbildung	4 (10,3 %)	3 (15,0 %)	1 (5,3 %)	
Beruf-betrieblich	12 (30,8 %)	3 (15,0 %)	9 (47,4 %)	
Beruflich-schulisch, Fachschule	8 (20,5 %)	3 (15,0 %)	5 (26,3 %)	
Fachhochschule, Universität	14 (35,9 %)	11 (55,0 %)	3 (15,8 %)	
Anderer Abschluss	1 (2,6 %)	0 (0,0 %)	1 (5,3 %)	
Keine Angabe	2	1	1	
Internationale Standardklassifikation für das Bildungswesen (ISCED), n (%)				0,7
Niedrig [12]	4 (10,0 %)	3 (14,3 %)	1 (5,3 %)	
Mittel [13]	18 (45,0 %)	5 (23,8 %)	13 (68,4 %)	
Hoch [14]	18 (45,0 %)	15 (61,9 %)	5 (26,3 %)	
Keine Angabe	1	0	1	
Beschäftigungsstatus, n (%)				0,2
Beschäftigt	34 (82,9 %)	16 (76,2 %)	18 (90,0 %)	
Nicht beschäftigt	7 (17,1 %)	5 (23,8 %)	2 (10,0 %)	
Keine Angabe	0	0	0	
Krankheitsmerkmale				
Schweregrad [Bell], MW (SD)	35,9 (7,7)	34,3 (6,8)	37,5 (8,5)	0,4
Dauer [Monate], MW (SD)	21,4 (14,9)	26,8 (17,1)	15,8 (9,6)	0,8
Auslöser (Trigger), n (%)				0,2

Merkmal	Gesamt N = 41	Kontrollgruppe N = 21	Interventionsgruppe N = 20	SMD [a]
Covid-19	25 (64,1 %)	11 (57,9 %)	14 (70,0 %)	
Epstein-Barr-Virus	5 (12,8 %)	3 (15,8 %)	2 (10,0 %)	
Andere Atemwegsinfektionen [15]	5 (12,8 %)	3 (15,8 %)	2 (10,0 %)	
Andere Infektionen [16]	1 (2,6 %)	1 (5,3 %)	0 (0,0 %)	
Keine Infektion [17]	3 (4,9 %)	1 (5,3 %)	0 (0,0 %)	
Unbekannt, keine Angabe	2	2	0	

MW: Mittelwert (arithemtisch), SD: Standardabweichung, n: Absolute Häufigkeit.
[a] Standardisierte Mittelwertsdifferenz (berechnet auf den relativen Häufigkeiten für kategoriale Merkmale).
[1] Arterieller Hypertonus, Herzrhythmusstörungen, usw.
[2] Reizdarmsyndrom, Gastritis, usw.
[3] Endometriose, Blasenstörungen, usw.
[4] Schilddrüsenüber- oder -unterfunktion, Diabetes mellitus, usw.
[5] Arthrose, Fibromyalgie Syndrom, usw.
[6] Migräne, Polyneuropathie, usw.
[7] Asthma, chronische Sinusitis, usw.
[8] Neurodermitis, Ekzem, usw.
[9] Depression, Angststörungen, usw.
[10] Tinnitus, Lip-/Lymphödem, u. a.
[11] Inklusive von Ehepartner:in getrennt lebend.
[12] Definiert als: Kein Schulabschluss, noch keinen Schulabschluss, Haupt-, Volks-, oder Realschulabschluss oder Abschluss an einer polytechnischen Oberschule; keine beruflich-betrieblich oder beruflich-schulische Ausbildung.
[13] Definiert als: Beruflich-betriebliche Ausbildung, beruflich-schulische Ausbildung, Fachhochschulreife, Abitur, allgemeine oder fachgebundene Hochschulreife, oder erweiterte Oberschule; keinen Abschluss an einer Fachschule, Fachakademie, Fachhochschule, Ingenieurschule, Universität oder Hochschule.
[14] Definiert als: Abschluss an einer Fachschule, Fachakademie, Fachhochschule, Ingenieurschule, Universität oder Hochschule.
[15] Pneumonie, Bronchitis, Sinusitis, Otitis, Influenza, usw.
[16] Meningitis, Pyelonephritis, Borreliose, Herpes, Zytomegalovirus.
[17] Impfung, körperliche Belastung.

Anlage 6: Anzahl der imputierten Werte je Outcome und Auswertepopulation

Tabelle E1: Anzahl der imputierten Datensätze je Auswertepopulation.

Endpunkt	Auswertepopulation C	P	A	B
Bell Disability Scale	8	14	39	17
Chalder				
Mental	8	20	45	23
Körperlich	8	17	42	20
Gesamtscore	8	23	48	26
CCC				
Kopfschmerzen	8	13	38	16
Muskelschmerz	8	13	38	16
COMPASS-31				
Blase	8	16	43	19
Gastroinestinaltrakt	8	24	50	27
Orthostase	8	18	44	21
Pupillomotorik	7	23	49	26
Sekretomotorik	8	24	51	27
Vasomotorik	8	20	45	23
Gesamtscore	7	47	76	50
SF-36				
Körperliche Funktionsfähigkeit	8	20	45	23
Körperliche Schmerzen	8	20	45	23
Emotionales Wohlbefinden	8	21	46	24
Allgemeine Gesundheitswahrnehmung	7	18	43	21
Änderung des Gesundheitszustands	8	16	41	19
Emotionale Rollenfunktion	8	19	44	22
Körperliche Rollenfunktion	8	17	42	20
Soziale Funktion	8	24	49	27
Vitalität	8	19	44	22
Handgriffkraft				
Fmean (Session 1)	7	24	54	27
Fmean (Session 2)	7	26	57	29
Fmax (Session 1)	7	24	54	27
Fmax (Session 2)	7	26	57	29
Fatigue Ratio (Session 1)	7	24	54	27
Fatigue Ratio (Session 2)	7	26	57	29
Recovery Ratio	7	26	57	29
Schellong-Test (Anlehn-Stehtest)				
Nach 0 Minuten Stehen	7	27	57	30
Nach 2 Minuten Stehen	7	31	61	34
Nach 5 Minuten Stehen	7	39	69	42
Nach 10 Minuten Stehen	7	46	78	49
Durchschnittliche Schrittzahl pro Tag	2	93	51	96

Endpunkt	Auswertepopulation C	P	A	B
IPAQ				
Gehen/leichte Aktivität	1	1	1	1
Moderate Aktivität	0	74	76	75
Anstrengende Aktivität	4	60	92	63
Blutparameter				
Albumin	1	92	57	95
Anorganisches Phosphorat	1	92	57	95
Calcium	1	92	57	95
Creatinkinase	1	92	57	95
Ferritin	1	92	57	95
Interleukin-8 nach Erylyse	1	97	46	100
Laktatdehydrogenase	1	92	57	95
NT-pro-BNP	1	92	57	95
PHQ-9	8	20	45	23
ESS				
EQ-5D-5L Indexwert	7	21	46	24