

تکرارپذیری نسبی و مطلق آزمون Timed Up and Go در سالمندان ساکن اجتماع و جوانان سالم

(مقاله پژوهشی برگرفته از رساله دکتری)

*فرهاد آزادی^۱، محمد پرنیان پور^۲، حسن شاکری^۳، انوشیروان کاظم‌نژاد^۴، احمد علی اکبری کامرانی^۵، امیرمسعود عرب^۶، ایرج عبدالهی^۲

چکیده:

هدف: تکرارپذیری نسبی و مطلق از ویژگی‌های سایکومتریک آزمون هستند که مبنای بسیاری از تصمیم‌گیری‌های بالینی قرار می‌گیرند. در بسیاری مواقع، فقط تکرارپذیری نسبی مبنای تصمیم‌گیری و توجه است؛ درحالی‌که تکرارپذیری مطلق هم اهمیت بسیاری دارد.

روش بررسی: ۱۱ سالمند ساکن اجتماع در دامنه سنی ۶۵ سال و بیشتر ($69/64 \pm 3/58$ سال) و ۲۰ جوان سالم ساکن اجتماع در دامنه سنی ۲۰ تا ۳۵ سال ($28/80 \pm 4/15$ سال)، با استفاده از نسخه سه‌گانه آزمون Timed Up and Go دو بار با فاصله زمانی ۲ تا ۵ روز ارزیابی شدند.

یافته‌ها: به‌طورکلی، غیرهمگن بودن جمعیت موضوع مطالعه، باعث افزایش ضریب هم‌بستگی درون طبقه‌ای شده و این ضریب در سالمندان، بیش از جوانان بوده و افزودن فعالیت ثانویه، باعث کاهش آن می‌شود.

در این تحقیق، شاخص‌های تکرارپذیری مطلق با استفاده از منابع داده‌ها و معادلات مختلف به نتایج کم‌وبیش مشابهی منجر شد. در کل، سالمندان در آزمون مجدد، نیازمند تغییر بیشتری در مقایسه با جوانان هستند تا تغییرات مشاهده شده تغییر واقعی و نه تصادفی قلمداد شود. سهم خطای تصادفی در متغیرهای تحقیق شده در سالمندان، اندکی بیش از جوانان بوده و افزودن فعالیت ثانویه باعث افزایش آن می‌شود. به‌نظر می‌رسد غیرهمگن بودن جمعیت، باعث تعدیل شاخص‌های تکرارپذیری مطلق می‌شود.

نتیجه‌گیری: نتایج این تحقیق نشان داد محققان و درمانگران هنگام توجه به تکرارپذیری نسبی، باید به نکاتی همچون همگن بودن جمعیت موضوع مطالعه توجه کنند. از طرفی، در تصمیم‌گیری‌های بالینی، توجه به تکرارپذیری مطلق در کنار تکرارپذیری نسبی هم لازم است.

کلیدواژه‌ها: آزمون Timed Up and Go، تکرارپذیری نسبی، تکرارپذیری مطلق، جوان، سالمند.

تاریخ پذیرش: ۱۳۹۱/۰۷/۱۰

تاریخ دریافت: ۱۳۹۱/۰۴/۰۱

۱. دانشجوی دکتری فیزیوتراپی، دانشگاه علوم بهزیستی و توانبخشی، مرکز تحقیقات مسائل اجتماعی و روانی سالمندان دانشگاه علوم بهزیستی و توانبخشی تهران (نویسنده مسئول)
۲. عضو هیئت علمی و استاد گروه مکانیک دانشگاه Hanyang کره جنوبی و دانشگاه صنعتی شریف
۳. عضو هیئت علمی و استادیار گروه فیزیوتراپی دانشگاه علوم بهزیستی و توانبخشی
۴. عضو هیئت علمی و استاد گروه آمار زیستی دانشگاه تربیت مدرس
۵. عضو هیئت علمی و دانشیار مرکز تحقیقات مسائل اجتماعی روانی سالمندان دانشگاه علوم بهزیستی و توانبخشی تهران
۶. عضو هیئت علمی و دانشیار گروه فیزیوتراپی دانشگاه علوم بهزیستی و توانبخشی

* آدرس نویسنده مسئول: ۹۹۹

* تلفن: ۹۸ ۹۱۲۱۹۳۰۷۱۹

* رایانامه: fa_azadi@yahoo.com

مقدمه

در این آزمون، فرد از روی یک صندلی با ارتفاع معمولی برمی‌خیزد؛ سه متر راه می‌رود؛ می‌چرخد و همان مسیر را برمی‌گردد و دوباره روی صندلی می‌نشیند. در واقع، این‌ها بخشی از فعالیت‌های زیربنایی روزمره است. مدت زمان انجام این توالی حرکتی ثبت می‌شود. انجمن سالمندی آمریکا و بریتانیا و انجمن سالمندشناسی اروپای شمالی^۱ این آزمون را برای ارزیابی خطر زمین خوردن سالمندان توصیه کرده‌اند (۵، ۷، ۹).

همچنین، مطالعات متعددی نشان داده‌اند فعالیت‌های چندگانه بهتر از فعالیت‌های منفرد، مشکلات تعادل را نمایان می‌سازد (۱۰-۱۲). در راستای چنین مطالعاتی، نسخ دیگری از آزمون معرفی شدند: TUG cognitive و TUG manual (۱۱).

در عصر سالمندشدن جمعیت به‌سر می‌بریم (۱) و زمین خوردن از چالش‌های مهم پیش روی سلامتی سالمندان است (۲-۴) که سلامت و استقلال‌شان را تهدید می‌کند (۲).

برای ارزیابی فعالیت جسمانی و خطر زمین خوردن سالمندان، آزمون‌های متعددی معرفی شده‌اند (۵) که از شناخته‌ترین آن‌ها، آزمون Get Up and Go است. در نسخه اصلی، عملکرد فرد به‌صورت کیفی، با استفاده از مقیاس پنج امتیازی ارزیابی می‌شود. به‌منظور غلبه بر جنبه کیفی، پودسیالود و ریچاردسون به مدت زمان انجام آزمون به‌عنوان متغیری کمی توجه کردند؛ یعنی Timed Up and Go (TUG) (۶-۸).

محققان و درمانگران می‌توانند تغییرات بالینی در وضعیت نمونه یا بیمار را از تغییراتی افتراق دهند که ممکن است به دلیل خطای اندازه‌گیری باشد (۱۴، ۵).

همچنین در تکرارپذیری مطلق، به‌منظور بررسی سهم خطای تصادفی آزمون، از «ضریب واریانس»^۵ و «درصد حداقل تغییر قابل تشخیص» هم استفاده می‌شود. این دو شاخص در تصمیم‌گیری‌های بالینی می‌تواند به درمانگران و محققان کمک کند (۱۷، ۱۶).

در پیشینه تحقیق، از بین آزمون‌ها، فقط تکرارپذیری نسبی و مطلق آزمون TUG original در جمعیت‌های خاصی انجام شده که عبارت‌اند از: بیماران مبتلا به پارکینسون (۱۶) و پارکینسونیسم (۱۸)، سالمندان ساکن اجتماع (۱۹)، بیماران آلزایمر (۲۰)، سالمندان آسایشگاهی (۵) و سالمندان آفریقایی آمریکایی (۲۱). درخصوص دیگر آزمون‌های این تحقیق، مطالعه‌ای صورت نگرفته است.

همان‌طور که می‌دانیم، محققان مختلف از منابع داده‌ای مختلف، یعنی داده‌های آزمون یا داده‌های مختلط (میانگین مجموع داده‌های آزمون و آزمون مجدد) و معادلات مختلف، تکرارپذیری مطلق را بررسی کرده‌اند. حال سؤال این است: آیا مقادیر حاصل شده با هم متفاوت خواهند بود؟ (۱۹، ۱۵) این مقاله چندین هدف دارد:

در اولین هدف، ما تلاش می‌کنیم تأثیر غیرهمگن بودن افراد تحت مطالعه را بر ضریب هم‌بستگی درون طبقه‌ای بررسی کنیم تا هرچه بیشتر، درمانگران و محققان را با این مفاهیم آشنا سازیم.

در دومین هدف، شاخص‌های تکرارپذیری مطلق را که شامل «خطای معیار اندازه‌گیری»، «حداقل تغییر قابل تشخیص»، «ضریب واریانس» و «درصد حداقل تغییر قابل تشخیص» هستند، با استفاده از منابع داده‌ها و معادلات مختلف مطالعه می‌کنیم. در این راستا، تأثیر منابع داده‌ها و معادلات مختلف گروه تحت مطالعه و فعالیت ثانویه را بر هر یک از شاخص‌های ذکر شده بررسی می‌کنیم.

سومین هدف تحقیق حاضر، کمی کردن فعالیت ثانویه ذهنی «شمردن معکوس» حین آزمون TUG cognitive است و در هر یک از اهداف یاد شده، به جنبه‌های مختلفی می‌پردازیم که در مقالات به آن توجه شده است.

در TUG cognitive، فرد به‌طور هم‌زمان، فعالیت ذهنی همچون شمردن اعداد را به‌صورت معکوس و در TUG manual، فرد به‌طور هم‌زمان، فعالیت حرکتی همچون حمل لیوان پر از آب را انجام می‌دهد (۱۱).

از ویژگی‌های سایکومتریک هر آزمون، تکرارپذیری^۱ آن است و به این دلیل اهمیت ویژه‌ای دارد که مبنای تصمیم‌گیری‌های بالینی است. به دلیل وجود خطا در تکرار هر اندازه‌گیری، همواره مقداری تفاوت و تغییر وجود دارد. این خطاها علل مختلفی دارند؛ اما به‌طور کلی، به دو دسته خطاهای سیستمی و تصادفی تقسیم می‌شوند (۵). تکرارپذیری روش علمی ارزیابی خطاهای تصادفی است (۵). متأسفانه، درک و اندازه‌گیری آن آسان نیست (۱۳) و به دو دسته تکرارپذیری نسبی و مطلق تقسیم می‌شود (۱۴). تکرارپذیری نسبی به‌وسیله ضریب هم‌بستگی درون طبقه‌ای^۲ اندازه‌گیری می‌شود.

از مباحث ضروری درمانگران و محققان این است که در آزمون مجدد، فرد به دنبال درمان یا چه مقدار از نتایج به‌دست آمده ناشی از خطای تصادفی و چه مقدار نشان دهنده تغییر واقعی است. ضریب هم‌بستگی درون طبقه‌ای، بیانگر توانایی آزمون در افتراق طبقه افراد است و نه شاخصی که اندازه‌گیری سهم خطای درخور انتظار را در هر بار آزمون فرد نشان دهد. این شاخص تحت تأثیر خصوصیات جمعیت موضوع مطالعه است؛ یعنی اگر تفاوت بین افراد زیاد باشد، ضریب هم‌بستگی درون طبقه‌ای بزرگ خواهد شد. حتی اگر تفاوت بین هر بار آزمون افراد، بزرگ باشد و برعکس (۱۵).

برخلاف تکرارپذیری نسبی، تکرارپذیری مطلق دقت خود آزمون را بررسی می‌کند. تکرارپذیری مطلق به وسیله «خطای معیار اندازه‌گیری»^۳ و «حداقل تغییر قابل تشخیص»^۴ اندازه‌گیری می‌شود (۱۵، ۱۴).

زیاد بودن ضریب هم‌بستگی درون طبقه‌ای آزمون، الزاماً به معنای مناسب بودن آن نیست، اگر خطای معیار اندازه‌گیری نشان دهد آزمون مدنظر مناسب نیست (۱۳).

در بسیاری مطالعات فقط به تکرارپذیری نسبی آزمون توجه می‌شود. با این حال، تکرارپذیری نسبی به‌تنهایی، به محققان و درمانگران کمکی نمی‌کند. آن‌ها نمی‌توانند تصمیم بگیرند که چقدر تغییر در نمرات آزمون، بیانگر تغییری واقعی در وضعیت نمونه یا بیمار است. با دانستن شاخص‌های تکرارپذیری مطلق، همچون خطای معیار اندازه‌گیری و حداقل تغییر قابل تشخیص،

1. Reliability 2. Intraclass Correlation Coefficient (ICC) 3. Standard Error of Measurement (SEM)
4. Minimal Detectable Changes (MDC) 5. Coefficient of Variation (CV)

روش بررسی

روش بررسی به این صورت بود: ۱۱ سالمند ساکن اجتماع با میانگین سن ۶۹/۶۴ و انحراف معیار ۳/۵۸ و دامنه ۶۵ تا ۷۶ سال و ۲۰ جوان سالم ساکن اجتماع با میانگین سن ۲۸/۸۰ و انحراف معیار ۴/۱۵ و دامنه ۲۰ تا ۳۵ سال، پس از اخذ موافقت آگاهانه، وارد مطالعه شدند. به علاوه، برای دستیابی به هدف اول این مطالعه، همه ۳۱ نمونه شرکت کننده در این تحقیق، به عنوان گروه مستقل مجزا به نام «گروه مختلط» فرض شد و اطلاعات آن‌ها با دو گروه شرکت کننده دیگر مقایسه شد. همه نمونه‌ها داوطلبانی بودند که در مطالعه اصلی شرکت کرده بودند.

معیارهای ورود در گروه سالمندان عبارت بود از: داشتن ۶۵ سال و بیشتر؛ داشتن زندگی مستقل براساس پرسش‌نامه «کاتز»^۱ که فعالیت‌های روزمره پایه‌ای را اندازه‌گیری می‌کرد (۲۲)؛ وضعیت شناختی و حافظه مساوی و بالاتر از ۲۳ (۲۳) براساس نسخه فارسی معاینه مختصر وضعیت شناختی^۲ (۲۴)؛ توانایی انجام آزمون‌ها با وسیله یا بدون وسیله کمکی؛ داشتن توانایی انجام دستوره‌های ساده. همچنین، معیارهای ورود در گروه جوانان چنین بود: داشتن سن ۲۰ تا ۳۵ سال و مبتلا نبودن به هرگونه بیماری شناخته شده‌ای که نتایج آزمون‌ها را متأثر سازد.

در ابتدا، همه نمونه‌ها به طور کامل در جریان اهداف و مراحل مختلف تحقیق قرار می‌گرفتند و در صورت توافق با شرکت در تحقیق، از آن‌ها خواسته می‌شد فرم رضایت آگاهانه‌ای را مطالعه و امضا کنند که کمیته اخلاق دانشگاه علوم بهزیستی و توانبخشی آن را تأیید کرده بود. سپس، نمونه‌ها، پرسش‌نامه‌ای مشتمل بر اطلاعات پایه‌ای همچون سن، جنس، تعداد اقامت داروهای استفاده شده و سطح تحصیلات را پر می‌کردند. استفاده از وسایل کمکی همچون عصا، سمک و... هم در پرسش‌نامه قید می‌شد. قد و وزن افراد با استفاده از متر و ترازوی دیجیتالی اندازه‌گیری و سپس، شاخص توده بدنی هم محاسبه می‌شد.

وضعیت سطح تحصیلات فرد در پنج مقطع طبقه‌بندی شد: ۱. بی‌سواد، یعنی هیچ‌گونه آموزشی دریافت نکرده بودند؛ ۲. مقطع اول که ۱ تا ۵ سال آموزش دیده بودند؛ ۳. مقطع دوم که ۶ تا ۹ سال آموزش دیده بودند؛ ۴. مقطع سوم که ۱۰ تا ۱۲ سال آموزش دیده بودند؛ ۵. مقطع چهارم که بیش از ۱۲ سال آموزش دیده بودند (۲۶، ۲۵).

در این تحقیق، از سه آزمون TUG_{original} و TUG_{manual} و TUG_{cognitive} استفاده شد. نمونه‌ها در جلسات آزمون اولیه و آزمون مجدد، یکبار این آزمون‌ها را انجام می‌دادند. ترتیب انجام دادن آزمون‌ها در هر جلسه و بین جلسات آزمون اولیه و آزمون مجدد تصادفی و فاصله زمانی بین آزمون اولیه و آزمون مجدد، ۵ تا ۲ روز بود. قبل از انجام هر یک از آزمون‌ها، مراحل انجام آن‌ها به صورت شفاهی توضیح داده می‌شد و حین انجام آزمون‌ها، فرد هیچ‌گونه دستوری دریافت نمی‌کرد. همچنین، به منظور آشناساختن هر یک از نمونه‌ها، قبل از آزمون اصلی، یکبار آزمون‌های TUG_{original} و TUG_{cognitive} را انجام می‌دادند. ابزار به کار برده شده، شامل متر، ترازو، کرنومتر و یک صدلی بدون دسته با ارتفاع نشیمنگاه ۴۵ سانتی‌متر بود. مدت انجام آزمون برحسب ثانیه، از زمانی که به فرد گفته می‌شد «۱، ۲، برو» تا زمانی اندازه‌گیری می‌شد که فرد پس از انجام مراحل آزمون، به صورت عمودی روی صدلی می‌نشست.

در آزمون TUG_{cognitive}، از نمونه‌ها خواسته می‌شد با صدای بلند، هم‌زمان با آزمون اصلی و به طور معکوس، سه تا سه تا از یک عدد بین ۷۰ تا ۱۰۰ که به آن‌ها گفته می‌شد، بشمرند. صدای نمونه‌ها حین انجام آزمون با استفاده از دستگاه ضبط صدا ضبط می‌شد. بعد از آزمون، با گوش دادن به فایل‌های صوتی ضبط شده، تعداد اعداد بیان شده و نیز تعداد خطاها مشخص می‌شد. در آزمون TUG_{manual}، از نمونه‌ها خواسته می‌شد حین انجام آزمون، لیوانی پر از آب را حمل کنند.

داده‌ها به وسیله نسخه ۲۰ نرم‌افزار SPSS تجزیه و تحلیل شدند. در همه تجزیه و تحلیل‌های آماری، سطح معناداری ۵ درصد تعیین شد. از آزمون کلموگروف اسمیرنوف^۳ برای بررسی توزیع داده‌ها استفاده شد. در متغیرهای توزیع نرمال، برای تعیین تفاوت بین میانگین گروه‌ها از آزمون تی مستقل و در متغیرهای توزیع غیرنرمال، از آزمون Mann-Whitney U استفاده شد. به دلیل وجود تأثیر نمونه‌ها و تأثیر تکرار اندازه‌گیری (آزمون اولیه و آزمون مجدد)، تکرارپذیری نسبی با استفاده از مدل تصادفی دوگانه ضریب هم‌بستگی درون طبقه‌ای (ICC_{۱,۲}) تخمین زده شد که فلیتز و شروت^۴ (۲۷) و نیز مک‌گراو و وانگ^۵ (۲۸) توضیح داده بودند.

به منظور اعمال اختلاف نتایج بین آزمون اولیه و آزمون مجدد، در مخرج معادله واریانس ضریب هم‌بستگی درون طبقه‌ای از توافق مطلق^۶ استفاده شد. همچنین، به دلیل اینکه نمره هر

1. Katz 2. (MMSE) Mini-Mental State Examination 3. Kolmogrov-Smirnov 4. Shrout & Fleiss
5. Mc Grow & Wong 6. Absolute Agreement

جذر مجموع میانگین خطاها در جدول ANOVA استفاده می‌شود.

$$\text{SEM} = (E)\sqrt{MS} \quad (\text{معادله } ۳)$$

هاپکینز^۲ پیشنهاد کرده از خطای موجود در مدل دوگانه برای محاسبه انحراف معیار اندازه‌گیری استفاده شود؛ زیرا مدل یگانه، هر دو خطای تصادفی و سیستمی را با هم لحاظ می‌کند. به این ترتیب، نتایج، سوگیری بسیاری داشته و تفسیر آن‌ها به دلیل این بسیار مشکل خواهد بود که سهم هر یک ناشناخته می‌ماند (۲۹). وقتی آثار آزمون با هر بار انجام دادن بزرگ باشد، سوگیری مدل یگانه درخور توجه خواهد بود (۱۵).

در این تحقیق، انحراف معیار اندازه‌گیری با استفاده از معادلات ۱ و ۳ محاسبه شد.

«انحراف معیار اندازه‌گیری» برای محاسبه «حداقل تغییر قابل تشخیص» سودمند است (۲۰، ۱۷). در غالب منابع، «حداقل تغییر قابل تشخیص» با استفاده از معادله زیر محاسبه می‌شود:

$$\text{MDC} = \text{SEM} \times z \times \sqrt{2} \quad (\text{معادله } ۴)$$

در مواقعی که توزیع داده‌ها نرمال نبود، به جای z ، از مقادیر t در جدول t استفاده کردیم.

$$\text{MDC} = \text{SEM} \times t \times \sqrt{2} \quad (\text{معادله } ۵)$$

در معادله «حداقل تغییر قابل تشخیص» به منظور پیشگیری از ناسازگاری‌های احتمالی از $\sqrt{2}$ استفاده می‌شود. این موضوع به فاصله اطمینان وسیع‌تر و محتاطانه‌تری منجر می‌شود (۱۶، ۱۵). هانگ^۳ و همکارانش (۱۶) از معادله زیر برای محاسبه «حداقل تغییر قابل تشخیص» استفاده کردند:

$$\text{MDC} = z \times \sqrt{2} \times \text{SEM} \quad (\text{معادله } ۶)$$

برای مقایسه تکرارپذیری مطلق بین متغیرها، «ضریب واریانس» با استفاده از معادله زیر محاسبه شد:

$$\text{CV} = \text{SD} / \text{mean} \times 100 \quad (\text{معادله } ۷)$$

علاوه بر «ضریب واریانس»، «حداقل تغییر قابل تشخیص» هم می‌تواند به صورت درصدی بیان شود. «درصد حداقل تغییر قابل تشخیص» در نکات زیر استفاده می‌شود:

۱. برای تعیین تغییرات واقعی نسبی بعد از درمان یا بین اندازه‌گیری‌های مکرر در گذر زمان؛

۲. برای نمایان‌ساختن مقدار نسبی خطای تصادفی اندازه‌گیری.

آزمون، نتیجه یک‌بار اندازه‌گیری بود و نه میانگین چندین بار اندازه‌گیری، از مقدار منفرد^۱ در محاسبات استفاده شد. برای توصیف میزان تکرارپذیری داده‌ها، از تقسیم‌بندی Munro استفاده شد که ۰ تا ۰/۲۵ هم‌بستگی جزئی یا خیلی کم، ۰/۲۶ تا ۰/۴۹ هم‌بستگی کم، ۰/۵۰ تا ۰/۶۹ هم‌بستگی متوسط، ۰/۷۰ تا ۰/۸۹ هم‌بستگی بالا و ۰/۹۰ تا ۱/۰۰ هم‌بستگی خیلی بالا طبقه‌بندی شده است (۱۴).

قبل از محاسبه تکرارپذیری مطلق، به منظور بررسی وجود یا نبود سوگیری و خطای سیستمی، از آزمون تی زوج درخصوص داده‌های آزمون اولیه و آزمون مجدد استفاده شد. برای محاسبه تکرارپذیری مطلق، «انحراف معیار اندازه‌گیری» و «حداقل اختلاف قابل تشخیص» محاسبه شد.

«انحراف معیار اندازه‌گیری» با استفاده از دو فرمول محاسبه شد: ضریب هم‌بستگی درون طبقه‌ای و بدون استفاده از ضریب هم‌بستگی درون طبقه‌ای.

از لحاظ تئوری محاسبه مستقیم انحراف معیار اندازه‌گیری مستلزم محاسبه انحراف معیار تعداد زیادی اندازه‌گیری از فردی است که این فرآیند در عمل قابل انجام نیست و در عمل انحراف معیار اندازه‌گیری تخمین زده می‌شود. اغلب منابع، از معادله زیر برای محاسبه انحراف معیار اندازه‌گیری استفاده می‌کنند.

$$\text{SEM} = \text{SD} \sqrt{1 - \text{ICC}} \quad (\text{معادله } ۱)$$

در معادله مذکور، SD انحراف معیار مجموع مقادیر آزمون اولیه و آزمون مجدد و ICC ضریب تکرارپذیری است (۱۵). آشکال مختلف ضریب هم‌بستگی درون طبقه‌ای، به‌خصوص در صورت وجود خطای سیستمی، به انحراف معیار اندازه‌گیری‌های متفاوت منجر خواهد شد. به‌منظور پرهیز از چنین مشکلاتی، انحراف معیار اندازه‌گیری از روش دوم محاسبه می‌شود:

$$\text{SEM} = (w)\sqrt{MS} \quad (\text{معادله } ۲)$$

به دلیل استقلال این شکل از انحراف معیار اندازه‌گیری از ضریب هم‌بستگی درون طبقه‌ای، استفاده از آن بیشتر در تفسیر مقادیر انحراف معیار اندازه‌گیری مطالعات مختلف مناسب است. محققان و درمانگران باید آگاه باشند که میانگین مجموع مربعات منتج از مدل‌های یگانه و دوگانه با هم متفاوت خواهند بود. استفاده از مدل یگانه، نیازمند استفاده از جذر میانگین مربعات تفاوت واریانس درون فردی است؛ درحالی‌که در مدل دوگانه، از

«درصد حداقل تغییر قابل تشخیص» به صورت زیر محاسبه می شود:

$$\text{MDC}\% = \text{MDC} / \text{mean} \times 100 \quad (17) \quad (\text{معادله ۸})$$

«درصد حداقل تغییر قابل تشخیص» کمتر از ۳۰ درصد پذیرفتنی است و درصد حداقل تغییر قابل تشخیص زیر ۱۰ درصد، عالی قلمداد می شود (۱۶).

یافته ها

جدول ۱ مشخصات دموگرافیک نمونه را نشان می دهد. در گروه سالمندان، فقط تعداد خطا حین آزمون مجدد TUG_{cognitive} توزیع نرمال نبود (P value=۰/۰۳). در گروه جوانان سالم هم فقط تعداد خطا حین آزمون اولیه TUG_{cognitive} توزیع نرمال نبود (P value=۰/۰۱). در گروه مختلط هم فقط تعداد خطا حین آزمون اولیه و مجدد TUG_{cognitive} توزیع نرمال نداشتند، به ترتیب (P value=۰/۰۲) و (P value=۰/۰۱).

جدول ۱. مشخصات دموگرافیک نمونه.

متغیر	سالمندان	جوانان سالم
جنسیت (درصد) مرد (زن)	۷۲/۷ (۲۷/۳)	۶۵ (۳۵)
سن (سال) (میانگین ± انحراف معیار)	۳/۵۸ ± ۶۹/۶۴	۴/۱۵ ± ۲۸/۸۰
وزن (کیلوگرم) (میانگین ± انحراف معیار)	۱۲/۰۵ ± ۶۹/۰۵	۱۱/۸۰ ± ۶۹/۷۲
قد (متر) (میانگین ± انحراف معیار)	۰/۰۹ ± ۱/۶۳	۰/۱۰ ± ۱/۷۰
شاخص توده جسمی (میانگین ± انحراف معیار)	۳/۶۴ ± ۲۵/۹۱	۲/۷۱ ± ۲۳/۸۸
تعداد داروی مصرفی (میانگین ± انحراف معیار)	۲/۱۱ ± ۲/۶۴	۰/۷۶ ± ۰/۴۵
سطح تحصیلات (میانگین ± انحراف معیار)	۰/۶۹ ± ۳/۵۵	۰/۶۰ ± ۳/۶۰

جدول ۲ مقادیر میانگین و انحراف معیار و دامنه آزمون های اولیه و مجدد را در گروه های تحت مطالعه نشان می دهد.

جدول ۲. آمار توصیفی آزمون های تحقیق در سه گروه تحت مطالعه.

متغیر	گروه سالمندان		گروه جوانان		گروه مختلط	
	آزمون اولیه	آزمون مجدد	آزمون اولیه	آزمون مجدد	آزمون اولیه	آزمون مجدد
TUG _{original} میانگین	۱۰/۰۸	۹/۹۵	۷/۶۶	۷/۶۸	۸/۵۲	۸/۴۸
انحراف معیار	۲/۸۷	۳/۰۱	۱/۱۳	۱/۳۴	۲/۲۲	۲/۳۲
دامنه	۱۵۳۴-۶/۱۰	۱۷/۱۳-۶/۵۰	۹/۴۰-۵/۶۶	۱۰/۲۸-۵/۳۴	۱۵-۵/۶۶	۱۷-۵/۳۴
TUG _{manual} میانگین	۱۲/۳۱	۱۱/۴۴	۸/۶۳	۸/۸۵	۹/۹۴	۹/۷۸
انحراف معیار	۳/۴۴	۲/۶۹	۱/۳۹	۱/۴۵	۲/۸۹	۲/۳۱
دامنه	۱۷/۴۴-۷/۴۷	۱۷/۱۶-۸/۳۱	۱۱/۷۲-۶/۲۱	۱۱/۱۵-۶/۳۵	۱۷-۶/۲۱	۱۷-۶/۳۵
TUG _{cognitive} میانگین	۱۴/۸۵	۱۴/۹۹	۱۰/۰۷	۹/۳۰	۱۱/۷۶	۱۱/۳۲
انحراف معیار	۵/۳۰	۴/۶۲	۲/۳۰	۱/۵۳	۴/۲۵	۴/۰۳
دامنه	۲۴/۳۵-۷/۶۱	۲۲/۸۱-۸/۶۰	۱۴/۳۸-۵/۹۷	۱۱/۸۷-۶/۲۱	۲۴-۵/۹۷	۲۲-۶/۲۱
طول زنجیره اعداد میانگین	۵/۵۵	۵/۸۲	۴/۱۰	۴/۲۵	۴/۶۱	۴/۸۱
انحراف معیار	۱/۹۲	۱/۷۲	۱/۲۹	۰/۹۷	۱/۶۷	۱/۴۷
دامنه	۳-۹	۳-۸	۲-۶	۲-۶	۲-۹	۲-۸
تعداد خطا میانگین	۰/۶۴	۰/۴۵	۰/۶۵	۰/۶۰	۰/۶۵	۰/۵۵
انحراف معیار	۰/۹۲	۰/۸۲	۰/۹۳	۰/۷۵	۰/۹۲	۰/۷۷
دامنه	۰-۳	۰-۲	۰-۳	۰-۳	۰-۳	۰-۳

نتایج جدول ۲ نشان می‌دهد در هر دو گروه سالمندان و جوانان، فعالیت ثانویه باعث افزایش زمان انجام آزمون می‌شود و تأثیر فعالیت ثانویه ذهنی، بیش از فعالیت ثانویه حرکتی است. همچنین، براساس نتایج جدول ۲، «طول زنجیره اعداد» بیان شده در گروه سالمندان، بیش از گروه جوانان است. این موضوع،

به دلیل میانگین زمان طولانی‌تر انجام آزمون سالمندان در مقایسه با جوانان است. آزمون تی مستقل و Mann-Whitney U نشان داد سالمندان و جوانان در همه متغیرها به جز تعداد خطا در آزمون اولیه و مجدد TUG cognitive اختلاف آماری معناداری داشتند (جدول ۳).

جدول ۳. نتایج آزمون تی مستقل برابری میانگین‌ها.

متغیر	معناداری (دو طرفه)
آزمون اولیه TUG original	۰/۰۰۲
آزمون مجدد TUG original	۰/۰۰۷
آزمون اولیه TUG manual	۰/۰۰۰
آزمون مجدد TUG Manual	۰/۰۰۱
آزمون اولیه TUG Cognitive	۰/۰۰۱
آزمون مجدد TUG cognitive	۰/۰۰۰
طول زنجیره اعداد در آزمون اولیه TUG cognitive	۰/۰۱۸
طول زنجیره اعداد در آزمون مجدد TUG cognitive	۰/۰۰۳
تعداد خطا در آزمون اولیه TUG cognitive*	۰/۹۵
تعداد خطا در آزمون مجدد TUG cognitive*	۰/۴۸

* با استفاده از آزمون Mann-Whitney U محاسبه شده است.

بین میانگین مقادیر آزمون اولیه و آزمون مجدد در همه گروه‌ها، اختلاف آماری معناداری وجود نداشت. این موضوع، نشان دهنده وجود نداشتن خطاهای سیستمی همچون خطای ناشی از یادگیری یا خستگی بود (جدول ۴).

جدول ۴. نتایج آزمون تی زوج.

متغیر	سالمندان	جوانان
TUG cognitive	۰/۸۸	۰/۸۶
TUG string no	۰/۵۷	۰/۵۲
TUG error no	۰/۳۴	۰/۸۱

با توجه به نتایج جدول ۳، اگر همه نمونه‌های موضوع مطالعه، اعم از سالمند و جوان را در گروهی به‌عنوان «گروه مختلط» مطالعه کنیم، این گروه تفاوت‌های بین‌فردی بسیاری خواهد داشت. با این فرض، ضریب همبستگی درون طبقه‌ای برای همه گروه‌ها محاسبه شد (جدول ۵).

جدول ۵. نتایج ضریب همبستگی درون طبقه‌ای و انحراف معیار اندازه‌گیری.

گروه	متغیر	ضریب همبستگی درون طبقه‌ای	انحراف معیار اندازه‌گیری با استفاده از داده‌های آزمون اولیه (SD $\sqrt{1-ICC}$)	انحراف معیار اندازه‌گیری با استفاده از داده‌های مختلط (SD $\sqrt{1-ICC}$)	انحراف معیار اندازه‌گیری (E \sqrt{MS})
سالمندان	TUG original	۰/۸۹	۰/۹۷	۰/۹۵	۱/۰۴
	TUG manual	۰/۶۴	۲/۰۸	۱/۸۳	۱/۸۵
	TUG cognitive	۰/۸۲	۲/۲۵	۲/۰۶	۲/۱۹
	طول زنجیره اعداد	۰/۶۵	۱/۱۳	۱/۰۵	۱/۱۰
جوانان	تعداد خطا	۰/۷۶	۰/۴۵	۰/۴۲	۰/۴۳
	TUG original	۰/۶۱	۰/۷۱	۰/۷۲	۰/۷۹

گروه	متغیر	ضریب هم‌بستگی درون طبقه‌ای	انحراف معیار اندازه‌گیری با استفاده از داده‌های آزمون اولیه ($SD\sqrt{1 - ICC}$)	انحراف معیار اندازه‌گیری با استفاده از داده‌های مختلط ($SD\sqrt{1 - ICC}$)	انحراف معیار اندازه‌گیری ($E\sqrt{MS}$)
میانگین	TUG manual	۰/۵۷	۰/۹۲	۰/۹۲	۰/۹۴
	TUG cognitive	۰/۵۰	۱/۶۲	۱/۳۹	۱/۳۴
	طول زنجیره اعداد	۰/۵۹	۰/۸۳	۰/۷۲	۰/۷۴
	تعداد خطا	۰/۳۹	۰/۷۳	۰/۶۵	۰/۶۷
	TUG original	۰/۸۶	۰/۸۳	۰/۸۴	۰/۸۷
	TUG manual	۰/۷۴	۱/۴۷	۱/۳۲	۱/۳۶
	TUG cognitive	۰/۸۳	۱/۷۶	۱/۷۰	۱/۶۸
	طول زنجیره اعداد	۰/۷۰	۰/۹۲	۰/۸۶	۰/۸۶
	تعداد خطا	۰/۵۲	۰/۶۴	۰/۵۸	۱/۲۱

جدول ۵ همچنین، مقادیر تخمینی «انحراف معیار اندازه‌گیری» را با استفاده از منابع مختلف داده‌ای، یعنی داده‌های آزمون اولیه یا داده‌های مختلط (میانگین مجموع داده‌های آزمون اولیه و آزمون مجدد) و استفاده از معادلات مختلف نشان می‌دهد.

جدول ۶ حداقل تغییر قابل تشخیص تخمینی با استفاده از منابع مختلف داده‌ها و معادلات مختلف.

گروه	متغیر	حداقل تغییر قابل تشخیص با استفاده از داده‌های آزمون ($SEM \times Z \times \sqrt{2}$)	حداقل تغییر قابل تشخیص با استفاده از معادله مختلط ($SEM \times Z \times \sqrt{2}$)	حداقل تغییر قابل تشخیص با استفاده از معادله (\sqrt{MSE})	حداقل تغییر قابل تشخیص با استفاده از معادله ($Z \times (\sqrt{2} \times SEM)$)
میانگین	TUG original	۱/۹۱	۲/۶۴	۲/۸۸	۲/۷۱
	TUG manual	۴/۰۷	۵/۰۶	۵/۱۳	۳/۷۵
	TUG cognitive	۴/۴۲	۵/۷۱	۶/۰۷	۳/۹۸
	طول زنجیره اعداد	۲/۲۲	۲/۹۲	۳/۰۵	۲/۸۵
	تعداد خطا	۰/۸۸	۱/۱۶	۱/۱۹	۱/۸۰
	TUG original	۱/۳۹	۲/۱۱	۲/۱۹	۲/۴۲
	TUG manual	۱/۷۹	۲/۵۶	۲/۶۱	۲/۶۶
	TUG cognitive	۳/۱۸	۳/۸۶	۳/۷۱	۳/۲۷
	طول زنجیره اعداد	۱/۶۲	۲/۰۱	۲/۰۵	۲/۳۶
	تعداد خطا	۱/۴۳	۱/۸۱	۱/۸۶	۲/۴۲
میانگین	TUG original	۲/۳۰	۲/۳۴	۲/۴۱	۲/۵۵
	TUG manual	۴/۰۹	۳/۶۷	۳/۸۷	۳/۱۹
	TUG cognitive	۴/۸۶	۴/۷۱	۴/۶۶	۳/۶۱
	طول زنجیره اعداد	۲/۵۳	۲/۳۷	۲/۳۸	۲/۵۶
	تعداد خطا	۱/۷۶	۱/۶۱	۳/۳۵	۲/۱۱

جدول ۷ اشکال مختلف «ضریب پراکنش» و «درصد حداقل تغییر قابل تشخیص» را برای مقایسه تکرارپذیری مطلق متغیرها نشان می‌دهد.

جدول ۷. مقادیر ضریب واریانس و درصد تغییر قابل تشخیص با استفاده از منابع داده‌ها و معادلات مختلف.

گروه	متغیر	ضریب پراکنش با استفاده از ضریب پراکنش با استفاده از داده‌های اولیه و معادله		درصد تغییر قابل تشخیص با استفاده از داده‌های مختلط و معادله		درصد تغییر قابل تشخیص با استفاده از داده‌های مختلط و معادله		درصد تغییر قابل تشخیص با استفاده از داده‌های مختلط و معادله	
		$CV = \left(\frac{SD}{Mean}\right) \times 100$	$CV = \left(\frac{SD}{Mean}\right) \times 100$	$CV = \left(\frac{SD}{Mean}\right) \times 100$	$CV = \left(\frac{SD}{Mean}\right) \times 100$	$CV = \left(\frac{SD}{Mean}\right) \times 100$	$CV = \left(\frac{SD}{Mean}\right) \times 100$	$CV = \left(\frac{SD}{Mean}\right) \times 100$	$CV = \left(\frac{SD}{Mean}\right) \times 100$
سالمندان	TUG original	۲۸/۴۹	۲۸/۷۰	۲۶/۱۹	۲۶/۳۸	۲۸/۷۹	۲۷/۰۲		
	TUG manual	۲۷/۹۶	۲۵/۶۴	۴۶/۴۹	۴۲/۶۴	۴۳/۱۸	۳۱/۵۵		
	TUG cognitive	۳۵/۶۹	۳۲/۵۳	۴۱/۹۷	۳۸/۲۵	۴۰/۶۹	۲۶/۶۶		
	طول زنجیره اعداد	۳۴/۵۶	۳۱/۳۸	۵۶/۶۷	۵۱/۴۶	۵۳/۶۶	۵۰/۱۱		
جوانان	تعداد خطا	۱۴۵/۲۷	۱۵۷/۲۸	۱۹۷/۲۶	۲۱۳/۵۷	۲۱۸/۵۱	۳۲۹/۴۴		
	TUG original	۱۴/۷۰	۱۵/۹۲	۲۵/۴۳	۲۷/۵۷	۲۸/۵۵	۳۱/۵۶		
	TUG manual	۱۶/۱۵	۱۶/۱۰	۲۹/۳۶	۲۹/۲۶	۲۹/۸۰	۳۰/۴۶		
	TUG cognitive	۲۲/۸۷	۲۰/۳۱	۴۴/۸۳	۳۹/۸۲	۳۸/۳۵	۳۳/۷۶		
بچه‌ها	طول زنجیره اعداد	۳۱/۵۵	۲۷/۰۶	۵۶/۰۰	۴۷/۰۲	۴۹/۱۳	۵۶/۴۷		
	تعداد خطا	۱۴۳/۵۸	۱۳۴/۰۵	۳۱۰/۸۵	۲۹۰/۲۰	۲۷۹/۱۴	۳۵۷/۷۳		
	TUG original	۲۶/۰۷	۲۶/۵۰	۲۷/۰۴	۲۷/۴۸	۲۸/۳۶	۲۹/۹۳		
	TUG manual	۲۹/۱۲	۲۶/۳۵	۴۱/۱۵	۳۷/۲۴	۳۸/۲۵	۳۲/۳۶		
مختلط	TUG cognitive	۳۶/۱۸	۳۵/۶۷	۴۱/۳۵	۴۰/۷۷	۴۰/۳۵	۳۱/۲۹		
	طول زنجیره اعداد	۳۶/۱۴	۳۳/۱۶	۵۴/۸۶	۵۰/۳۴	۵۰/۶۱	۵۴/۴۳		
	تعداد خطا	۱۴۴/۷۷	۱۴۰/۵۶	۲۷۲/۲۵	۲۶۹/۹۳	۵۰۲/۰۱	۳۵۴/۰۸		

بحث

ضریب هم‌بستگی درون طبقه‌ای بیانگر توانایی آزمون در افتراق بین افراد است و نه شاخص اندازه‌گیری خطای مورد انتظار حاصل از اندازه‌گیری‌های مکرر که برای درمانگران و... ضروری است (۱۵).

نتایج جدول ۵ نشان می‌دهد ضریب هم‌بستگی درون طبقه‌ای سالمندان، بیش از جوانان است. این مسئله ممکن است ناشی از ظرفیت حرکتی و استراتژی‌های با تنوع کمتر سالمندان نسبت به جوانان باشد. همین موضوع باعث می‌شود موقعیت و طبقه آن‌ها در گروه در هر بار تکرار آزمون، تغییر چندانی پیدا نکند؛ درحالی‌که این مسئله درباره جوانان مصداق نمی‌یابد.

همچنین براساس نتایج جدول ۵، اگر تفاوت‌های بین افراد گروه زیاد باشد، ضریب هم‌بستگی درون طبقه‌ای افزایش می‌یابد؛ برای مثال، در آزمون TUG manual، تفاوت‌های بین افراد در «گروه مختلط» نسبت به گروه سالمندان و گروه جوانان سالم بیشتر است. بر این اساس، ضریب هم‌بستگی درون طبقه‌ای در این گروه نیز در مقایسه با دو گروه دیگر، بیشتر است؛ یعنی

به ترتیب، ۰/۷۴ و ۰/۶۴ و ۰/۵۷ است. درباره متغیرهای TUG cognitive و طول زنجیره اعداد هم همین موضوع مصداق می‌یابد. درباره TUG original هم مقدار ضریب هم‌بستگی درون طبقه‌ای گروه مختلط به‌رغم غیرهمگن بودن افراد، بین دو گروه سالمندان و جوانان قرار دارد؛ اما به دلیل تفاوت اندک آن با مقدار مربوط در گروه جوانان، به نظر می‌رسد علت این موضوع حجم کم نمونه تحت مطالعه باشد. به‌عکس، درباره تعداد خطا در آزمون TUG cognitive، به دلیل نبود اختلاف معنادار بین افراد در «گروه مختلط» می‌توان این گروه را همگن قلمداد کرد و همان‌طور که در جدول ۵ مشهود است، ضریب هم‌بستگی درون طبقه‌ای گروه مزبور، بین دو گروه دیگر قرار دارد؛ یعنی به ترتیب، ۰/۵۲ و ۰/۷۶ و ۰/۳۹ است.

با توجه به نتایج، محققان و درمانگران باید در تفسیر ضریب هم‌بستگی درون طبقه‌ای محتاط باشند. ضریب هم‌بستگی درون طبقه‌ای، به طبقه‌بندی نسبی نمونه حین اندازه‌گیری‌های مکرر توجه می‌کند. اگر تفاوت بین افراد زیاد باشد، این موضوع ممکن است تفاوت اندازه‌گیری‌های مکرر را به‌طور کاذب بپوشاند (۱۵). همچنین، نتایج این مطالعه نشان داد افزودن فعالیت ثانویه به

برای مثال، با توجه به نتایج جدول ۶، درباره آزمون TUG_{original}، در سالمندان تفاوت بین آزمون مجدد با آزمون اولیه باید مساوی یا بیش از ۳/۳ ثانیه باشد تا این تفاوت تغییری واقعی و نه تغییری تصادفی قلمداد شود. در گروه جوانان، این مقدار ۲/۳۴ ثانیه است. درخصوص آزمون TUG_{cognitive}، هرگونه تغییر کمتر از ۷/۰۱ ثانیه در آزمون مجدد نسبت به آزمون اولیه در گروه سالمندان و ۳/۹۷ ثانیه در گروه جوانان را نباید تغییر واقعی قلمداد کرد.

مقادیر «حداقل تغییر قابل تشخیص» بزرگ‌تر سالمندان در مقایسه با جوانان در همه متغیرها به جز «تعداد خطا حین آزمون TUG_{cognitive}» نشان می‌دهد در متغیرهای مزبور، سالمندان در مقایسه با جوانان، نیازمند تغییرات بیشتری در آزمون مجدد هستند تا تغییر واقعی قلمداد شود.

از طرفی، با توجه به نتایج جدول ۶، به نظر می‌رسد استفاده از منابع مختلف داده‌های آزمون اولیه یا داده‌های مختلط یا استفاده از معادلات مختلف در محاسبه «حداقل تغییر قابل تشخیص»، نتایج کم‌وبیش یکسانی را فراهم ساخته است. این موضوع ناشی از نبود وجود خطای سیستمی و نیز اثر تعدیلی «انحراف معیار» و «ضریب هم‌بستگی درون طبقه‌ای» برهم است.

همچنین، با توجه به جدول ۶ به نظر می‌رسد غیرهمگن بودن نمونه موضوع مطالعه، باعث تعدیل «حداقل تغییر قابل تشخیص» می‌شود. با توجه به نتایج جدول ۷، مقادیر خطای تصادفی برخی متغیرهای تحقیق، بیش از ۳۰ درصد است؛ اما با اندکی اغماض، می‌توان نتیجه گرفت تقریباً همه متغیرها خطای تصادفی معقولی داشتند؛ به جز «تعداد خطا حین آزمون TUG_{cognitive}».

همچنین، نتایج جدول ۷ نشان می‌دهد در کل، سهم خطای تصادفی متغیرهای تحقیق شده در سالمندان، اندکی بیش از جوانان است و نیز انجام فعالیت ثانویه سهم خطای تصادفی، آزمون‌های موضوع مطالعه را افزایش می‌دهد؛ اما منابع داده‌ها و معادلات مختلف، نتایج کم‌وبیش مشابهی را فراهم می‌سازند.

تکرارپذیر بودن هر آزمونی مقدم بر معتربودن آن است؛ اما بی‌توجهی به دیگر ویژگی‌های سایکومتریک آزمون همچون اعتبار آزمون، از محدودیت‌های تحقیق حاضر است.

از دیگر محدودیت‌های این تحقیق، حجم کم نمونه مطالعه شده است؛ به همین دلیل، توصیه می‌شود تعمیم نتایج با احتیاط انجام شود. از آنجایی که تکرارپذیری هر آزمونی، بستگی به ویژگی‌های

فعالیت اصلی باعث کاهش ضریب هم‌بستگی درون طبقه‌ای می‌شود؛ اما بررسی میزان اثر فعالیت ثانویه حرکتی و ذهنی بر شاخص مزبور، نیازمند مطالعات بیشتر است.

نکته دیگری که باید به آن توجه شود، این است که در این تحقیق، از مدل تصادفی دوگانه ضریب هم‌بستگی درون طبقه‌ای (ICC_{2,1}) استفاده کردیم. تعدد مدل‌های کاربردی در محاسبه ضریب هم‌بستگی درون طبقه‌ای در مطالعات مختلف نشان می‌دهد مقایسه نتایج تحقیقات مختلف با هم مشکل بوده است و این مسئله نباید هنگام تفسیر و مقایسه نتایج مطالعات، از نظر محققان و درمانگران دور بماند.

برخلاف ضریب هم‌بستگی درون طبقه‌ای، «انحراف معیار اندازه گیری»، «حداقل تغییر قابل تشخیص»، «ضریب پراکنش» و «درصد حداقل تغییر قابل تشخیص»، دقت خودآزمون را کمی می‌کنند (۱۵). همان‌طور که در جدول ۵ مشهود است، نتایج «انحراف معیار اندازه‌گیری» با استفاده از معادلات ۱ و ۳ مشابه هستند؛ زیرا «انحراف معیار» و «ضریب هم‌بستگی درون طبقه‌ای» تمایل دارند اثر همدیگر را تعدیل کنند. با این حال، این تعدیل به‌طور کامل صورت نمی‌گیرد و محاسبه انحراف معیار اندازه‌گیری با استفاده از معادله ۱، تا حدودی از تفاوت‌های بین فردی متأثر می‌شود (۱۵).

همچنین، در جدول ۵ نشان داده شده که محاسبه انحراف معیار اندازه‌گیری با استفاده از داده‌های آزمون اولیه یا داده‌های مختلط (میانگین مجموع داده‌های آزمون اولیه و آزمون مجدد) به نتایج مشابه‌ای منجر شده است؛ زیرا بین مقادیر آزمون اولیه و آزمون مجدد این تحقیق، اختلاف آماری معناداری وجود نداشت. درخصوص تأثیر غیرهمگن بودن گروه بر نتایج «انحراف معیار اندازه‌گیری»، نتایج این تحقیق نشان می‌دهد نتایج در گروه مختلط، در مقایسه با گروه سالمندان و جوانان تعدیل شده است. نحوه تفسیر مقادیر «حداقل تغییر قابل تشخیص» بدین صورت است: در اندازه‌گیری‌های مجدد، فقط تغییراتی را که به اندازه «حداقل تغییر قابل تشخیص» با هم اختلاف دارند، باید تغییری واقعی در وضعیت فرد در راستای بهبودی یا نبود بهبودی قلمداد کرد. اگر بخواهیم دقیق‌تر بگوییم، همه افرادی که در آزمون مجدد حداقل بزرگ‌تر یا مساوی «حداقل تغییر قابل تشخیص» بودند، با آزمون اولیه تفاوت داشته‌اند. ۹۵ درصد آن‌ها بیانگر تفاوت معنادار و واقعی هستند (۱۷).

جوانان بوده است و با افزودن فعالیت ثانویه، این مقدار افزایش می‌یابد. نتایج «ضریب پراکنش» و «درصد حداقل تغییر قابل تشخیص» نشان داد سهم خطای تصادفی آزمون‌های TUG در گروه سالمندان بیش از جوانان بوده و با افزودن فعالیت ثانویه، سهم آن افزایش می‌یابد.

بررسی تکرارپذیری نسبی و مطلق و توجه به جنبه‌های مختلف هر یک از این شاخص‌ها، محققان و درمانگران را بیشتر با ابعاد آزمون آشنا ساخته است و می‌تواند به آن‌ها در اتخاذ تصمیمی بالینی یا تحقیقی مناسب‌تر کمک کند.

قدردانی

تیم تحقیقاتی از جناب آقای دکتر حسن آذرآباد، ریاست محترم پلی‌کلینیک شهرداری منطقه ۶ تهران و جناب آقای مهدی احمدی، مدیریت محترم مرکز درمانی مذکور تشکر فراوان می‌کند که فضا و امکانات لازم را در اختیار قرار دادند و در راستای انجام این تحقیق، نهایت همکاری‌های لازم را مبذول کردند.

نمونه مطالعه شده و وضعیت‌هایی دارد که در آن شرایط اندازه‌گیری انجام شده است، نتایج این تحقیق ممکن است برای سالمندان ناتوان یا مقیم آسایشگاه تعمیم‌پذیر نباشد. همچنین، فاصله زمانی به‌کار رفته در این تحقیق، ۲ تا ۵ روز بوده است که اگر این فاصله تغییر کند، ممکن است خطای اندازه‌گیری افزایش یابد. تأثیر فاصله زمانی بین آزمون اولیه و آزمون مجدد باید در تحقیقات آینده مدنظر قرار گیرد. مطالعه با یک‌سری طولانی‌تر آزمون‌های مجدد ممکن است تخمین پایدارتری از خطا را ارائه دهد. پیشنهاد می‌شود پژوهشگران دیگر، این فرایند را با استفاده از حجم نمونه‌های بزرگ‌تر انجام دهند و در سالمندان ناتوان و آسایشگاهی هم بررسی کنند.

نتیجه‌گیری

نتایج این تحقیق نشان داد غیرهمگن بودن گروه موضوع مطالعه، باعث افزایش ضریب هم‌بستگی درون طبقه‌ای می‌شود و سالمندان ضریب هم‌بستگی بیشتری در مقایسه با جوانان دارند. درخصوص شاخص‌های تکرارپذیری مطلق و منابع داده‌ای و معادلات مختلف، نتایج کم‌وبیش یکسانی را فراهم می‌سازند؛ اما مقادیر حداقل خطای قابل تشخیص در گروه سالمندان بیش از

REFERENCES

منابع

1. Tajvar M, Arab M, Montazeri A. Determinants of health-related quality of life in elderly in Tehran, Iran. *BMC Public Health*. 2008;8:323.
2. Najafi B, Aminian K, Loew F, Blanc Y, Robert PA. Measurement of stand-sit and sit-stand transitions using a miniature gyroscope and its application in fall risk evaluation in the elderly. *Biomedical Engineering, IEEE Transactions on*. 2002;49(8):843-51.
3. Noury N, Fleury A, Rumeau P, Bourke AK, Laighin GO, Rialle V, et al. Fall detection-principles and methods. *ConfProcIEEE Eng MedBiolSoc*. 2007;2007: pp:1663-6.
4. O'Sullivan M, Blake C, Cunningham C, Boyle G, Finucane C. Correlation of accelerometry with clinical balance tests in older fallers and non-fallers. *Age Ageing*. 2009;38(3):308-13.
5. Nordin E, Rosendahl E, Lundin-Olsson L. Timed "Up & Go" test: reliability in older people dependent in activities of daily living--focus on cognitive state. *Phys Ther*. 2006;86(5):646-55.
6. Bohannon RW. Reference values for the timed up and go test: a descriptive meta-analysis. *J Geriatr Phys Ther*. 2006;29(2):64-8.
7. Giné-Garriga M, Guerra M, Mari-Dell'Olmo M, Martin C, Unnithan VB. Sensitivity of a modified version of the "timed get up and go" test to predict fall risk in the elderly: a pilot study. *Arch Gerontol Geriatr*. 2009;49(1):e60-66.
8. Wall JC, Bell C, Campbell S, Davis J. The Timed Get-up-and-Go test revisited: measurement of the component tasks. *J Rehabil Res Dev*. 2000;37(1):109-13.
9. Whitney JC, Lord SR, Close JCT. Streamlining assessment and intervention in a falls clinic using the Timed Up and Go Test and Physiological Profile Assessments. *Age Ageing*. 2005;34(6):567-71.
10. Rankin JK, Woollacott MH, Shumway-Cook A, Brown LA. Cognitive influence on postural stability: a neuromuscular analysis in young and older adults. *J Gerontol A Biol Sci Med Sci*. 2000;55(3):M112-119.
11. Shumway-Cook A, Brauer S, Woollacott M. Predicting the probability for falls in community-dwelling older adults using the Timed Up & Go Test. *Phys Ther*. 2000;80(9):896-903.
12. Shumway-Cook A, Woollacott M. Attentional demands and postural control: the effect of sensory context. *J Gerontol A Biol Sci Med Sci*. 2000;55(1):M10-16.
13. Denegar CR, Ball DW. Assessing reliability and precision of measurement: an introduction to intraclass correlation and standard error of measurement. *J Sport Rehabil*. 1993;2(1):35-42.
14. Domholdt E. *Rehabilitation Research: Principles and Applications*. Elsevier Science Health Science Division; 2004.
15. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *J Strength Cond Res*. 2005;19(1):231-40.
16. Huang S-L, Hsieh C-L, Wu R-M, Tai C-H, Lin C-H, Lu W-S. Minimal detectable change of the timed "up & go" test and the dynamic gait index in people with Parkinson disease. *Phys Ther*. 2011;91(1):114-21.
17. Mazaheri M, Salavati M, Negahban H, Parnianpour M. Test-retest reliability of postural stability measures during quiet standing in patients with a history of nonspecific low back pain. *Biomedical Engineering: Applications, Basis and Communications*. 2010;22(03):255-62.
18. Steffen T, Seney M. Test-retest reliability and minimal detectable change on balance and ambulation tests, the 36-item short-form health survey, and the unified Parkinson disease rating scale in people with parkinsonism. *Phys Ther*. 2008;88(6):733-46.
19. Wang CY, Sheu CF, Protas EJ. Test-retest reliability and measurement errors of six mobility tests in the community-dwelling elderly. *Asian J Gerontol Geriatr*. 2009;4:8-13.
20. Ries JD, Echternach JL, Nof L, Gagnon Blodgett M. Test-retest reliability and minimal detectable change scores for the timed "up & go" test, the six-minute walk test, and gait speed in people with Alzheimer disease. *Phys Ther*. 2009;89(6):569-79.
21. Mangione KK, Craik RL, McCormick AA, Blevins HL, White MB, Sullivan-Marx EM, et al. Detectable changes in physical performance measures in elderly African Americans. *Phys Ther*. 2010;90(6):921-7.
22. Katz S, Downs TD, Cash HR, Grotz RC. Progress in development of the index of ADL. *Gerontologist*. 1970;10(1):20-30.
23. Woodford HJ, George J. Cognitive assessment in the elderly: a review of clinical methods. *QJM*. 2007;100(8):469-84.
24. Foroughan M, Jafari Z, Shirin Bayan P, Ghaem Magham Z, Rahgozar M. Validation of mini-mental state examination (MMSE) in the elderly population of Tehran *Advances in Cognitive Science* 2008;10(2):29-37.
25. Donyavi T, Naieni KH, Nedjat S, Vahdaninia M, Najafi M, Montazeri A. Socioeconomic status and mortality after acute myocardial infarction: a study from Iran. *Int J Equity Health*. 2011;10(1):9.
26. Montazeri A, Goshtasebi A, Vahdaninia M. Educational inequalities in self-reported health in a general Iranian population. *BMC Res Notes*. 2008;1:50.
27. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86(2):420-8.
28. McGraw KO, Wong SP. Forming inferences about some intraclass correlation coefficients. *Psychological Methods*. 1996;1(1):30.
29. Hopkins WG. Measures of reliability in sports medicine and science. *Sports medicine*. 2000;30(1):1-15.