

# اعتبارسنجی نسخه فارسی مقیاس رتبه‌بندی برنامه‌های کاربردی

## همراه (مارس) در حوزه بارداری

دکتر رسول نوری<sup>۱</sup>، راحله سموعی<sup>۲</sup>، دکتر فیروزه زارع فراشبندی<sup>۳</sup>، محبوبه شیروانیان دهکردی<sup>۴\*</sup>

۱. استادیار مدیریت اطلاعات سلامت، مرکز تحقیقات فناوری اطلاعات در امور سلامت، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران.
۲. استادیار علوم بهداشتی، مرکز تحقیقات عوامل اجتماعی مؤثر بر سلامت، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران.
۳. دانشیار کتابداری و اطلاع‌رسانی، مرکز تحقیقات فناوری اطلاعات در امور سلامت، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران.
۴. کارشناس ارشد کتابداری و اطلاع‌رسانی پزشکی، دانشکده مدیریت و اطلاع‌رسانی پزشکی، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران.

تاریخ دریافت: ۱۴۰۳/۰۳/۰۵ تاریخ پذیرش: ۱۴۰۳/۰۶/۱۰

### خلاصه

**مقدمه:** برنامه‌های کاربردی همراه در دوران بارداری می‌توانند به زنان باردار کمک کنند تا مراقبت‌های سلامتی را برای خود و نوزادشان ارتقاء دهند. مقیاس رتبه‌بندی برنامه‌های کاربردی سلامت همراه (MARS) یکی از ابزارهای معتبر به زبان انگلیسی برای سنجش کیفیت این برنامه‌هاست که نسخه فارسی آن در برنامه‌های حوزه بارداری در این مطالعه اعتبارسنجی شد.

**روش کار:** در این مطالعه مقطعی، مقیاس رتبه‌بندی برنامه‌های کاربردی سلامت همراه (MARS)، به فارسی ترجمه و به انگلیسی بازترجمه و با فرمت اولیه مقایسه شد. فرمت فارسی مقیاس برای ارزیابی ۴۴ برنامه کاربردی همراه فارسی در حوزه بارداری استفاده شد. روایی صوری، روایی محتوا و روایی ملاکی (همگرا و واگرا)، پایایی به روش آلفای کرونباخ و توافق ارزیابان سنجش شد. ۱۱ نفر از متخصصان حوزه انفورماتیک پزشکی و مدیریت اطلاعات سلامت در مرحله روایی محتوا و دو نفر کارشناس مامایی در مرحله ارزیابی برنامه‌های کاربردی حوزه مامایی همکاری داشتند. تجزیه و تحلیل با استفاده از نرم‌افزار آماری SPSS (نسخه ۲۵) و آزمون‌های ضریب همبستگی پیرسون و درون طبقه‌ای و آزمون تی دانشجویی انجام شد.

**یافته‌ها:** در مرحله سنجش روایی محتوا هر ۲۳ آیتم مقیاس تأیید و برخی ویرایش‌های مورد نیاز جهت حفظ وضوح و سادگی آن‌ها انجام شد. روایی همگرا و واگرا تأیید شد. ضریب همبستگی بین نمره زیرمقیاس‌ها (مشغولیت، عملکردی، زیبایی‌شناسی و اطلاعات) با نمره کل مارس برای ارزیابی اول از ۰/۸۵ تا ۰/۹۲ و برای ارزیابی دوم از ۰/۷۷ تا ۰/۹۲ بود. بین نظر ارزیابان در رتبه‌بندی کیفیت تمام زیرمقیاس‌ها در ۴۴ برنامه بارداری رابطه مثبت و معناداری وجود داشت. آلفای کرونباخ برای کل مقیاس در دو ارزیابی ۰/۹۳ و ۰/۹۱ محاسبه شد.

**نتیجه‌گیری:** با توجه به شاخص‌های محاسبه شده در این مطالعه، مقیاس رتبه‌بندی برنامه‌های کاربردی سلامت همراه، ابزاری روا و پایا در حوزه بارداری است. نتایج این مطالعه برای متخصصان، طراحان و ارزیابان برنامه‌های کاربردی همراه در حوزه بارداری قابل استفاده است و زنان باردار از بهره‌برداران اصلی آن هستند.

**کلمات کلیدی:** بارداری، برنامه‌های کاربردی، رتبه‌بندی، سلامت همراه

\* نویسنده مسئول مکاتبات: محبوبه شیروانیان دهکردی؛ دانشکده مدیریت و اطلاع‌رسانی پزشکی، دانشگاه علوم پزشکی اصفهان، اصفهان، ایران. تلفن:

۰۳۱-۳۷۹۲۵۱۵۱ پست الکترونیک: mahboubehshirvanian@gmail.com

## مقدمه

بارداری و دوره پس از زایمان، زمان تغییرات سریع پزشکی، اجتماعی و رفتاری برای زنان و خانواده‌هایشان است (۱). داشتن آگاهی نسبت به این تغییرات و دریافت اطلاعات لازم و متناسب با هر مرحله از بارداری می‌تواند رفتار سلامت مادران را بهبود بخشد و در نتیجه خطر ابتلاء به عوارض بارداری را کاهش دهد. اغلب زنان نیازمند حمایت اطلاعاتی و مشاوره‌های مداوم در این دوران می‌باشند. برنامه‌های کاربردی همراه، ابزار مفیدی برای دسترسی به این اطلاعات است. مطالعات گذشته نشان داده‌اند که استفاده از فناوری برای جستجوی اطلاعات بهداشتی در بین زنان باردار رایج است و زنان علاقه‌مند و راحت‌تر هستند که اطلاعات مرتبط با بارداری و زایمان را از طریق تلفن همراه خود دریافت کنند (۲). با توجه به تعداد برنامه‌های موجود برای زنان باردار، دانستن این نکته مهم است که چه برنامه‌هایی در دسترس است و چگونه باید از بین آن‌ها انتخاب کرد و زنان به دنبال چه اطلاعات و ویژگی‌هایی هستند (۳). اگرچه مطالعات متعددی نشان داده‌اند که رسانه‌های اجتماعی و سلامت همراه در ارتقای سلامت جسمانی مادر از جمله مدیریت وزن، کنترل دیابت بارداری و کنترل آسم، بهبود سلامت روان و همچنین افزایش آگاهی در مورد بارداری مؤثر بوده است (۴)؛ با این وجود برنامه‌های نامعتبر و نامناسبی نیز در دسترس هستند که ممکن است سلامت بارداری را به خطر اندازند. بررسی کیفیت برنامه‌های بارداری توسط براون و همکاران (۲۰۱۹) نشان داد که با وجود در دسترس بودن تعداد زیاد برنامه‌های بارداری، تعداد کمی از آن‌ها با کیفیت بالا هستند (۵). بنابراین لازم است کیفیت اطلاعات برنامه‌ها پیش از استفاده کاربران ارزیابی شود. علاوه بر کیفیت اطلاعات، دیگر ویژگی‌های کیفی برنامه‌ها مانند کاربردوستی، رضایت، قابلیت استفاده و زیبایی‌شناسی نیز باید در نظر گرفته شود (۶). مقیاس رتبه‌بندی برنامه‌های کاربردی همراه (مارس)<sup>۱</sup>، ابزاری است که همه این ویژگی‌ها را در نظر گرفته است و برنامه‌ها را

در چهار بُعد کیفیت عینی (مشغولیت<sup>۲</sup>، عملکردی<sup>۳</sup>، زیبایی‌شناسی<sup>۴</sup> و اطلاعات<sup>۵</sup>) و یک بُعد ذهنی<sup>۶</sup> ارزیابی می‌کند. نسخه اصلی ابزار، سطوح بالایی از پایایی در سنجش کیفیت برنامه‌های کاربردی مرتبط با سلامت ذهنی (موجود در اپل استور استرالیا) را نشان داده است (۷). مفید بودن استفاده از ابزار مارس برای ارزیابی کیفیت برنامه‌های دیگری مانند برنامه‌های مدیریت دیابت در چین (۸)، برنامه‌های روماتولوژی در آلمان (۹)، برنامه‌های بارداری در استرالیا (۵) و برنامه‌های مدیریت درد در اسپانیا (۱۰) نشان داده شده است. برزگری و همکاران (۲۰۲۲) (۱۱) مارس را به‌عنوان ابزاری قابل اعتماد و معتبر برای ارزیابی برنامه‌های مرتبط با اعتیاد معرفی کردند. در حال حاضر مارس یکی از پرکاربردترین ابزارهای موجود برای سنجش برنامه‌های کاربردی همراه به‌شمار می‌رود (۱۲) که امکان ارزیابی روی موضوعات و برنامه‌های مختلف را دارد. با توجه به اهمیت و حساسیت اطلاع‌رسانی‌ها در حوزه بارداری و عدم مشاهده ارزیابی استاندارد برنامه‌های موجود در فروشگاه‌های الکترونیکی ایران که به زبان فارسی در دسترس کاربران می‌باشد، مطالعه حاضر با هدف اعتبارسنجی نسخه فارسی ابزار مارس در برنامه‌های حوزه بارداری فارسی انجام شد.

## روش کار

فرآیند مطالعه حاضر که به‌صورت مقطعی انجام شد، شامل: ترجمه و بازترجمه ابزار مارس و سنجش روایی و پایایی آن در حوزه بارداری بود.

### ترجمه مقیاس رتبه‌بندی برنامه‌های کاربردی همراه

در آغاز نسخه اصلی ابزار مارس توسط دو مترجم از زبان انگلیسی به فارسی ترجمه شد. در این مرحله برای برخی از واژه‌ها معادل فارسی یافت نشد، بنابراین از املائی اصلی انگلیسی همراه با معنی ارجاعی‌شان در زبان فارسی (به‌صورت پاورقی) استفاده شد. پس از ترجمه انگلیسی به فارسی با همکاری ۲ نفر فارسی -

<sup>2</sup> Engagement

<sup>3</sup> Functionality

<sup>4</sup> Aesthetics

<sup>5</sup> Information

<sup>6</sup> Subjective dimension

<sup>1</sup> Mobile Application Rating Scale (MARS)

گرفت. در این مطالعه برنامه‌های مرتبط با بارداری و زایمان و مراقبت‌های مربوط به این دوران موردنظر بود، بنابراین یک نفر کارشناس مامایی و یک نفر کارشناس ارشد مامایی ارزیابی برنامه‌ها را به‌عهده داشتند. برنامه‌های منتخب با یک دستگاه تبلت مدل سامسونگ Sm-t595 و یک دستگاه گوشی هوشمند مدل Xiaomi Redmi 6 Pro بارگیری شدند. در این مرحله دو برنامه به‌علت باز نشدن بر روی یکی از دستگاه‌ها از مطالعه خارج شدند. ۴۴ برنامه توسط دو ارزیاب با استفاده از ابزار مارس به‌طور مستقل ارزیابی شدند. هر یک از برنامه‌ها در چهار بُعد کیفیت عینی (مشغولیت، عملکردی، زیبایی‌شناسی و اطلاعات) و یک بُعد کیفیت ذهنی رتبه‌بندی شدند. به‌طور کلی مقیاس رتبه‌بندی کیفیت برنامه‌های کاربردی شامل پنج زیرمقیاس<sup>۱</sup> با ۲۳ آیتم است: زیرمقیاس مشغولیت (۵ آیتم: سرگرمی، جذابیت، سفارشی‌سازی، تعامل و گروه هدف)، زیرمقیاس عملکردی (۴ آیتم: کارایی، سهولت استفاده، راهبری و طراحی ظاهری)، زیرمقیاس زیبایی‌شناسی (۳ آیتم: طرح‌بندی، گرافیک و جاذبه دیداری)، زیرمقیاس اطلاعات (۷ آیتم: صحت توصیف برنامه، اهداف، کیفیت اطلاعات، کمیت اطلاعات، اطلاعات دیداری، اعتبار و مبنای شواهد) و زیرمقیاس کیفیت ذهنی (۴ آیتم: توصیه برنامه به دیگران، دفعات استفاده در سال، پرداخت هزینه و امتیاز ستاره به برنامه). هر آیتم ۵ پاسخ احتمالی دارد که از ۱ (ناکافی) تا ۵ (عالی) نمره می‌گیرد. چند آیتم در زیرمقیاس اطلاعات (۳ آیتم: اهداف، کیفیت اطلاعات و کمیت اطلاعات)، یک گزینه «بدون پاسخ» دارند. این گزینه به ارزیابان این امکان را می‌دهد که زمانی که اطلاعات در برنامه موجود نیست، این گزینه را انتخاب کنند. پس از آن که رتبه‌بندی ۴۴ برنامه توسط ارزیابان انجام شد، پاسخ‌های عددی به آیتم‌ها در هر زیرمقیاس با یکدیگر جمع بسته شد، سپس میانگین گرفته شد (ماکزیمم نمره میانگین برای هر زیرمقیاس ۵ است). نمرات میانگین زیرمقیاس‌ها (مشغولیت، عملکردی، زیبایی‌شناسی و اطلاعات) با هم جمع و سپس تقسیم بر ۴ شد. حاصل این تقسیم، نمره‌ای بین

زبان مسلط به زبان انگلیسی (یک نفر کارشناس ارشد آموزش زبان انگلیسی) ترجمه‌های اولیه به یک ترجمه واحد ترکیب شد. پس از آن یک نفر دکترای زبان انگلیسی، بازترجمه از فارسی به انگلیسی را انجام داد. در این مرحله نسخه اولیه مارس فارسی ایجاد شد و پس از ویرایش‌های لازم نسخه نهایی آماده و در این مطالعه برای ارزیابی نمونه‌ای از برنامه‌های کاربردی همراه مرتبط با بارداری استفاده شد.

### شناسایی برنامه‌های کاربردی بارداری

ابتدا برای شناسایی برنامه‌های کاربردی سلامت همراه که به زبان فارسی و مرتبط با بارداری بودند، جستجو در فروشگاه اینترنتی کافه بازار که یکی از بزرگ‌ترین فروشگاه‌های اینترنتی ایران است، انجام شد. از اصطلاحات «بارداری»، «زایمان» و «حاملگی» برای جستجو استفاده شد. در این مرحله ۲۱۲ برنامه شناسایی شد (۲۵ فروردین ۱۴۰۰). سپس با ورود به بخش توصیف هر یک از برنامه‌ها، مشخصات آن‌ها استخراج شد. این مشخصات شامل نام توسعه‌دهنده، تعداد نصب فعال، رتبه ستاره، تعداد نظر کاربر، حجم برنامه، تاریخ بروزرسانی، هزینه و منابع برنامه بود. معیارهای ورود به مطالعه شامل: مرتبط بودن اطلاعات برنامه کاربردی با موضوع بارداری، به زبان فارسی بودن، قابلیت نصب بر روی سیستم اندروید، دسترسی رایگان، دریافت حداقل ۲/۵ امتیاز ستاره توسط کاربران و دارا بودن حداقل ۵۰ تعداد نظر کاربر در صفحه توصیف برنامه بود. برنامه‌ها بر اساس معیارهای ورود غربالگری شدند: ۱۵ برنامه با هزینه و ۴۸ برنامه به‌صورت خرید درون برنامه بودند که حذف شدند. از ۱۴۹ برنامه باقی‌مانده، ۱۰۳ برنامه که تعداد نظر کاربر کمتر از ۵۰ داشتند، حذف شدند. تمام ۴۶ برنامه باقی‌مانده، رتبه ستاره کاربر بالاتر از ۲/۵ دریافت کرده بودند؛ بنابراین همه معیارهای ورود را داشتند و برای رتبه‌بندی در نظر گرفته شدند. در نهایت ۴۴ برنامه به‌صورت سرشماری و با توجه به معیارهای ورود برای ارزیابی در نظر گرفته شدند.

### ارزیابی برنامه‌های کاربردی بارداری

انتخاب ارزیاب‌ها مطابق با مطالعات مشابه بر اساس تخصص آن‌ها در حوزه برنامه‌های مورد مطالعه صورت

<sup>1</sup> subscale

صفر تا ۵ است که نمره کل مارس یا میانگین کیفیت برنامه محسوب می‌شود. این محاسبات برای تک‌تک برنامه‌ها انجام شد. امتیازات محاسبه شده برای ۴۴ برنامه توسط هر یک از ارزیاب‌ها جداگانه در یک جدول اکسل جمع‌آوری شد و برای تحلیل آماری مورد استفاده قرار گرفت.

### سنجش روایی و پایایی

برای تعیین روایی صوری و محتوای مقیاس از افراد متخصص در حوزه انفورماتیک پزشکی (۵ نفر) و مدیریت اطلاعات سلامت (۶ نفر) خواسته شد جهت سنجش اعتبار مقیاس رتبه‌بندی کیفیت برنامه‌های کاربردی، هر یک از آیتم‌های مقیاس را بررسی کرده و از نظر ضروری بودن، ساده بودن، واضح بودن و مرتبط بودن با موضوع ارزیابی نمایند. نتایج ارزیابی ۱۱ پاسخ‌دهنده در یک صفحه اکسل جمع‌آوری شد، سپس با توجه به امتیازات محاسبه شده و بر اساس فرمول لاوشه CVI، CVR برای هر آیتم محاسبه شد. حداقل میزان قابل قبول CVR بر اساس تعداد متخصصین (۱۱ نفر)، طبق تکنیک لاوشه برای هر یک از آیتم‌های مقیاس ۰/۵۹ بود (۱۳).

فرمول ۱: محاسبه نسبت روایی محتوایی CVR

$$CVR = \frac{ne - N/2}{N/2}$$

در این فرمول ne تعداد افرادی است که وجود آیتم موردنظر را ضروری می‌دانند و N تعداد کل افراد متخصصی که آیتم مورد نظر را ارزیابی کرده‌اند.

فرمول ۲: محاسبه شاخص روایی محتوایی CVI

$$CVI = \frac{n}{N}$$

مقدار CVI برای هر آیتم کمتر از ۰/۷۰ به معنای عدم تأیید، ۰/۷۰ تا ۰/۷۹ مشروط و بیشتر از ۰/۷۹ تأیید آیتم می‌باشد (۱۴). برای تعیین روایی همگرا و واگرا در مقیاس رتبه‌بندی کیفیت برنامه‌های بارداری، میزان همبستگی هر یک از آیتم‌ها با زیرمقیاس مربوط به آن و همچنین سایر زیرمقیاس‌ها اندازه‌گیری شد. همسانی درونی مقیاس به روش آلفای کرونباخ و توافق بین ارزیابان به روش همبستگی پیرسون و درون طبقه‌ای اندازه‌گیری شد. بر اساس مطالعات قبلی (مارس ایتالیایی و مارس عربی)، پایایی به‌عنوان عالی

( $ICC_s \geq 0.70$ )، خوب ( $0.76 - 0.89$ )، متوسط ( $0.50 \leq ICC_s < 0.70$ ) و ضعیف ( $ICC_s < 0.50$ ) تفسیر شد (۶، ۱۵).

بر اساس نمرات میانگین ابعاد و نمرات کلی کیفیت برنامه‌ها که از رتبه‌بندی ۴۴ برنامه توسط ارزیابان حاصل شد، همسانی درونی مقیاس به روش آلفای کرونباخ و توافق بین ارزیابان به روش همبستگی پیرسون و درون طبقه‌ای اندازه‌گیری شد.

### تجزیه و تحلیل داده‌ها

آمار توصیفی با استفاده از شاخص‌های میانگین، انحراف معیار، فراوانی و درصد بیان شد. برای گزارش داده‌ها ضریب همبستگی پیرسون و ضریب همبستگی درون طبقه‌ای و برای تعیین همسانی درونی مقیاس از روش آلفای کرونباخ در نرم‌افزار SPSS (نسخه ۲۵) استفاده گردید و برای سنجش اختلاف بین میانگین‌ها در دو ارزیابی از آزمون تی دانشجویی برای گروه‌های مستقل استفاده شد.

در این مطالعه ضمن اخذ مجوز اخلاق (IR.MUL.RESEARCH.REC.1399.223) از معاونت تحقیقات و فناوری دانشگاه علوم پزشکی اصفهان، اصول محرمانگی اطلاعات مربوط به متخصصین مشارکت کننده در پژوهش رعایت شد.

### یافته‌ها

در مرحله سنجش روایی همه آیتم‌های مقیاس CVR قابل قبول داشتند. همچنین با توجه به امتیازات محاسبه شده برای CVI، همه آیتم‌ها مورد تأیید واقع شدند. با توجه به نظرات و پیشنهادات متخصصان، برخی از آیتم‌ها ویرایش شدند (آیتم ۱ «سرگرمی»، آیتم ۲ «جذابیت»، آیتم ۳ «سفارشی‌سازی» و آیتم ۴ «تعامل» از زیرمقیاس مشغولیت، آیتم ۷ «سهولت استفاده» از زیرمقیاس عملکردی و آیتم ۱۵ «کیفیت اطلاعات» از زیرمقیاس اطلاعات). برای مثال در مورد آیتم ۱، نظر متخصصان بر این بود که تفاوت گزینه‌ها مشخص نیست، بنابراین برای مشخص شدن میزان کیفیت، گزینه‌ها به شکل زیر تغییر یافتند:

مشغولیت، عملکردی، زیبایی‌شناسی، اطلاعات و کیفیت ذهنی و نمره کل مارس تفاوت معنادار نبود. همسانی درونی مقیاس همچنین از طریق تعیین رابطه بین نمره هر یک از زیرمقیاس‌ها با نمره کل مارس اندازه‌گیری شد. ضریب همبستگی بین نمره زیرمقیاس‌ها (مشغولیت، عملکردی، زیبایی‌شناسی و اطلاعات) با نمره کل مارس برای ارزیابی اول از ۰/۸۵ تا ۰/۹۲ و برای ارزیابی دوم از ۰/۷۷ تا ۰/۹۲ بود. زیرمقیاس مشغولیت نسبت به سایر زیرمقیاس‌ها (عملکردی، زیبایی‌شناسی و اطلاعات) بیشترین همبستگی با نمره کل مارس را داشت. همچنین ضریب همبستگی بین رتبه‌بندی ستاره ارزیابی با نمره کل مارس و ضریب همبستگی بین رتبه‌بندی ستاره کاربر (درفروشگاه برنامه) با نمره کل مارس اندازه‌گیری شد. بین نمره زیرمقیاس‌ها و همچنین رتبه ستاره ارزیابی با نمره کل مارس رابطه مثبت و معناداری وجود داشت، ولی بین رتبه ستاره کاربر با نمره کل مارس رابطه وجود نداشت (جدول ۱).

گزینه اول: «کسل کننده» به «کاملاً کسل کننده»  
گزینه دوم: «عمدتاً خسته‌کننده» به «تا حدی کسل کننده»  
گزینه سوم: «خوب، سرگرم‌کننده» به «نسبتاً سرگرم کننده»  
گزینه چهارم: «نسبتاً سرگرم کننده» به «خوب، سرگرم‌کننده»  
از بین ۴۴ برنامه که توسط دو ارزیاب با استفاده از مارس فارسی رتبه‌بندی شدند، بالاترین میانگین رتبه دو ارزیاب (۳/۷۶) مربوط به برنامه «نی نی پلاس: بارداری، زایمان، نوزاد» و پس از آن رتبه ۳/۶۷ و ۳/۵۵ به ترتیب مربوط به برنامه‌های «مادرشو: بارداری، زایمان و کودکیاری» و «مامانا پلاس (بارداری، زایمان، کودک)» بود. پایین‌ترین رتبه (۱/۸۴) مربوط به برنامه «باید و نبایدهای بارداری» بود.  
در مقایسه میانگین نمرات برنامه‌های رتبه‌بندی شده توسط دو ارزیاب از طریق آزمون تی دانشجویی مستقل، بین میانگین نمرات هیچ یک از زیرمقیاس‌های

جدول ۱- همسانی درونی پرسشنامه از طریق تعیین رابطه بین نمره هر یک از زیرمقیاس‌ها با نمره کل مارس

	مشغولیت	عملکردی	زیبایی‌شناسی	اطلاعات	رتبه ستاره ارزیاب	رتبه ستاره کاربر
ارزیاب ۱	۰/۹۲۲	۰/۸۲۴	۰/۹۰۳	۰/۸۵۶	۰/۸۷۲	-۰/۱۶۹
۱	<۰/۰۰۰۱	<۰/۰۰۰۱	<۰/۰۰۰۱	<۰/۰۰۰۱	<۰/۰۰۰۱	۰/۲۷۳
ارزیاب ۲	۰/۹۲۷	۰/۷۷۵	۰/۹۰۳	۰/۹۰۰	۰/۸۲۶	۰/۰۴۷
۲	<۰/۰۰۰۱	<۰/۰۰۰۱	<۰/۰۰۰۱	<۰/۰۰۰۱	<۰/۰۰۰۱	۰/۷۶۴

آن، نسبت به سایر زیرمقیاس‌ها تأیید شد. آلفای کرونباخ به دست آمده برای کل مقیاس در هر دو ارزیابی در حد عالی بود (۰/۹۳۹ و ۰/۹۱۲). آلفای کرونباخ برای ارزیابی اول در چهار زیرمقیاس (مشغولیت، عملکردی، زیبایی‌شناسی و اطلاعات) بین ۰/۷۵ تا ۰/۹۰ و برای ارزیابی دوم بین ۰/۴۵ تا ۰/۸۶ بود (جدول ۲).

طبق یافته‌های جدول ۱، بین نمره ابعاد مختلف مقیاس (مشغولیت، عملکردی، زیبایی‌شناسی، اطلاعات) و نمره کل ابعاد برای دو ارزیاب رابطه مثبت و معناداری وجود داشت ( $p < ۰/۰۰۰۱$ )، ولی بین رتبه ستاره کاربر و نمره کل رابطه وجود نداشت. روایی همگرا و واگرا با توجه به همبستگی بیشتر هر آیتام با نمره زیرمقیاس مربوط به

جدول ۲- همسانی درونی مقیاس رتبه‌بندی برنامه‌های کاربردی بارداری به روش آلفای کرونباخ

زیرمقیاس	ضریب آلفای کرونباخ	
	ارزیاب ۱	ارزیاب ۲
مشغولیت (n = ۵)	۰/۸۸۵	۰/۸۶۳
عملکردی (n = ۴)	۰/۹۰۴	۰/۴۵۲
زیبایی‌شناسی (n = ۳)	۰/۹۰۵	۰/۸۴۶
اطلاعات (n = ۷)	۰/۷۵۴	۰/۶۷۴
کل (n = ۱۹)	۰/۹۳۹	۰/۹۱۲

کیفیت تمام ابعاد برنامه‌های بارداری توافق وجود داشت. ICC در زیرمقیاس‌های عملکردی و اطلاعات متوسط، در زیرمقیاس مشغولیت عالی و برای نمره کل مارس در حد خوب گزارش شد (جدول ۳).

طبق یافته‌های جدول ۲، آلفای کرونباخ برای ارزیاب اول در همه ابعاد در گستره ۰/۷۵ تا ۰/۹۰ و برای کل مقیاس ۰/۹۳ به دست آمد و آلفای کرونباخ برای ارزیاب دوم در همه ابعاد در گستره ۰/۴۵ تا ۰/۸۶ و برای کل مقیاس ۰/۹۱ محاسبه شد. بین نظر ارزیابان در رتبه‌بندی

جدول ۳- توافق بین ارزیابان در ابعاد مختلف مقیاس رتبه‌بندی برنامه‌های بارداری به روش همبستگی درون طبقه‌ای

زیر مقیاس	همبستگی	کران بالا	کران پایین
مشغولیت	۰/۷۷	۰/۸۷	۰/۵۸
عملکردی	۰/۴۱	۰/۶۷	۰/۷۰
زیبایی‌شناسی	۰/۶۷	۰/۸۲	۰/۳۹
اطلاعات	۰/۵۷	۰/۷۶	۰/۲۱
کل	۰/۷۴	۰/۸۵	۰/۵۳
کیفیت ذهنی	۰/۶۷	۰/۸۲	۰/۴۰

برنامه، میزان همبستگی بین رتبه‌بندی کاربران واقعی و نمره کل مارس از ضعیف تا متوسط بوده است (۶). استویانوف و همکاران (۲۰۱۵) گزارش کردند که از بین ۵۰ برنامه سلامت روان که در تحقیقشان مورد ارزیابی قرار داده‌اند، تنها ۱۵ مورد رتبه‌بندی کاربر را دریافت کرده بودند که همبستگی این رتبه‌بندی (رتبه‌بندی ستاره کاربر در آی تونز) با نمره کل مارس در حد متوسط بوده است (۷). به گفته محققان ایتالیایی، قابل اعتماد بودن برنامه‌هایی با رتبه‌بندی کم ممکن است با بررسی‌های جعلی توسعه‌دهندگان برنامه به خطر بیافتد، این موضوع ممکن است تا حدی بیان کننده علت افزایش شدت همبستگی بین رتبه‌بندی کاربران و نمره کل مارس با افزایش خط برش تعداد رأی به برنامه باشد (۶). استویانوف وارد نکردن زیرمقیاس کیفیت ذهنی در محاسبه میانگین کلی کیفیت برنامه را به دلیل داشتن ماهیت ذهنی، باعث تقویت عینیت مارس به‌عنوان یک معیار اندازه‌گیری کیفیت برنامه دانسته است. با این وجود همبستگی بالای بین نمره کل کیفی مارس با رتبه‌بندی ستاره مارس نشان می‌دهد که کیفیت کلی برنامه توسط ارزیابان درک شده است و احتمالاً رتبه‌بندی ستاره مارس تحت تأثیر ۱۹ آیتم قبلی انجام شده است (۷). این موضوع همچنین می‌تواند یکی از دلایل تفاوت بین رتبه ستاره ارزیاب و رتبه ستاره کاربر باشد. با توجه به نتایج به‌دست آمده از

طبق یافته‌های جدول ۳، توافق بین ارزیابان با استفاده از روش همبستگی درون طبقه‌ای در کل ۰/۷۴ و در حد خوب بود. میزان توافق در بُعد عملکردی و اطلاعات متوسط و در بُعد مشغولیت عالی بود.

### بحث

یافته‌های حاصل از تعیین رابطه بین نمره هر یک از زیرمقیاس‌ها (مشغولیت، عملکردی، زیبایی‌شناسی، اطلاعات و کیفیت ذهنی) با نمره کل مارس نشان داد که رابطه مثبت و معناداری بین زیرمقیاس‌ها وجود دارد. رتبه‌بندی ستاره مارس (آیتم ۱۹) برای هر دو ارزیاب همبستگی بالایی با نمره کل مارس داشت، در حالی‌که بین نمره رتبه‌بندی ستاره کاربر با نمره کل مارس رابطه وجود نداشت. این یافته همسو با مطالعات خارجی بود. در مارس عربی گزارش شده که رتبه‌بندی ۵ ستاره از فروشگاه‌های برنامه با هر یک از زیرمقیاس‌ها و نمره کل مارس عربی به‌طور قابل توجهی نامرتب بوده است (۱۵). در مارس آلمانی نیز عنوان شده که هیچ ارتباطی بین رتبه‌بندی کاربران و نمره کل مارس آلمانی وجود نداشته است (۱۶). همچنین محققان ایتالیایی با بررسی همبستگی بین رتبه‌بندی ستاره مارس و نمره کل مارس با رتبه‌بندی ستاره کاربر دریافتند که همبستگی معناداری بین رتبه‌بندی کاربر و نمره کل مارس برای برنامه‌هایی با بیشتر از ۵ رأی وجود نداشته، اما با افزایش خط برش تعداد رأی به

الکترونیکی دارد (۱۹). در مجموع ضرایب همبستگی بین آیت‌ها با زیرمقیاس‌ها و زیرمقیاس‌ها با نمره کل مارس نشان داد که مقیاس از همسانی درونی خوبی برخوردار است و مقیاس رتبه‌بندی برنامه‌های کاربردی همراه می‌تواند به‌عنوان یک ابزار پایا برای ارزیابی برنامه‌های کاربردی در زمینه بارداری مورد استفاده قرار گیرد.

از محدودیت‌های مطالعه این بود که برنامه‌های کاربردی تحت سیستم IOS، برنامه‌های کاربردی در سایر فروشگاه‌های اندروید و همچنین برنامه‌های کاربردی با پرداخت هزینه، وارد مطالعه نشدند.

### نتیجه‌گیری

نتایج این مطالعه نشان داد که نسخه فارسی مقیاس رتبه‌بندی برنامه‌های کاربردی همراه، ابزاری قابل اعتماد برای رتبه‌بندی کیفیت برنامه‌های کاربردی مرتبط با بارداری می‌باشد. با این حال تحقیقات بیشتری برای ارزیابی ویژگی‌های روان‌سنجی آن در سایر حوزه‌های سلامت موردنیاز است. علاوه بر نسخه اصلی مارس که بیشتر برای کارشناسان و محققان سلامت همراه قابل استفاده است، نسخه کاربر مارس نیز در دسترس است و در صورت انطباق با زبان فارسی، برای کاربران نهایی برنامه‌های کاربردی همراه می‌تواند مفید و قابل استفاده باشد.

### تشکر و قدردانی

این مقاله حاصل بخشی از پایان‌نامه کارشناسی ارشد با کد علمی ۳۹۹۰۸۶ می‌باشد که با حمایت دانشگاه علوم پزشکی اصفهان انجام شد؛ بدین‌وسیله از حوزه معاونت پژوهشی دانشگاه علوم پزشکی اصفهان جهت حمایت از این مطالعه، تشکر و قدردانی می‌گردد.

### تضاد منافع

نویسندگان اظهار می‌دارند که تضاد منافی وجود ندارد.

### حمایت مالی

این مطالعه با حمایت مالی دانشگاه علوم پزشکی اصفهان انجام شده است.

مطالعه حاضر و مطالعات پیشین می‌توان نتیجه گرفت رتبه‌بندی ستاره کاربر به‌دلیل آنکه تحت تأثیر معیارهای پراکنده ذهنی کاربران صورت گرفته است، نمی‌تواند بیان‌کننده کیفیت واقعی برنامه، به‌خصوص در رابطه با کیفیت اطلاعات باشد، زیرا کاربران معمولاً اطلاعاتی را جستجو می‌کنند که دانش کمتری نسبت به آن دارند و امکان دارد تشخیص آن‌ها از صحت اطلاعات در حد کارشناسان آموزش دیده نباشد؛ بنابراین به سختی می‌توان به رتبه‌بندی ستاره کاربر اعتماد کرد. محاسبه روایی همگرا و واگرا در مقیاس رتبه‌بندی کیفیت برنامه‌های کاربردی بارداری نشان داد اکثر آیت‌ها با زیرمقیاس مربوطه خود همبستگی بیشتری نسبت به سایر زیرمقیاس‌ها داشتند. بیشترین میزان همبستگی آیت‌های زیرمقیاس مشغولیت با سایر مقیاس‌ها (عملکردی، زیبایی‌شناسی، اطلاعات و کیفیت ذهنی)، در آیت ۱ (سرگرمی) با زیرمقیاس زیبایی‌شناسی مشاهده شد. از این مشاهده می‌توان نتیجه گرفت که سطح جاذبه دیداری و گرافیک با کیفیت بالا ممکن است میزان مشغولیت را افزایش دهد. محققان آلمانی نیز از همبستگی بالا بین زیبایی‌شناسی و مشغولیت در مارس آلمانی به نتایج مشابهی در این زمینه دست یافته‌اند. آن‌ها احتمال می‌دهند که این دو زیرمقیاس عینی، سازه‌های نزدیک و مرتبط به هم را ارزیابی می‌کنند (۱۶). مارس مرتبط با برنامه‌های اعتیاد نیز در ایران همگرایی خوب را گزارش کرده است (۱۱). در مجموع می‌توان گفت برنامه‌هایی که از نظر جنبه‌های زیبایی‌شناسی و ویژگی‌های تعاملی انتظارات کاربر را برآورده می‌کنند، بیشتر استفاده می‌شوند (۱۷). از همبستگی بالای آیت ۷ (سهولت استفاده) با زیرمقیاس زیبایی‌شناسی می‌توان این نتیجه را گرفت که چیدمان و اندازه برای دکمه‌ها، آیکون‌ها و منوها و همچنین گرافیک مطلوب به‌کار رفته در برنامه‌های کاربردی می‌تواند در سهولت استفاده از این برنامه‌ها مؤثر باشد. کاربران معمولاً برنامه‌هایی را انتخاب می‌کنند که طراحی خوب و قابلیت استفاده آسان دارند (۱۸).

استنالت و همکار (۲۰۰۶) معتقدند طراحی و چیدمان، نقش مهمی در مشغولیت با سیستم‌های یادگیری

## ملاحظات اخلاقی

این مطالعه به تصویب کمیته اخلاق  
IR.MUI.RESEARCH.REC.1399.223  
دانشگاه علوم پزشکی اصفهان رسید.

## مشارکت نویسندگان

طراحی و اجرای این مطالعه با همکاری همه  
نویسندگان انجام شد.

## منابع

- Hussain T, Smith P, Yee LM. Mobile phone-based behavioral interventions in pregnancy to promote maternal and fetal health in high-income countries: systematic review. *JMIR mHealth and uHealth* 2020; 8(5):e15111.
- Osma J, Barrera AZ, Ramphos E. Are pregnant and postpartum women interested in health-related apps? Implications for the prevention of perinatal depression. *Cyberpsychology, Behavior, and Social Networking* 2016; 19(6):412-5.
- Frid G, Bogaert K, Chen KT. Mobile health apps for pregnant women: systematic search, evaluation, and analysis of features. *Journal of medical Internet research* 2021; 23(10):e25667.
- Chan KL, Chen M. Effects of social media and mobile health apps on pregnancy care: meta-analysis. *JMIR mHealth and uHealth* 2019; 7(1):e11836.
- Brown HM, Bucher T, Collins CE, Rollo ME. A review of pregnancy iPhone apps assessing their quality, inclusion of behaviour change techniques, and nutrition information. *Maternal & child nutrition* 2019; 15(3):e12768.
- Domnich A, Arata L, Amicizia D, Signori A, Patrick B, Stoyanov S, et al. Development and validation of the Italian version of the Mobile Application Rating Scale and its generalisability to apps targeting primary prevention. *BMC Medical Informatics and Decision Making* 2016; 16:1-10.
- Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR mHealth and uHealth* 2015; 3(1):e3422.
- Gong E, Zhang Z, Jin X, Liu Y, Zhong L, Wu Y, et al. Quality, functionality, and features of Chinese mobile apps for diabetes self-management: systematic search and evaluation of mobile apps. *JMIR mHealth and uHealth* 2020; 8(4):e14836.
- Knitz J, Tascilar K, Messner EM, Meyer M, Vossen D, Pulla A, et al. German mobile apps in rheumatology: review and analysis using the Mobile Application Rating Scale (MARS). *JMIR mHealth and uHealth* 2019; 7(8):e14991.
- Salazar A, de Sola H, Failde I, Moral-Munoz JA. Measuring the quality of mobile apps for the management of pain: systematic search and evaluation using the mobile app rating scale. *JMIR mHealth and uHealth* 2018; 6(10):e10718.
- Barzegari S, Sharifi Kia A, Bardus M, Stoyanov SR, GhaziSaedi M, Rafizadeh M. The Persian version of the Mobile Application Rating Scale (MARS-Fa): translation and validation study. *JMIR formative research* 2022; 6(12):e42225.
- Terhorst Y, Philippi P, Sander LB, Schultchen D, Paganini S, Bardus M, et al. Validation of the mobile application rating scale (MARS). *Plos one* 2020; 15(11):e0241480.
- Lawshe CH. A quantitative approach to content validity. *Personal Psychology* 1975; 28(4):563-575.
- Waltz CF. *Nursing research: Design, statistics, and computer analysis*. FA Davis. 1981.
- Bardus M, Awada N, Ghandour LA, Fares EJ, Gherbal T, Al-Zanati T, et al. The Arabic version of the Mobile App Rating Scale: development and validation study. *JMIR mHealth and uHealth* 2020; 8(3):e16956.
- Messner EM, Terhorst Y, Barke A, Baumeister H, Stoyanov S, Hides L, et al. The German version of the Mobile App Rating Scale (MARS-G): development and validation study. *JMIR mHealth and uHealth* 2020; 8(3):e14479.
- Lazard AJ, Brennen JS, Belina SP. App designs and interactive features to increase mHealth adoption: user expectation survey and experiment. *JMIR mHealth and uHealth* 2021; 9(11):e29815.
- Bardus M, van Beurden SB, Smith JR, Abraham C. A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. *International Journal of Behavioral Nutrition and Physical Activity* 2016; 13:1-9.
- Stenalt MH, Godsk M. The pleasure of e-learning: Towards aesthetic e-learning platforms. In *Proceedings of the 12th international conference of european university information systems* 2006:210-212. University of Tartu.





# Validation of the Persian Version of the Mobile Application Rating Scale (MARS) in the field of pregnancy

Rasool Nouri<sup>1</sup>, Rahele Samouei<sup>2</sup>, Firoozeh Zare-Farashbandi<sup>3</sup>, Mahboobeh Shirvanian-Dehkordi<sup>4\*</sup>

1. Assistant Professor of Health Information Management, Health Information Technology Research Center, Isfahan University of Medical Sciences, Isfahan, Iran.
2. Assistant Professor of Health Sciences, Social Determinants of Health Research Center, Isfahan University of Medical Sciences, Isfahan, Iran.
3. Associate Professor of Library and Information Sciences, Health Information Technology Research Center, Isfahan University of Medical Sciences, Isfahan, Iran.
4. M.Sc. in Medical Library and Information Science, School of Management and Medical Information Science, Isfahan University of Medical Sciences, Isfahan, Iran.

Received: May 25, 2024 Accepted: Aug 31, 2024

## Abstract

**Introduction:** Mobile applications during pregnancy can help pregnant women improve health care for themselves and their infants. In this regard, using valid and reliable evaluation tools is a good solution. Mobile application rating scale (MARS) is one of the valid tools in English to measure the quality of these programs, whose Persian version was validated in the field of pregnancy in this study.

**Methods:** The cross-sectional study, the Mobile App Rating Scale (MARS) was translated into Persian and back-translated into English and compared with the original format. In the next step, the Persian version of MARS was used to evaluate 44 Persian mobile applications in the field of pregnancy. Face validity, content validity (CVR, CVI) and criterion validity (convergent and divergent) were calculated by Cronbach's alpha method and evaluators' agreement. Eleven specialists in the field of medical informatics and health information management participated in the content validity stage and two midwifery experts participated in the evaluation stage of midwifery applications. Pearson and intra-class correlation coefficient and student t-test were used in SPSS software (version 25).

**Results:** In the content validity assessment stage, all 23 items were confirmed, and some edits needed to maintain the clarity and simplicity of the items was done. Convergent and divergent validity was confirmed. There was correlation coefficient between the score of the subscales (engagement, functionality, aesthetics, and information) with the total score of MARS from 0.85 to 0.92 for the first evaluator and from 0.77 to 0.92 for the second evaluator. There was a positive and significant relationship between the evaluators' opinion in ranking the quality of all subscales of 44 pregnancy apps. Cronbach's alpha for the whole scale was calculated to be 0.93 and 0.91 in two assessments.

**Conclusion:** According to the indicators calculated in this study, Mobile App Rating Scale (MARS) is a valid and reliable tool in the field of pregnancy. The results of this study can be used by specialists, designers and evaluators of mobile applications in the field of pregnancy and pregnant women are its main beneficiaries.

**Keywords:** Applications, Mobile health, Pregnancy, Rating

► Please cite this article as:

Nouri R, Samouei R, Zare-Farashbandi F, Shirvanian-Dehkordi M. Validation of the Persian Version of the Mobile Application Rating Scale (MARS) in the field of pregnancy. *Iran J Obstet Gynecol Infertil* 2024; 27(6):1-9. DOI: 10.22038/ijogi.2024.73111.5750