

کاربرد مدل‌های پیوندی برای تعیین میزان توافق داده‌های ترتیبی در بررسی عملکرد انترن‌های دندانپزشکی در تشخیص نیاز به درمان ارتودنسی

دکتر علیرضا اکبرزاده باغبان* - دکتر محمدرضا صفوی** - دکتر مهتاب‌نوری*** - دکتر لادن اسلامیان** - دکتر غلامرضا بابایی**** - دکتر انوشیروان کاظم‌نژاد**** - دکتر سقراط فقیه‌زاده****

*- استادیار گروه آموزشی آمار زیستی دانشگاه علوم پزشکی شهید بهشتی.

** - دانشیار گروه آموزشی ارتودنسی دانشکده دندانپزشکی دانشگاه علوم پزشکی شهید بهشتی.

*** - استادیار گروه آموزشی ارتودنسی دانشکده دندانپزشکی دانشگاه علوم پزشکی شهید بهشتی.

**** - استاد گروه آموزشی آمار زیستی دانشگاه تربیت مدرس.

چکیده

زمینه و هدف: یکی از مسائل مهم در اندازه‌گیری شاخصها، پایایی آزماینده‌هایی است که آن شاخصها را اندازه می‌گیرند. شاخص نیاز به درمان ارتودنسی نیز از این امر مستثنی نمی‌باشد. هدف از این مطالعه تعیین عملکرد انترن‌های دندانپزشکی در تشخیص نیاز به درمان ارتودنسی با استفاده از شاخص اصلاح شده سه سطحی IOTN در دو بار اندازه‌گیری با فاصله زمانی دو هفته‌ای می‌باشد. روش بررسی: در مطالعه حاضر که از نوع مقطعی تحلیلی می‌باشد، شش انترن دندانپزشکی به طور مستقل این شاخص را در دو زمان مجزا، با استفاده از خط کش استاندارد DHC، روی ۴۲ کست استاندارد مطالعه ارتودنسی اندازه‌گیری کردند. از ضریب کاپا و مدل‌های پیوند برای تحلیل داده‌ها استفاده شد.

یافته‌ها: ضریب کاپا برای اندازه‌گیری توافق کلی هریک از انترن‌ها در دو بار اندازه‌گیری، با کمترین مقدار ۰/۵۴، بیشترین مقدار ۰/۷۷ و میانگین ۰/۶۵ بیانگر پایایی مناسب انترن‌ها بود. ۳۳٪ از انترن‌هایی که در تفکیک سطوح ۱ و ۲ این شاخص مشکل داشتند، دارای مشکل قابل‌ذکری در تفکیک سطوح دو و سه نبودند. سایر انترن‌ها نیز دارای قابلیت تشخیص بسیار بالا در تفکیک همه سطوح بودند. کمترین مقدار قابلیت تشخیص برای سطوح یک و دو برابر ۰/۴۸ و بیشترین مقدار برای سطوح دو و سه برابر ۰/۹۹ به دست آمد. نتیجه‌گیری: از آنجایی که تفکیک سطوح پایین شاخص نیاز به درمان ارتودنسی برای انترن‌ها مشکل است، قبل از تحقیقات میدانی وسیع آموزش دقیق آنها توصیه می‌شود.

کلیدواژه‌ها: قابلیت تشخیص - مدل پیوند - کاپا - نیاز به درمان ارتودنسی

پذیرش مقاله: ۱۳۸۶/۴/۱۴

اصلاح نهایی: ۱۳۸۶/۱/۲۳

وصول مقاله: ۱۳۸۵/۹/۱۹

e.mail:akbarzad@sbmu.ac.ir

نویسنده مسئول: گروه آموزشی آمار زیستی دانشکده دندانپزشکی دانشگاه علوم پزشکی شهید بهشتی

مقدمه

ضرورت بررسی شاخصها در علوم پزشکی بر کسی پوشیده نیست و ارتودنسی به عنوان یکی از قدیمیترین شاخه‌های دندانپزشکی از این امر مستثنی نمی‌باشد. نیاز به درمان ارتودنسی در جامعه ما الگوی شناخته شده‌ای نداشته و برآورد دقیقی از میزان این نیاز در سطح جامعه در دست نیست. از طرفی هر چند هدف دندانپزشک در کلینیک درمان تک تک افراد است ولی به عنوان عضوی از

ارایه‌دهندگان خدمات بهداشتی، هدف نهایی وی باید درمان جامعه و بهبود شرایط موجود در کل جامعه باشد. شاخص نیاز به درمان ارتودنسی (Index of Orthodontic Treatment Need) ابزار مناسبی برای اندازه‌گیری میزان نیاز به درمان در جامعه بوده و دارای کارآئیهای متعددی از جمله برنامه‌ریزی، تقسیم منابع و بهبود استانداردهای درمانی می‌باشد. (۱)

روش بررسی

مطالعه انجام شده از نوع مقطعی تحلیلی (Analytic cross-sectional) می‌باشد. در این مطالعه، IOTN که یکی از شاخصهای معروف در درمانهای ارتودنسی می‌باشد، به کمک سی مشخصه و به صورت بیشترین امتیاز (Maximum Score) اندازه‌گیری شده است. در این مطالعه از شاخص اصلاح شده سه سطحی نیاز به درمان ارتودنسی با سطوح عدم نیاز یا نیاز جزئی به درمان، نیاز متوسط به درمان و نیاز قطعی به درمان استفاده گردیده است. (۱۳)

در این مطالعه، ۴۲ کست که به تصادف از قسمت بایگانی بخش ارتودنسی دانشکده دندانپزشکی دانشگاه علوم پزشکی شهید بهشتی انتخاب شده بودند در دو زمان (به فاصله دو هفته) توسط هر کدام از شش انترن دندانپزشکی با استفاده از خط کش استاندارد DHC مورد ارزیابی قرار گرفتند. قرار بود پس از تایید روایی و پایایی این انترن‌ها، از آنها در برآورد میزان نیاز به درمان ارتودنسی دانش‌آموزان دبیرستانی شهر تهران استفاده شود.

با توجه به اینکه حجم نمونه در مطالعات مربوط به ارزیابی روایی و پایایی نوعاً کم است، این تعداد می‌تواند نمونه قابل قبولی بوده و زیربنای انجام استنباطهای آماری قرار گیرد. (۱۴)

در این مطالعه برای ارزیابی توافق درونی انترن‌ها از ضریب توافق کاپا و برای اندازه‌گیری قابلیت تشخیص آنها در تفکیک سطوح شاخص نیاز به درمان ارتودنسی از مدل پیوند مربع امتیازات (Square scores association model) و مدل توافق به علاوه پیوند مربع امتیازات (Agreement plus square scores association model) استفاده شده است. مدل‌های مذکور از مدل‌های پیوند (Association models) موجود نتیجه می‌شوند تنها با این تفاوت که در مدل‌های موجود امتیازاتی که به سطوح شاخص ترتیبی اختصاص می‌یابد به صورت خطی و در مدل‌های جدید به صورت توان دوم می‌باشد. (۱۵-۱۷)

قابلیت تشخیص در ارتباط با دو سطح مجاور بیان می‌دارد که آن دو سطح تا چه میزان برای انترن‌های دندانپزشکی قابل تفکیک می‌باشند. (۱۸-۱۹)، دامنه تغییرات این پارامتر

اگرچه در تحقیقهای قبلی از شاخص نیاز به درمان ارتودنسی در ارزیابی تک تک بیماران اشکالاتی گرفته شده اما این شاخص در بررسیهای جامعه‌نگر همچنان اعتبار روزافزونی داشته و به علت سرعت اندازه‌گیری آن، طرفداران بسیاری دارد. (۲-۳)

این شاخص دارای دو جزء سلامت دندانی (Dental health) و زیبایی (Aesthetic) می‌باشد و افراد با توجه به بدترین ناهنجاری دندانی خود در یکی از پنج گروه، نیازی به درمان ندارند، به درمان جزئی نیاز دارند، در مرز نیاز درمانی قرار دارند، به درمان نیاز دارند، شدیداً به درمان نیاز دارند، قرار می‌گیرند. (۴)، این شاخص در بسیاری از کشورهای اروپایی و آسیایی که در آنها بیمه درمان ارتودنسی مطرح می‌باشد، مورد استفاده قرار می‌گیرد. (۵-۶)

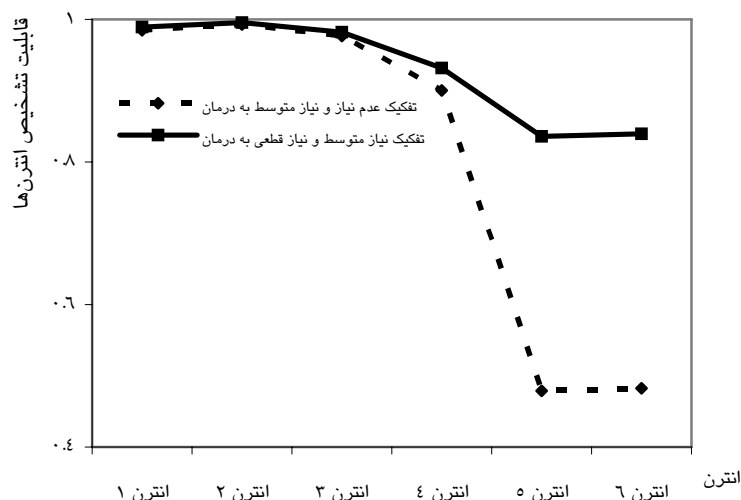
قبل از اندازه‌گیری هر شاخصی در تحقیقهای میدانی، ارزیابی روایی و پایایی آن الزامی است. معمولاً برای انجام این ارزیابی از ضریب توافق کاپا استفاده می‌شود. اما استفاده از این ضریب و تکیه صرف به نتایج حاصل از آن، خالی از اشکال نمی‌باشد و توصیه می‌شود که در کنار این محاسبه، ساختار توافق داده‌ها نیز مدلبندی شود تا اطلاعات نهفته در داده‌ها بیش از پیش استخراج گردد. (۷-۱۱)

یکی دیگر از مسائل مهم در اندازه‌گیری شاخصها، سطح آگاهی آزماینده‌ها می‌باشد. در مورد شاخص نیاز به درمان ارتودنسی Richmond و همکاران (۱۲) نشان داده‌اند که ضریب توافق کاپا در تشخیص نیاز به درمان ارتودنسی در بین ارتودنتیست‌ها ۰/۷۰ و در بین پزشکان عمومی ۰/۵۸ است. لذا تعیین روایی این شاخص باید در مقایسه با بالاترین سطح تخصص یعنی ارتودنتیست‌ها انجام گیرد.

در این مطالعه از زاویه توافق درونی آزماینده‌ها (Intra-rater agreement) و همچنین قابلیت تشخیص رده‌ها (Category distinguishability) به مقوله پایایی شاخص نیاز به درمان ارتودنسی، پرداخته می‌شود. هدف این مقاله تعیین عملکرد انترن‌های دندانپزشکی در تشخیص نیاز به درمان ارتودنسی با استفاده از شاخص IOTN در دو بار اندازه‌گیری با فاصله زمانی دو هفته می‌باشد.

یک و دو، و حداقل مقدار ۰/۹۳ برای گروه‌های مجاور دو و سه بودند. به علاوه در مورد این انترن‌ها تفاوت قابل ملاحظه‌ای در قابلیت تشخیص گروه‌های یک و دو با گروه‌های دو و سه وجود نداشت، یعنی آنها در تفکیک گروه‌های یک و دو همان قدر مهارت داشتند که در تفکیک گروه‌های دو و سه. زیرا میانگین قابلیت تشخیص این افراد برای گروه‌های یک و دو برابر ۰/۹۶ و برای گروه‌های دو و سه برابر ۰/۹۷ بدست آمد.

قابلیت تشخیص برای دو انترن دیگر - که مدل پیوند مربع امتیازات برای نرخ‌گذاری آنها مناسب تشخیص داده شد - پایینتر بود. بخصوص برای گروه‌های مجاور یک و دو، این انترن‌ها دارای میانگین قابلیت تشخیص ۰/۴۸ بودند. این آزماینده‌ها در تفکیک سطوح دو و سه، با حداقل قابلیت تشخیص ۰/۸۳، و میانگین قابلیت تشخیص ۰/۸۴ وضعیت بهتری نسبت به سطوح یک و دو داشتند. در مجموع این انترن‌ها با میانگین کلی ۰/۶۶ دارای قابلیت تشخیص پایینتری نسبت به انترن‌های قبل با میانگین کلی ۰/۹۷ بودند. نمودار ۱ جزئیات بیشتری را ارائه می‌کند.



مدل اول: توافق به علاوه پیوند مربع امتیازات (برای چهار انترن اول)
مدل دوم: پیوند مربع امتیازات (برای دو انترن آخر)

نمودار ۱: قابلیت تشخیص شش انترن دندانپزشکی به تفکیک
مدل اول و دوم در تعیین سطوح شاخص اصلاح شده سه
سطحی IOTN

مانند ضریب تعیین (R^2) در مدل‌های رگرسیون بین صفر و یک می‌باشد، بنابراین هر قدر قابلیت تشخیص دو رده مجاور برای یک انترن به عدد یک نزدیکتر باشد گفته می‌شود وی در تفکیک دو رده مذکور بهتر عمل کرده است.

برای استخراج جداول و محاسبه ضرایب کاپا از نرم‌افزار SPSS ویرایش ۱۳، برای برازاندن مدل‌ها از نرم‌افزار SAS ویرایش هشت و برای محاسبه قابلیت تشخیص با استفاده از فرمول‌نویسی و رسم نمودار، از نرم‌افزار EXCEL ۲۰۰۰ استفاده شد.

یافته‌ها

جدول ۱ خلاصه‌ای از کارهای انجام شده روی داده‌های مربوط به شش انترن دندانپزشکی را نشان می‌دهد. ضریب توافق کاپا بین دو بار اندازه‌گیری انترن‌ها با تغییر از ۰/۵۴ (برای انترن شماره پنج) تا ۰/۷۷ (برای انترن شماره یک) و میانگین ۰/۶۵ مقادیر قابل قبولی از توافق را ارائه می‌کند و بیانگر پایداری خوب آنهاست. (۲۰)

جدول ۱: ضریب کاپا و قابلیت تشخیص برای هر کدام از

انترن‌های دندانپزشکی			
شماره انترن	مقدار کاپا	قابلیت تشخیص برای سطوح عدم نیاز و نیاز متوسط به درمان	قابلیت تشخیص برای سطوح نیاز متوسط و نیاز قطعی به درمان
*۱	۰/۷۷	۰/۹۸۴۸	۰/۹۸۹۴
*۲	۰/۵۸	۰/۹۹۳۲	۰/۹۹۵۵
*۳	۰/۷۳	۰/۹۷۶۳	۰/۹۸۲۳
*۴	۰/۶۰	۰/۹۰۰۲	۰/۹۳۱۹
*۵	۰/۵۴	۰/۴۷۸۸	۰/۸۳۶۳
*۶	۰/۶۸	۰/۴۸۲۵	۰/۸۳۹۶

* مدل توافق به علاوه پیوند مربع امتیازات

** مدل پیوند مربع امتیازات

چهار انترن اول - که مدل توافق به علاوه پیوند مربع امتیازات برای نرخ‌گذاری آنها مناسب بود - دارای قابلیت تشخیص بالایی، با حداقل مقدار ۰/۹ برای گروه‌های مجاور

گمراه کننده باشد. اتفاقاً در داده‌های رتبه‌ای - که اساس کار این مطالعه می‌باشد، این فرض به ندرت برقرار است. (۷)،
ب- ضریب کاپا فقط سطح توافق روی هم مشاهده شده را با سطح توافق روی هم مورد انتظار مقایسه می‌کند و به الگوی توافق کاری ندارد. (۸)،

ج- این شاخصها فقط برای جداول مربعی که در آنها هر دو آزماینده سطوح یکسانی از نرخ‌گذاری را تشخیص دهند قابل محاسبه می‌باشد. یعنی در صورتی که یکی از سطوح مقیاس ترتیبی، توسط یک آزماینده تشخیص داده شود ولی توسط آزماینده دیگر تشخیص داده نشود، این ضرایب قابل محاسبه نخواهند بود. (۹)،

د- تفسیر ضریب کاپا مشکل است و این ضریب به شدت متأثر از توزیع‌های حاشیه‌ای آزماینده‌هاست. (۱۰)،

ه- کاپا موزون که از آن فقط برای بررسی توافق بین نرخ‌گذاریهای ترتیبی استفاده می‌شود، بیشتر از نسبت دقیق توافق نسبت به انتخاب وزنها حساس است. به این معنی که با دادن مجموعه‌های متفاوتی از وزنها به سلول‌های جدول توافقی مقادیر مختلفی برای این ضریب به دست می‌آید، لذا تفسیر درست از این ضریب دشوار می‌باشد. (۱۱)

در این مطالعه ۲۳٪ از انترن‌هایی که در تفکیک سطوح یک و دو شاخص نیاز به درمان ارتودنسی مشکل داشتند، دارای مشکل قابل‌ذکری در تفکیک سطوح دو و سه نبودند. سایر انترن‌ها نیز دارای قابلیت تشخیص بسیار بالا در تفکیک همه سطوح مجاور هم بودند. بنابراین در حالی که تشخیص نیازهای شدید به درمان ارتودنسی ساده است، تفکیک سطوح پایین شاخص مذکور می‌تواند خالی از اشکال نباشد. لذا قبل از اینکه در یک کار میدانی وسیع از انترن‌های دندانپزشکی خواسته شود که این شاخص را اندازه‌گیری نمایند، باید آموزشهای لازم را با تمرکز بر سطوح پایین مقیاس به آنها داده شود. این اطلاع اضافه‌تر در مورد تعیین پایایی اندازه‌گیری شاخص توسط انترن‌ها، فقط به کمک مدل‌های پیوند حاصل می‌شود نه شاخصهای توافق.

در بین شش آزماینده ضریب کاپا مربوط به آزماینده‌های دو و چهار مشابه آزماینده‌های پنج و شش می‌باشد. در حالی که با توجه به مدل‌های برازش شده، انترن‌های پنج و شش قابلیت تشخیص پایینی برای تفکیک سطوح اول و دوم شاخص دارا می‌باشند (به ترتیب ۰/۴۷۸۸ و ۰/۴۸۲۵).

بحث

اگر چه برای هیچ کدام از انترن‌ها توافق کامل یا تقریباً کاملی (به عنوان مثال با ضریب توافق حداقل ۰/۹) بین نرخ‌گذاریها در دو زمان دیده نمی‌شود، ولی این ضریب با تغییر از ۰/۵۴ تا ۰/۷۷ مقادیر قابل قبولی از پایایی آزماینده‌ها را ارائه می‌کند. این مقدار توافق تقریباً برابر با توافق ارائه شده توسط Richmond و همکاران (۲۱) و اندکی پایینتر از مقدار ارائه شده توسط Cooper و همکاران (۶) می‌باشد.

در تحقیقهای مشابه، برای بررسی روایی و پایایی اندازه‌گیری نیاز به درمان ارتودنسی با استفاده از IOTN، از کاپا و کاپا موزون استفاده شده است. (۲۱-۲۳)، اگرچه این ضرایب میزان کلی توافق را نشان می‌دهند اما با توجه به اشکالاتی که در مقالات مختلف به آنها وارد شده (۷-۱۱) و نظر به اینکه نمی‌توان به کمک آنها در مورد قابلیت تشخیص آزماینده‌ها در بین سطوح یک شاخص اظهار نظر کرد، استفاده از این ضرایب چندان خالی از ایراد نیست. لذا لازم است که علاوه بر محاسبه این ضرایب، با استفاده از مدل‌های آماری به مقوله قابلیت آزماینده‌ها در تفکیک رده‌های مقیاس ترتیبی نیز پرداخته شود تا بتوان در مراحل بعدی آموزش، ضعف آزماینده‌ها را در اندازه‌گیری سطحی از شاخص که دچار مشکل می‌باشند، تعیین و جبران کرد.

پاره‌ای از مشکلات کاپا و کاپا موزون عبارتند از:

الف- صورت آماره کاپا را می‌توان به فرمی نوشت که نشان دهد این ضریب بر مبنای مجموع تفاوت‌های بین فراوانی مشاهده شده و فراوانی مورد انتظار سلول‌های روی قطر اصلی، تحت مدل استقلال می‌باشد. به عبارت دیگر کاپا برای اندازه‌گیری توافق اصلاح شده، مدل استقلال را به عنوان مدل مبنا مورد استفاده قرار می‌دهد. بنابراین در صورتی که فرض استقلال برقرار نباشد، استفاده از این آماره می‌تواند

نتیجه‌گیری

ضروری است. در آموزش انترن‌های دندانپزشکی جهت

اندازه‌گیری نیاز به درمان ارتودنسی، تاکید بیشتر باید در مورد نحوه تعیین سطوح پایین شاخص صورت پذیرد.

استفاده از مدل‌های پیوند جهت تکمیل یافته‌های حاصل از ضریب توافق کاپا در بحث روایی و پایایی آزماینده‌ها

REFERENCES

1. Shaw WC, Richmond S, O'Brien KD, Brook P, Stephens CD. Quality control in orthodontics: Indices of treatment need and treatment standards. Br Dent J. 1991 Feb;170(3):107-112.
2. So LL, Tang EL. A comparative study using the occlusal index and the index of orthodontic treatment need. Angle Orthod. 1993 Spring; 63(1): 57-64, 65-66.
3. Jones CM, Woods K, O'Brien KD, Winard C, Tylor GO. Index of orthodontic treatment need, its use in dental epidemiology survey calibration exercise. Community Dent Health. 1996 Dec;13(4): 208-210.
4. Brook PH, Shaw WC. The development of an index for orthodontic treatment priority. Eur J Orthod. 1989 Aug; 11(3):309-320.
5. Waring D, Jones JW. Does the GDP need to know about IOTN. Dent Update. 2003 Apr; 30(3):123-130.
6. Cooper S, Mandall NA, Dibiasi D, Shaw WC. The reliability of the Index of orthodontic treatment need over time. J Orthod. 2000 Mar; 27(1):47-53.
7. Kraemer HC, Periaikov VS, Noda A. Tutorial in biostatistics, kappa coefficients in medical research. Stat Med. 2002 Nov;21(1):2109-2129.
8. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problem of two paradoxes. J Clin Epidemiol. 1990 June;43(6):543-549.
9. May SM. Modelling observer agreement – an alternative to kappa. J Clin Epidemiol. 1994 Nov; 47(11):1315-1324.
10. Tanner MA, Young MA. Modelling agreement among raters. J Am Statis Assoc. 1985 Mar;80(389):175-180.
11. Perkins SM, Becker MP. Assessing rater agreement using marginal association models. Stat Med. 2002 June; 21(12):1743-1760.
12. Richmond S, O'Brien KD, Roberts CT. Dentist's variation in determination of orthodontic treatment need. Br J Orthod. 1994 Feb;21(1):65-68.
13. Richmond S, Shaw WC, O'Brien KD, Buchanan IB, Stephens CD, Andrews M, et al. The relationship between the Index of orthodontic treatment need and the consensus opinion of a panel of 74 dentists. Br dent J. 1995 May; 178(10):370-374.
14. Fleiss JL. The design and analysis of clinical experiments. 1st ed. New York: John Wiley & Sons; 1999,8.
15. Goodman LA. Simple models for the analysis of association in cross-classifications having ordered categories. J Am Statis Assoc. 1979 Sep;74(367):537-551.
16. Agresti A. A model for agreement between ratings on an ordinal scale. Biometrics. 1988 Jun;44(2):539-548.
۱۷. اکبرزاده‌باغبان، ع؛ بابایی، غ؛ کاظم‌نژاد، الف. یک خط مشی جدید در مدل‌بندی توافق داده‌های ترتیبی. دو ماهنامه علمی پژوهشی دانشگاه شاهد (دانشور)، دی ماه ۱۳۸۳؛ سال دوازدهم، شماره ۵۴: ۱۹-۲۸.

18. Darroch JN, McCloud PI. Category distinguishability and observer agreement. *Austral J Statist.* 1986;28(3):371-388.
19. Becker MP, Agresti A. Log - linear modelling of pairwise interobserver agreement on a categorical scale. *Stat Med.* 1992 Jan; 11(1):101-114.
20. Byrt T. How good is that agreement? (Letter to editor). *Epidemiology.* 1996 Sep;7(5):561.
21. Richmond S, Roberts CT, Anderws M. Use of the index of orthodontic treatment need (IOTN) in assessing the need for orthodontic treatment pre- and post- appliance therapy. *Br J Orthod.* 1994 May;21(2):175-184.
22. Burden DJ, Pine CM, Burnside G. Modified IOTN: An orthodontic treatment need index for use in oral health surveys. *Com Dent Oral Epidemiol.* 2001 Jan;29(3):220-225.
23. Kerosuo H, Enezi SA, Kerosuo E, Abdulkarim E. Association between normative and self perceived orthodontic treatment need among Arab high school students. *Am J Orthod Dentofacial Orthop.* 2004 Mar;125(3):373-378.