

KEBUGARAN JASMANI MENURUT INSTRUMEN GPAQ DIBANDINGKAN DENGAN VO2MAX PADA WANITA UMUR 25 SAMPAI 54 TAHUN

Physical Fitness According to GPAQ Instruments Compared with Vo2Max in Women Age 25 to 54 Years Old

Olwin Nainggolan, Lely Indrawati, Julianty Pradono

Puslitbang Upaya Kesehatan Masyarakat - Balitbangkes - Kemenkes RI, Jl. Percetakan Negara No. 29 Jakarta

Naskah Masuk: 7 Mei 2018, Perbaikan: 25 Mei 2018, Layak Terbit: 9 Juni 2018

<http://dx.doi.org/10.22435/hsr.v2i14.752>

ABSTRAK

Instrumen *Global Physical Activity Questionnaire (GPAQ)* adalah *tools* yang paling sering digunakan untuk mengukur aktivitas fisik dalam survei yang dilakukan oleh Kementerian Kesehatan. Analisis *Receiver Operating (ROC)* dilakukan untuk melihat kemampuan diagnostik instrumen GPAQ terhadap aktivitas fisik dengan nilai *VO2max* sebagai pembanding. Penelitian ini bertujuan untuk mengetahui sebaran nilai *METs* menit/minggu setiap domain aktivitas, nilai *cut off point METs*, luas daerah di bawah kurva (*AUC*), sensitivitas, spesifisitas, nilai prediksi positif dan nilai prediksi negatif dari pengukuran GPAQ dengan *gold standard* pemeriksaan *VO2max*. Pemeriksaan tingkat kebugaran menggunakan sepeda statis (*ergocycle*) dengan metoda *astrand*. Subyek penelitian adalah wanita umur 25 sampai dengan 54 tahun sebanyak 117 orang di Kecamatan Bogor Tengah. Penelitian ini merupakan penelitian deskriptif dan uji diagnostik dengan desain potong lintang. Pengukuran aktivitas fisik dilakukan dengan wawancara secara terstruktur menggunakan instrumen GPAQ versi 2 yang terdiri dari 16 pertanyaan pada 3 domain aktivitas yaitu pekerjaan, perjalanan dan rekreasi. Hasil penelitian menunjukkan seluruh responden tidak memiliki aktivitas kategori berat baik domain pekerjaan maupun rekreasi. Rerata nilai *METs* dengan domain pekerjaan level sedang $4271,69 \pm 2874,34$; domain perjalanan 1058 ± 1730 ; domain rekreasi kategori sedang dengan nilai *METs* $181,23 \pm 471,594$ dan nilai rerata total keseluruhan *METs* adalah $5511,11 \pm 3440,48$. Nilai *cut off point* instrumen GPAQ sebesar 4.668/minggu, sensitivitas = 50,0; spesifisitas = 43,3; nilai prediksi positif 75,9%; nilai prediksi negatif 29,3%. Dapat disimpulkan bahwa nilai rerata *METs* tiap domain aktivitas *over estimate*, tingkat aktivitas fisik menggunakan instrumen GPAQ tidak akurat dan tidak ada korelasi antara tingkat aktivitas fisik dengan instrumen GPAQ dibandingkan dengan kebugaran *VO2max* dengan $P = 0.451$. Perlu alternatif instrumen lain untuk penilaian aktivitas fisik dalam survei berbasis populasi di luar instrumen GPAQ.

Kata kunci: ROC, aktivitas fisik, VO2 Max, sensitivitas, spesifisitas

ABSTRACT

The Global Physical Activity Questionnaire (GPAQ) is the most commonly used tool for measuring physical activity in surveys. Receiver Operating Analysis (ROC) is performed to see the GPAQ instrumen's diagnostic ability to physical activity with VO2max as a comparison. The aims of study is to determine the distribution of METs minutes / weeks of each activity domain, cut off point METs, area under the curve (AUC), sensitivity, specificity, positive predictive value and negative predictive value of GPAQ instrumen measurement with gold standard VO2max. Examination of fitness level using a bicycle static (ergocycle) with astrand method. Research subjects were women aged 25 to 54 years as many as 117 people in Central Bogor District. It is descriptive research with cross sectional design. The measurement of physical activity is done by interview using GPAQ instrumen version 2 which consists of 16 questions on 3 domains of activities ie work, travel and recreation. The results showed that all respondents did not have heavy category activity either job or recreation domain. Average METs score with moderate job domain $4271,69 \pm 2874,34$; travel domain 1058 ± 1730 ; medium category recreation domain with METs value of $181.23 \pm 471,594$ and the overall average total METs were 5511.11 ± 3440.48 . Cut off point value of GPAQ instrumen is 4,668/week; sensitivity = 50,0; specificity = 43,3; a positive predictive

Korespondensi:

Olwin Nainggolan

Puslitbang Upaya Kesehatan Masyarakat - Badan litbangkes - Kemenkes RI

E-mail: olwin.n@gmail.com

value of 75.9%; negative predictive value of 29.3%. The concluded that the average value of METs for each domains over estimate activity, the level of physical activity using GPAQ instrumen is not accurate and there is no correlation between the level of physical activity with GPAQ instrumen compared with fitness V02max with $P = 0.451$. An alternative instrumen for the assessment of physical activity in a population-based survey is required.

Keywords: ROC, physical activity, V02max, sensitivity, specificity

PENDAHULUAN

Aktivitas fisik secara teratur merupakan perilaku gaya hidup yang penting untuk pengembangan dan pemeliharaan kesehatan individu dan populasi untuk mencapai kesejahteraan. Secara global, tingkat ketidakaktifan yang tinggi berkontribusi secara substansial terhadap beban penyakit global, dengan biaya ekonomi yang tinggi. Kebugaran merupakan kebutuhan bagi setiap manusia untuk dapat melakukan Aktivitas hidup secara optimal. Setiap Aktivitas yang kita lakukan memerlukan kondisi tubuh yang baik. Untuk mendapatkan kondisi tubuh yang baik diperlukan kebugaran yang cukup. Kita mengetahui untuk mencapai kebugaran tersebut salah satunya adalah dengan melakukan Aktivitas fisik maupun dengan berolahraga secara teratur¹

Aktivitas fisik berperan sangat penting untuk mencegah penyakit tidak menular kronis, namun data yang digunakan untuk menginformasikan hal tersebut masih kurang. Kementerian Kesehatan melalui Badan Litbang Kemenkes telah melakukan beberapa riset berskala nasional, dengan menyediakan data perilaku Aktivitas fisik masyarakat Indonesia, seperti pada Riset Kesehatan Dasar (Risekesdas) tahun 2007 dan 2013 dengan menggunakan instrumen GPAQ (*Global Physical Activity Questionnaire*). GPAQ merupakan kuesioner yang dikeluarkan WHO untuk mengetahui derajat Aktivitas fisik individu terutama pada negara berkembang².

Kuesioner Aktivitas Fisik Global (GPAQ) dikembangkan pada tahun 2002 oleh Organisasi Kesehatan Dunia (WHO) sebagai bagian dari pendekatan WHO *STEPwise* untuk pengamatan Aktivitas fisik *surveillance* faktor risiko penyakit kronis³. Penggunaan instrumen ini untuk pengawasan Aktivitas fisik nasional yang direkomendasikan oleh WHO dalam strategi global pada tahun 2004⁴. GPAQ terdiri dari 16 pertanyaan yang dirancang untuk memperkirakan tingkat Aktivitas fisik seseorang pada 3 domain Aktivitas yaitu (pekerjaan, perjalanan dan rekreasi) serta waktu yang dihabiskan untuk perilaku *sedentari*⁵. Instrumen GPAQ dikembangkan sebagai alat yang dapat digunakan untuk melakukan evaluasi

dan membandingkan tingkat Aktivitas fisik pada skala lokal dan internasional⁶.

Banyak penelitian untuk mengukur *validitas* dan *reliabilitas* instrumen GPAQ di dunia dengan menghasilkan banyak variasi. Hasil penelitian di beberapa negara menunjukkan perbedaan tingkat pendidikan bisa menyebabkan perbedaan tingkat validitas⁶. Instrumen GPAQ perlu dievaluasi lebih lanjut apabila akan terus digunakan dalam survei besar yang dilakukan di Indonesia. Belum ada penelitian yang dilakukan di Indonesia yang mereferensikan penggunaan instrumen GPAQ untuk mengukur perubahan Aktivitas fisik.

Uji diagnostik yang ideal adalah memberikan hasil Aktivitas fisik “cukup” pada semua subyek yang memiliki kebugaran baik, dan memberikan hasil “kurang” Aktivitas fisik pada subyek yang tidak bugar. Tetapi ada kemungkinan akan diperoleh hasil uji positif (Aktivitas fisik cukup) pada subyek yang tidak bugar (positif palsu) dan hasil negatif (kurang Aktivitas fisik) pada subyek bugar (negatif palsu). Penelitian ini bertujuan untuk membandingkan hasil yang diperoleh wawancara Aktivitas fisik menggunakan instrumen GPAQ dengan standar baku pemeriksaan kebugaran yaitu V02max. Analisis *Receiver Operating Curve* (ROC) berfungsi untuk melihat akurasi data Aktivitas fisik yang diperoleh dari hasil wawancara dengan menggunakan instrumen GPAQ dibandingkan dengan tingkat kebugaran yaitu VO2 max. Dalam penelitian ini akan dilihat sebaran nilai METs (*Metabolic Equivalent*) Aktivitas fisik responden di kecamatan Bogor Tengah dengan menggunakan instrumen GPAQ. Selanjutnya dilakukan analisis korelasi antara nilai METs dari instrumen GPAQ dengan nilai VO2max menggunakan analisis *Pearson correlation*. Dengan analisis ROC akan dihasilkan nilai AUC (*Area Under Curve*), *sensitivity*, *specificity*, Nilai Prediksi Positif (NPP), serta Nilai Prediksi Negatif (NPN) data Aktivitas fisik dengan instrumen GPAQ. Diharapkan dengan penelitian ini, menjadi salah satu bahan pertimbangan apakah kuesioner GPAQ masih layak digunakan sebagai instrumen pengumpulan data Aktivitas fisik di riset nasional

selanjutnya, khususnya di lingkungan Kementerian Kesehatan Republik Indonesia.

METODE

Penelitian ini merupakan penelitian observasional analitik dan uji diagnostik dengan rancangan potong lintang. Sampel diambil secara *purposive sampling* dari data penduduk wanita dengan rentang usia 25 sampai dengan 54 tahun di dua kelurahan yaitu Kelurahan Paledang dan Kelurahan Gedung di Kota Bogor. Penelitian dilakukan mulai bulan Juli sampai dengan Agustus tahun 2016.

Subyek

Sebanyak 117 wanita, berusia 25 tahun sampai dengan 54 tahun, direkrut dari 2 kelurahan yang dipilih secara *purposive* di Kelurahan Paledang serta Kelurahan Gudang Bogor di Kecamatan Bogor Tengah. Semua responden tidak dalam keadaan hamil, tidak memiliki cacat fisik dan mampu berdiri tegak untuk pengukuran antropometri. Responden yang terpilih tidak memiliki riwayat diabetes mellitus, hipertensi, obesitas, dan penyakit jantung. Diabetes mellitus ditegakkan dengan pemeriksaan gula darah 2 jam *post prandial*, hipertensi dengan pengukuran tekanan darah, obesitas dengan IMT yaitu pengukuran berat dan tinggi badan serta penyakit jantung dengan pemeriksaan EKG.

Analisis

Data Aktivitas fisik yang dikumpulkan menggunakan instrumen GPAQ versi 2 hanya dilakukan pada 3 domain Aktivitas yaitu Aktivitas pekerjaan, perjalanan dan rekreasi dalam hal ini adalah Aktivitas olah raga⁸. Pada tahap awal, data dilakukan analisis deskriptif untuk mengetahui sebaran nilai METs meliputi *mean* (rata-rata), *median*, minimum, maksimum serta standard deviasi pengukuran METs pada masing-masing domain instrumen GPAQ serta nilai *VO2max*. Kemudian dilakukan analisis korelasi dengan *pearson correlation* untuk melihat hubungan antara nilai METs dengan nilai *VO2max*. Selanjutnya dilakukan analisis ROC (Receiver Operating Characteristic), yaitu sebuah alat diagnostik yang dapat menguji keakuratan sebuah alat baru dalam hal ini data Aktivitas fisik dari pengukuran GPAQ dibandingkan dengan alat yang lama dalam hal ini

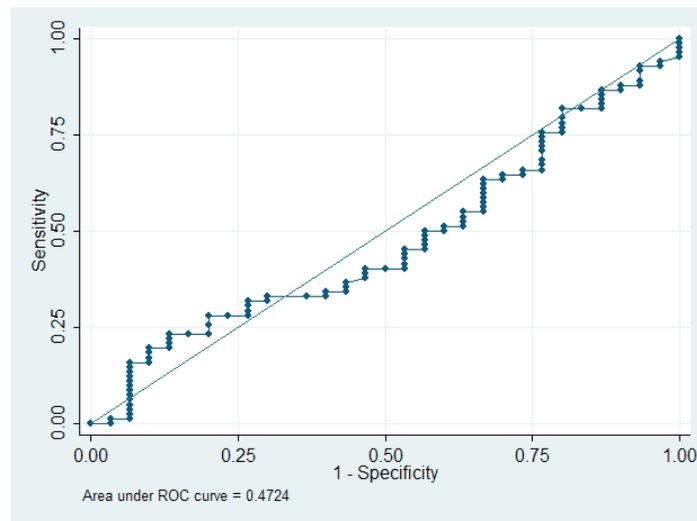
kebugaran tubuh (*VO2max*). Analisis ROC akan menghasilkan nilai yang menggambarkan keakuratan data GPAQ yang dicerminkan pada luas daerah di bawah kurva (*AUC*). Analisis ROC menghasilkan nilai titik potong (*cut off point*) Aktivitas fisik terhadap kebugaran. Dengan menggunakan uji diagnostik akan dihasilkan nilai ratio yaitu *sensitivitas dan spesifisitas* dari instrumen GPAQ serta Nilai Prediksi Positif (*NPP*) dan Nilai Prediksi Negatif (*NPN*). Analisis korelasi digunakan untuk menguji hubungan antara nilai METs dan kebugaran *Vo2max* dengan *Pearson correlation* pada $\alpha=0,05$. Keseluruhan analisis data menggunakan perangkat lunak *Stata* dengan serial number 81990548552 serta SPSS 16 dengan serial number 5061284.

Persetujuan Etik

Penelitian ini dilakukan setelah memperoleh persetujuan etik penelitian yang dikeluarkan oleh dari Komisi Etik Penelitian Kesehatan Badan Litbangkes dengan nomor persetujuan: LB.02.01/5.2/KE. 257/2016). Izin Persetujuan Setelah Penjelasan (PSP) dimintakan dari seluruh calon responden secara tertulis dan ditandatangani sebelum menjadi responden penelitian.

Teknik Pengambilan Sampel

Semua wanita dengan rentang usia 25–54 tahun yang terletak di Kelurahan Paledang dan Gudang diberikan kesempatan yang sama untuk berpartisipasi dalam penelitian ini. Kader Posbindu (Pos Pembinaan Terpadu) mencatat semua calon responden dengan rentang usia yang dibutuhkan, kemudian kader akan menanyakan apakah bersedia untuk ikut dalam penelitian ini. Sebelum mengikuti pengukuran kebugaran dengan metode *ergocycle*, calon responden diharuskan menjawab questioner *PAR-Q (Anamnesis Physical Activity Readiness Questionnaire)* yang terdiri dari 7 (tujuh) pertanyaan. Kuesioner *PAR-Q* bertujuan menjangkau orang-orang yang pernah didiagnosis penyakit jantung, hipertensi dan masalah persendian, sehingga memiliki fisik yang tidak mampu untuk melakukan Aktivitas fisik. Calon responden yang terpilih adalah responden yang keseluruhannya menjawab tidak. Calon responden yang telah lolos uji *PAR-Q*, selanjutnya akan dilakukan pemeriksaan kadar gula darah, kadar kolesterol, berat badan, tinggi badan dan



Grafik1. Luas Area Under Curve (AUC)

elektrokardiogram (EKG). Calon responden yang memenuhi seluruh syarat akan diikutsertakan dalam penelitian.

Tim enumerator direkrut dengan latar belakang pendidikan minimal S1 kesehatan, dan diberikan pelatihan untuk mengenal instrumen dan praktik wawancara. Setiap enumerator mengunjungi responden untuk melakukan wawancara. Untuk kontrol kualitas data, selama pengumpulan data, tim enumerator didampingi oleh supervisor. Untuk pemeriksaan VO_{2max} , dilakukan oleh tim dari Balai Kesehatan Masyarakat Bandung (BKOM) Kementerian Kesehatan dan responden dikumpulkan untuk melakukan tes kebugaran dengan menggunakan *ergocycle test* (tes mengayuh sepeda).

Tes Kebugaran *Ergocycle*

Tes kebugaran dilakukan dengan menggunakan sepeda merk Monark 828E dengan beban kayuh yang dapat disetel berdasarkan kebutuhan. Parameter pada tes ini adalah frekuensi denyut nadi dan besar beban, yang selanjutnya digunakan untuk menentukan besar VO_{2max} (*maximal oxygen consumption*). VO_{2max} adalah kecepatan maksimum tubuh menggunakan oksigen selama latihan fisik. Satuan absolut VO_{2max} adalah volume oksigen per menit (L/min) dan satuan relatif VO_{2max} berkaitan dengan berat badan (BB) yaitu volume oksigen per kg BB per menit (ml/kg/min).

Sebelum melakukan pemeriksaan, responden diminta untuk menandatangani pernyataan bahwa responden dalam keadaan sehat dan tidak memiliki penyakit yang terkait dengan kontra

indikasi penelitian. Setelah itu responden dilakukan pencatatan dan pengukuran berat badan, tinggi badan, dan tekanan darah. *Chest Band* atau *polar heart rate* (HR) meter dipasang, kemudian diukur *heart rate* istirahat responden. Selanjutnya responden diminta duduk pada sepeda *Monarch* kemudian dilakukan pemanasan kurang lebih lima menit dengan mengayuh sepeda dengan kecepatan konstan 50 rpm dengan beban dinaikkan secara bertahap hingga mencapai 300 kpm (1kp). Beban yang digunakan adalah 600 kpm (2kp) untuk laki-laki dan 450 kpm (1,5kp) untuk wanita dengan kecepatan 50 rpm konstan selama 5 menit. Saat *denyut nadi* mencapai 170 bpm atau ada keluhan atau tampak kesakitan seperti nyeri dada, sesak, maka percobaan dihentikan⁷.

HASIL

Dari tabel 1 terlihat gambaran sebaran nilai *METs* setiap unsur dari instrumen *GPAQ* meliputi domain pekerjaan, perjalanan, rekreasi serta total seluruh nilai *METs*. Jumlah responden yang dilakukan wawancara sebanyak 117 responden dan yang dapat dilakukan pemeriksaan VO_{2max} sebanyak 115 responden. Secara keseluruhan sebanyak 117 responden pada penelitian ini tidak mempunyai Aktivitas fisik domain pekerjaan dan domain rekreasi dalam kategori berat dan hanya masuk Aktivitas sedang. Nilai *METs* domain pekerjaan kategori sedang mempunyai rata-rata $4271,69 \pm 2874,34$ dan domain perjalanan dengan nilai rata-rata *METs* $1058,19 \pm 1730,57$. Untuk domain rekreasi kategori

Tabel 1. Karakteristik dan Sebaran Nilai METs pada Setiap Domain Aktivitas Fisik dan VO2max

Level Aktivitas Fisik	Mean	Median	Min	Max	SD	n
GPAQ (METs menit/minggu)						
Bekerja						
Berat	-	-	-	-	-	
Sedang	4271,69	3528,00	360	14280	2874,34	117
Perjalanan	1058,19	560,00	0	12880	1730	
Rekreasi						
Berat	-	-	-	-	-	
Sedang	181,23	0	0	3408	471,594	117
Total METs menit/minggu	5511,11	4900,00	560	16848	3440,48	
METs menit/minggu per kelompok umur						
<41 Tahun	4646,81	3600,00	720	15820	3216,41	46
>=41 Tahun	6091,43	5300,00	560	16848	3486,03	66
VO2max per kelompok umur						
<41 Tahun	33,14	30,52	13,80	75,80	12,32	46
>=41 Tahun	27,44	26,90	12,90	46,00	7,93	66
VO2max Total	29,78	17,60	12,90	75,80	10,31	115

Tabel 2. Nilai Sensitivitas dan Spesifisitas Optimum pada Berbagai Nilai *Cut-off* METs pada Aktivitas Fisik, Wanita Umur 25–54 Tahun di Kecamatan Bogor Tengah, 2016

METs menit/minggu	Sensitivitas	Spesifisitas
4.172	53,66	36,67
4.392	52,44	36,67
4.560	51,22	36,67
4.564	51,22	40,00
4.620	50,00	40,00
4.668	50,00	43,33
4.800	48,78	44,33
4.900	47,56	43,33
5.096	46,34	43,33
5.180	45,12	43,33
5.208	45,12	46,67

Tabel 3. Tabel Uji Diagnostik GPAQ dengan Gold Standard Pemeriksaan Kebugaran VO2max

Statistik diagnostik	%	95% CI
Prevalensi	73,0	64,0-81,1
Sensitivitas	50,0	38,7-61,3
Spesifisitas	56,7	37,4-74,5
Nilai Prediksi Positif	75,9	62,4-86,5
Nilai Prediksi Negatif	29,3	18,1-42,7

Tabel 4. Uji Korelasi antara Tingkat Kebugaran dengan Aktifitas Fisik

	Korelasi	VO2max	METs menit /minggu
METs menit /minggu	Pearson Correlation	0,070	1
	Signifikansi (P value)	0,465	

berat juga tidak ditemukan, dengan kategori sedang mempunyai rata-rata *METs* 5511,11 ± 3440,48. Dari pengukuran VO_{2max} dihasilkan nilai rata-rata 29,18 ml O_2 /kg BB/menit ± 10,31.

Grafik 1 memperlihatkan hasil analisis ROC menggunakan aplikasi STATA menunjukkan nilai *Area Under Curve (AUC)* sebesar 0,4724. Terlihat titik-titik yang menghubungkan nilai sensitivitas dan 1-Spesifisitas sebagian besar berada pada garis diagonal kurva. Kurva *AUC* merupakan gabungan dari sensitivitas dan spesifisitas yang tujuannya adalah untuk mengukur kualitas sebuah uji alat diagnostik secara umum. Kurva *AUC* juga dapat diinterpretasikan sebagai rata-rata sensitivitas untuk semua nilai spesifisitas yang sering juga disebut sebagai akurasi uji diagnostik.

Dari tabel 2, nilai titik potong Aktivitas fisik *METs* dapat diperoleh dengan memperhatikan nilai sensitivitas dan Spesifisitas yang paling tinggi. Penentuan nilai titik potong (*cut off points*) dengan memperhatikan keseimbangan antara nilai sensitivitas dengan spesifisitas. Nilai titik potong yang diambil adalah nilai dimana sensitivitas tinggi tetapi nilai spesifisitas juga tinggi. Oleh sebab itu peneliti mengambil keputusan bahwa sensitivitas nya adalah sebesar 50,00% dan spesifisitas sebesar 43,33% dengan nilai *cut off point METs* adalah sebesar 4.668/minggu. Bila *cut off point* Aktivitas fisik *GPAQ* dinaikkan sebesar 51,22%, maka spesifisitas akan turun menjadi 40,00%. Dengan demikian untuk nilai Aktivitas fisik nilai *METs* sebesar 4.668 ke bawah dianggap mempunyai Aktivitas fisik “kurang aktif” sedangkan nilai *METs* sebesar 4.668/minggu ke atas dianggap mempunyai Aktivitas fisik “aktif”.

Dari tabel 3 uji diagnostik *GPAQ* dengan baku emas pemeriksaan kebugaran menggunakan VO_{2max} didapatkan bahwa sensitivitas = 50,0%, spesifisitas = 56,7%. Nilai Prediksi Positif = 75,9%; Nilai Prediksi Negatif = 29,3%. Semua ukuran diagnostik diatas berada pada *prevalensi* Aktivitas cukup sebesar 73%.

Untuk melihat hubungan antara kedua variabel diuji dengan menggunakan uji korelasi dengan menggunakan uji *korelasi Pearson*. Dari uji *korelasi* antara tingkat kebugaran dengan Aktivitas fisik, terlihat tidak ada korelasi antara keduanya dengan nilai $r = 0,070$ dan $P \text{ value (Pearson Correlation)} = 0,4465 > 0,05$ yang artinya tidak ada korelasi antara Aktivitas fisik (menggunakan instrumen *GPAQ*) dengan tingkat kebugaran (pengukuran VO_{2max}).

PEMBAHASAN

Aktivitas fisik yang dilakukan secara terstruktur dan terencana disebut latihan jasmani, sedangkan Aktivitas fisik yang tidak dilakukan secara terstruktur dan terencana disebut Aktivitas fisik sehari-hari. Untuk menilai Aktivitas fisik, 4 (empat) dimensi utama yang menjadi fokus yaitu tipe, frekuensi, durasi, dan intensitas Aktivitas fisik. Tipe adalah jenis Aktivitas fisik seperti berjalan, bersepeda, olahraga; frekuensi Aktivitas fisik mengacu kepada jumlah sesi Aktivitas fisik per satuan waktu tertentu; durasi Aktivitas fisik merupakan lamanya waktu yang dihabiskan ketika melakukan Aktivitas fisik; dan intensitas Aktivitas fisik sering dinyatakan dengan istilah ringan, sedang, atau berat⁹. Tingkat kebugaran adalah ukuran dari kesanggupan seseorang untuk dapat melakukan Aktivitas sehari-hari. Semakin baik tingkat kebugaran seseorang, maka tingkat kesanggupan untuk melaksanakan Aktivitas cenderung semakin baik terutama dari segi fisik maupun stamina. VO_{2max} sangat menentukan tingkat kebugaran seseorang. Dapat disimpulkan bahwa semakin tinggi VO_{2max} seseorang maka tingkat kebugaran semakin tinggi¹⁰.

Nilai *METs menit/minggu* wanita usia 25 sampai dengan 54 tahun di Kota Bogor dengan menggunakan instrumen *GPAQ* sebagian besar cenderung *over estimate*. Dengan rata-rata nilai *METs* menit/minggu domain pekerjaan sebesar 4.271,69 ± 2.874,34; nilai *METs* menit/minggu domain perjalanan 1.058 ± 1.730 dan domain rekreasi 181,23 ± 471,594; besar nilai relatif tinggi jika dibandingkan dengan beberapa penelitian seperti penelitian Surya Dhimas Aditya di Jogjakarta dimana responden mahasiswa rata-rata *METs* menit/minggu Aktivitas pekerjaan hanya sebesar 306,67, domain Aktivitas perjalanan 345,33 dan Aktivitas rekreasi 394,13¹¹. Penelitian Mumu SJ *etal* (2017) di Bangladesh menemukan nilai rata-rata *METs* menit/minggu dengan domain pekerjaan, perjalanan dan rekreasi berturut-turut adalah sebesar 840; 840; dan 600. Jika dilihat dari rata-rata total nilai *METs* menit/minggu pada penelitian ini, jauh lebih tinggi jika dibandingkan dengan nilai total nilai *METs* menit/minggu penelitian Surya Dhimas Aditya yang hanya sebesar 472,43¹² maupun penelitian Mumu SJ *et al.*, sebesar 3320¹³.

Dari data terlihat bahwa semua responden penelitian ini tidak mempunyai Aktivitas domain pekerjaan serta rekreasi dengan kategori berat. Nilai *METs* diperoleh dari ketiga domain yaitu pekerjaan,

perjalanan serta rekreasi hanya dari kategori sedang. Hasil penelitian yang sama ditemukan pada penelitian Singh (2011) di India, dimana respondennya juga tidak mempunyai Aktivitas pekerjaan kelompok kategori berat, demikian juga penelitian Mumu (2017) di India, nilai *METs* = 0 pada domain pekerjaan serta rekreasi dengan kategori berat. Namun disadari bahwa nilai-nilai *METs* domain pekerjaan dan rekreasi kategori sedang pada penelitian ini sebagian besar cenderung *over estimate*. *METs* (*minute equivalents*) Aktivitas domain pekerjaan kategori sedang, domain perjalanan dan domain rekreasi cenderung tinggi, sehingga kalau dikategorikan dengan *cut off point* standard *WHO* total hanya ≥ 600 *METs* menit/minggu seluruh domain, hampir semua responden penelitian masuk kategori aktif dan hanya 1 orang responden yang masuk ke dalam kelompok kurang aktif.

Nilai *METs* yang di luar kewajaran, mungkin disebabkan oleh ketidakmampuan enumerator untuk menggali lebih dalam apa sebenarnya yang menjadi Aktivitas keseharian responden. Enumerator cenderung membacakan saja pertanyaan pada instrumen, tanpa mengerti kedalaman maksud tiap-tiap butir pertanyaan. Setiap pertanyaan yang ditujukan kepada responden, responden akan menjawab dan enumerator menuliskan saja jawabannya walau jawaban yang diberikan tidak masuk akal. Misalnya pertanyaan domain rekreasi, berapa lama dilakukan. Maksud dari rekreasi disini adalah kegiatan yang dilakukan dalam waktu senggang termasuk olahraga. Namun dari jawaban sebagian responden agak berlebihan misalnya dengan jawaban kegiatan rekreasi ini diatas 10 jam. Terlihat dari jawaban responden ada kecenderungan karena dibacakan saja oleh enumerator, tanpa ada penggalian lebih lanjut karena diterjemahkan dan diartikan oleh responden, maksud rekreasi disini adalah jalan-jalan. Oleh karena itu ada baiknya dipikirkan untuk mencari kata lain sebagai pengganti "rekreasi" disini, bukan langsung mengutip dari instrumen asli dalam bahasa Inggris.

METs (*Metabolic Equivalents*) biasanya digunakan untuk menggambarkan intensitas Aktivitas fisik, dan juga digunakan untuk analisis data GPAQ. *METs* (*Metabolic Equivalents*) adalah rasio tingkat metabolisme pada saat bekerja seseorang relatif terhadap tingkat metabolisme saat beristirahat. Satu *METs* didefinisikan sebagai energi yang dihabiskan untuk duduk dan berbaring, dan setara dengan konsumsi 1 kkal/kg/jam. Diperkirakan, jika dibandingkan dengan duduk, diam konsumsi kalori

seseorang, empat kali lebih tinggi saat sedang aktif, dan delapan kali lebih tinggi saat aktif dengan giat. Oleh karena itu, ketika dilakukan penghitungan pengeluaran energi keseluruhan seseorang dengan menggunakan data GPAQ, 4 *METs* dihabiskan untuk waktu Aktivitas sedang, dan 8 *METs* untuk waktu yang dihabiskan untuk Aktivitas yang berat⁸.

Dari nilai *cut off points* nilai *METs* dengan gold standard *VO2max* didapat nilai *METs* menit/minggu adalah sebesar 4.688. Nilai ini sangat besar jika dibandingkan dengan nilai *cut off points* dari *WHO* yang hanya sebesar 600 *METs* menit/minggu. Nilai *METs* menit/minggu penelitian ini mirip dengan yang dilakukan di Malaysia hasil penelitian Lingesh G *et al.* (2016) yang mendapati nilai rata-rata *METs* menit/minggu sangat tinggi dimana Aktivitas ringan dengan 923,33; Aktivitas sedang 1243 dan Aktivitas berat 9096,62. Hasil penelitian ini juga menyatakan bahwa tidak ada korelasi antara GPAQ dengan *accelerometer*. Secara keseluruhan hasil penelitiannya dikatakan bahwa instrumen GPAQ tidak handal dan tidak akurat untuk mengukur skor Aktivitas fisik untuk mengukur nilai *METs* menit/minggu seseorang¹⁴.

Dari hasil analisis korelasi juga menunjukkan tidak adanya korelasi antara nilai *METs* dengan nilai *VO2max*. Hasil penelitian yang sama dilaporkan oleh Gademan MGJ tahun 2013 menemukan bahwa adanya hubungan yang lemah antara Aktivitas fisik dengan metode *self reported* terhadap kadar *VO2max*¹⁵. Meskipun GPAQ digunakan secara luas, beberapa penelitian seperti Lingesh G *et al* di Malaysia tahun 2016 melaporkan bahwa data GPAQ tidak akurat dan tidak dapat diandalkan sebagaimana seharusnya, karena sangat bergantung pada pola gaya hidup dan kemampuan responden dalam mengingat dan menjawab kuesioner¹⁴. Penelitian lainnya Thuy AB *et al* tahun 2010 di Vietnam melaporkan bahwa GPAQ memiliki keandalan dan akurasi yang baik jika respondennya adalah mereka yang dengan pola kerja yang stabil dibandingkan dengan orang dengan pola Aktivitas fisik yang bervariasi¹⁵.

Nilai *cut off point* sebuah *outcome* yang datanya bersifat kontinyu dengan menggunakan analisis ROC tergantung dari bobot yang diberikan untuk hasil negatif palsu dan positif palsu. Bobot yang lebih, harus diberikan pada hasil negatif palsu (lebih merugikan) jika tujuan tes adalah untuk deteksi semua kasus yang ada. Sebaliknya bobot harus diberikan lebih pada hasil positif palsu (lebih merugikan) jika tujuan

tes adalah diagnosis pasti dari sebuah *outcome*¹⁷. Dalam menentukan titik potong ini harus dilakukan tawar-menawar, karena peningkatan sensitivitas akan menyebabkan penurunan spesifisitas, dan juga sebaliknya. Meskipun sensitivitas dan spesifisitas merupakan indikator utama dari sebuah instrumen diagnostika, namun nilai keduanya tidak selalu harus berada pada titik absolut dan tetap. Pada praktiknya, terjadi tarik ulur antara nilai sensitivitas dan spesifisitas, berdasarkan tujuan penggunaan instrumen diagnostika yang sedang diuji. Semakin tinggi nilai sensitivitas, maka akan diiringi dengan penurunan nilai spesifisitas, dan begitu pula sebaliknya. Tarik ulur antara sensitivitas pada berbagai titik potong inilah yang kemudian tergambarkan dalam kurva *ROC*¹⁷. Secara aplikatif, penggunaan titik potong ≥ 4.668 tentu akan cenderung lebih menguntungkan karena dalam proses skrining, persentase negatif palsu yang lebih rendah. Apabila diturunkan titik potong optimum menjadi ≥ 4.62 ; maka nilai sensitivitas yang diperoleh adalah 50,0% dan spesifisitas 40,0%.

Keakuratan sebuah alat/metode tergantung pada seberapa baik metode tersebut memisahkan masing-masing kelompok yang diuji. Akurasi diukur sebagai luas area di bawah kurva (*Area Under Curve*) *ROC*. Luas area sama dengan “1” (satu) mewakili tes yang dapat dianggap sempurna (*excellent*); area seluas $< 0,5$ dianggap sebagai kegagalan alat/metode yang diuji⁵. Hasil analisis *ROC* di atas menunjukkan bahwa luas *AUC* (*Area Under Curve*) adalah sebesar 0,4.724; Titik-titik yang menghubungkan nilai sensitivitas dan 1-spesifisitas sebagian besar berada pada garis diagonal. Shin (2009) menyatakan suatu alat uji diagnostik baru yang tidak memiliki nilai prediktif akan memiliki $AUC < 0,5$ sementara alat dengan kemampuan sempurna untuk memprediksi suatu *outcome* akan memiliki $AUC = 1$ ¹⁸. Dari grafik di atas dapat disimpulkan bahwa instrumen *GPAQ* tidak akurat, artinya instrumen *GPAQ* tidak mampu memisahkan antara orang yang mempunyai kebugaran baik dengan yang memiliki kebugaran tidak baik dengan baku emas *Vo2max*. Tilaki (2013) menyatakan bahwa kurva *AUC* memainkan peran sentral dalam mengevaluasi kemampuan diagnostik sebuah tes untuk membedakan keadaan subyek yang sebenarnya, menemukan nilai *cut off* optimal, serta dapat digunakan untuk membandingkan dua alat diagnostik alternatif yang baru pada subyek yang bersamaan¹⁹.

Alat ukur yang memiliki Nilai Prediktif Positif (NPP) tinggi akan terbukti banyak terjadi positif sejati (*true positif*) dan sedikit positif palsu (*false positif*), sedangkan alat ukur yang memiliki Nilai Prediktif Negatif (NPN) tinggi artinya banyak terjadi negatif sejati dan sedikit negatif palsu. Alat ukur memiliki *validitas* prediktif tinggi, jika skor Nilai Prediktif Positif dan Nilai Prediktif Negatif mendekati 100%²⁰. Nilai Prediksi Positif dan Negatif terhadap tes Aktivitas fisik *GPAQ* pada penelitian ini adalah 75,9% dan 29,3 %. Dengan demikian dari hasil tes Aktivitas fisik dengan menggunakan instrumen *GPAQ* memiliki nilai prediksi positif relatif tidak tinggi, ini berarti bahwa pengukuran Aktivitas fisik menggunakan instrumen *GPAQ* akan menghasilkan positif palsu cukup tinggi, karena nilai NPP hanya 75,9 %, sedangkan NPN hanya sebesar 29,3%; artinya hanya sebagian kecil hasil test menggunakan instrumen *GPAQ* akan menunjukkan hasil negatif sejati dan sebagian besar hasilnya akan berupa negatif palsu²¹.

Nilai sensitivitas instrumen *GPAQ* cukup rendah hanya sebesar 50,0%, ini berarti instrumen ini hanya dapat mendeteksi sebanyak 50,0% responden dengan tingkat Aktivitas fisik yang cukup disertai dengan tingkat *VO2max* dengan kategori bugar. Namun sebaliknya, sebanyak 50,0% responden akan dideteksi sebagai negatif palsu (responden dinyatakan mempunyai Aktivitas fisik “tidak aktif” tapi sebenarnya memiliki tingkat *VO2max* dengan kategori bugar). Dengan hasil demikian, jika kita menggunakan instrumen *GPAQ* untuk mengukur tingkat Aktivitas fisik di populasi, kita akan melewatkan banyak responden yang memiliki tingkat *VO2max* yang “bugar” tetapi mereka masuk ke dalam kategori Aktivitas fisik “tidak aktif”. Spesifisitas *GPAQ* sebesar 56,7%, berarti akan ada sebanyak 43,3% responden akan masuk ke dalam kategori Aktivitas “aktif” akan tetapi sebenarnya berdasarkan pemeriksaan *VO2max* masuk kelompok “tidak bugar”. Angka sebanyak 43,3% disebut dengan nilai positif palsu (1-spesifisitas), bahwa tes ini akan menempatkan banyak responden dalam kelompok aktif meskipun mereka sebenarnya “tidak bugar”. Dalam jargon *epidemiologi* dikatakan bahwa suatu skrining/penapisan dengan sensitivitas yang rendah akan meningkatkan jumlah ‘*negatif palsu*’ sedangkan jika suatu skrining/penapisan memiliki *spesifisitas* yang rendah akan menghasilkan banyak ‘*positif palsu*’²¹.

Pertanyaan apakah instrumen *GPAQ* benar dapat menangkap perilaku Aktivitas fisik responden atau tidak, dicerminkan dari nilai sensitivitas, spesifisitas,

nilai NPP dan NPN . Pengumpulan data Aktivitas fisik menggunakan kuesioner GPAQ ini dengan sistem *recall*, tidak akurat untuk mengukur tingkat Aktivitas fisik seseorang. Namun karena kemudahannya, metode ini masih digunakan dalam penelitian dalam skala besar dan serentak seperti survei skala nasional. Oleh karena itu perlu dicari penggunaan metode/instrumen yang lain untuk mengukur data Aktivitas fisik pada populasi yang besar. Pengukuran dari kumpulan Aktivitas fisik melalui kuesioner maupun *interview* menggunakan instrumen GPAQ sulit dilakukan, karena Aktivitas fisik terjadi pada berbagai tempat yang berbeda misalnya di tempat kerja, saat bepergian, di tempat khusus olahraga/ klub olahraga, pada waktu senggang maupun rekreasi. Lagi pula Aktivitas tersebut dapat bersifat sementara dan bukan merupakan kebiasaan sehari hari responden.

Studi untuk menguji validitas dan reliabilitas instrumen GPAQ telah dilakukan di beberapa negara, seperti yang dilaksanakan oleh *Bull et al* pada tahun 2009, dan baru-baru ini divalidasi kembali di Malaysia, Vietnam dan Amerika Serikat pada tahun 2009^{6,5,22}. Bukti validitas GPAQ di banyak Eropa ternyata masih sedikit dan membutuhkan penelitian lanjutan karena instrumen ini sangat dipengaruhi oleh norma, budaya, tingkat pendidikan serta perbedaan dalam nilai sosial²³.

Dalam praktiknya tidak mudah untuk menanyakan pertanyaan terdapat pada instrumen GPAQ kepada responden. Ada perbedaan interpretasi, baik pada diri trainer saat mengajar maupun kesalahan pada saat wawancara yang dilakukan oleh enumerator. Dilain pihak nilai *VO2max* merupakan nilai baku emas untuk mengukur tingkat kebugaran seseorang. Kelemahan menggunakan pengukuran *Vo2max* ini adalah alat yang digunakan berukuran besar (tidak *portable*), sulit untuk dibawa ke lapangan dan membutuhkan keahlian khusus bagi orang yang mengoperasikannya. Kristanti (2002) menyatakan bahwa sebenarnya sulit untuk menilai Aktivitas sehari hari masyarakat dengan hanya berdasarkan pengakuan semata tanpa pengukuran alat yang valid. Disamping itu Aktivitas pada hari kerja dan akhir pekan selalu akan ada perbedaan. Lebih lanjut Kristanti menyatakan bahwa perlu keahlian khusus untuk menanyakan semua pertanyaan dalam instrumen GPAQ, sehingga pelatihan dan kalibrasi terhadap pewawancara perlu dilakukan². Perlu pedoman yang berisi petunjuk pengisian kuesioner dan teknik/cara wawancara Aktivitas fisik

beserta banyak contoh-contoh yang lebih Konkret. Oleh sebab itu perlu keahlian khusus dan kesabaran tersendiri dalam menanyakan dan menelusuri berbagai kegiatan serta kebiasaan yang dilakukan sehari-hari oleh responden. Selain itu perekrutan enumerator haruslah benar-benar diperhatikan terutama yang mempunyai keahlian khusus untuk dapat menggali kebiasaan Aktivitas fisik sehari-hari dengan jumlah responden yang besar. Selanjutnya yang tidak kalah penting adalah pengembangan instrumen GPAQ yang berkesinambungan untuk mendapatkan instrumen pengukur Aktivitas fisik yang benar-benar valid dan reliabel.

KESIMPULAN

- Nilai *METs* menit/minggu menggunakan instrumen GPAQ pada wanita Bogor usia 25 sampai 54 tahun cenderung *over estimate* dan sebarannya sangat bervariasi, dimana rata-rata *METs* domain pekerjaan sedang sebesar $4.271,69 \pm 2.874,34$; domain perjalanan $1.058,19 \pm 1.058,19$ dan domain rekreasi sedang $181,23 \pm 471,594$.
- Nilai *AUC* instrumen GPAQ sebesar 0,4724, dengan nilai *AUC* < 0,5 maka instrumen GPAQ tidak memiliki kemampuan prediktif untuk membedakan antara orang yang bugar dengan yang tidak bugar.
- Sensitivitas dan spesifisitas Aktivitas fisik menggunakan kuesioner GPAQ dengan pemeriksaan kebugaran dengan *ergocycle* sebagai baku emas berturut-turut adalah sebesar 50,0% dan 43,3%.
- Nilai prediksi positif instrumen GPAQ adalah sebesar 75,9% dan nilai prediksi negatif adalah 23,9%.
- Tingkat Aktivitas fisik dengan instrumen GPAQ tidak berhubungan dengan kebugaran *VO2max* dengan menggunakan *ergocycle*.

SARAN

- Pemeriksaan perilaku Aktivitas fisik dengan kuesioner GPAQ **hanya dapat merekam** Aktivitas fisik responden sehari-hari di populasi.
- Jika instrumen GPAQ masih dipergunakan, dibutuhkan pedoman petunjuk teknis yang rinci dan nilai baku justifikasi untuk mengurangi pengukuran perkiraan yang terlalu besar (*overestimated*) yang disesuaikan dengan kebiasaan dan budaya masyarakat yang ada di Indonesia.

- Untuk mengukur kebugaran dalam populasi bisa digunakan pengukuran *V02max* dengan metode lain yang lebih mungkin misalnya dengan metode *Harvard*, *Cooper* maupun *Balke*.
- Perlu dicari instrumen yang lain untuk menangkap perilaku Aktivitas fisik di populasi, mudah digunakan dan murah tanpa mengurangi validitas dan reliabilitas.

UCAPAN TERIMAKASIH

Terimakasih kami ucapkan kepada Kepala Pusat Penelitian Upaya Kesehatan Masyarakat Kemenkes RI, Bapak Drg Agus Suprpto M.Kes, yang telah menyediakan dana DIPA untuk pendanaan penelitian ini. Juga kepada semua teman yang terlibat dalam penelitian ini antara lain : Dr. Hapsari Tjandarini SKM, M.Kes; Ika Dharmayanti, SKM, Env; Putisari SKM, MSc; Priska Arfines, SGZ, MSc; Sri Mulyati SKM, Mkes; Yudi Kristanto, S.Sos, MKM. Juga kepada semua pihak yang terlibat, pejabat Pemerintah Kota Bogor dan jajarannya serta seluruh warga kelurahan Paledang dan Gudang Bogor tempat lokasi penelitian ini dilaksanakan.

DAFTAR PUSTAKA

Adhitya S.D. 2016. Tingkat Aktivitas Fisik Operator Layanan Internet Mahasiswa Universitas Negeri Yogyakarta. Yogyakarta, Universitas Negeri.

Ariawan I. 2012. Receiver Operating Characteristic Curve. (s.l), (s.n).

Aulia Citra K, et al. 2015. Analisis Indeks Kebugaran Tubuh Seseorang. (s.l), (s.n)

Au TB, et al. 2010. Reliability and validity of the global physical activity questionnaire in Vietnam. *J Phys Act Health*, 7 (3), 410–8.

Budiman I. 2007. Perbandingan Tes Lari 15 Menit Balke dengan Tes Ergometer Sepeda Astrand. Bandung, Univ. Kristen Maranata.

Bull F, Bull FC, Maslin TS, Armstrong T. 2014. Global Physical Activity Questionnaire (GPAQ): Nine Country Reliability and Validity Study.

Ch. M. Kristanti. 2002. Kondisi Fisik Kurang Gerak, Instrumen Pengukuran. *Media Litbang Kesehat*, XII (1).

Cleland CL, et al. 2014. Validity of the Global Physical Activity Questionnaire (GPAQ) in assessing levels and change in moderate-vigorous physical activity and sedentary behaviour, 1–11. WHO. (s.a). STEP wise approach to surveillance (STEPS).. Available at: <http://www.who.int/chp/steps/en/> [accessed 2017 Nov 7].

Dahlan MS. 2009. Penelitian Diagnostik Dasar-Dasar Teoritis dan Aplikasi dengan Program SPSS dan Stata. 1st ed. Jakarta, Salemba Medika.

Gademan MGJ, et al. 2014. A poor association was found between self-reported physical activity and estimated maximal oxygen uptake of sedentary multiethnic women. *J Clin Epidemiol*, 67 (4), 462–7.

Gibney MJ, New SAL, Cassidy A, Vorster HH. 2009. Introduction to Human Nutrition. Second edi. Wiley Blackwell.

WHO. 2004. Global Strategy on diet. Physical activity and health, May.

Hoos T, Marshall S, Arredondo EM. 2013. *NIH Public Access*, 9 (5), 698–705.

Iskaningtyas DA. 2012. Model Prediksi V02 Max Anak Usia 10-11 Tahun Etnis Jawa (Desa Tersobo, Kebumen) Dari Tes Berjalan 1 mil Berdasarkan Jenis Kelamin, Denyut Nadi dan Waktu Tempuh. Jakarta, Universitas Indonesia.

Karimollah Hajian-Tilaki. 2013. Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. *Casp J Intern Med*, 4 (2), 627-35 c.

Lingesh G, et al. 211. Comparing physical activity levels of Malay version of the IPAQ and GPAQ with accelerometer in nurses.

Mumu SJ, Ali L, Barnett A, Merom D. 2017. Validity of the global physical activity questionnaire (GPAQ) in Bangladesh. *BMC Public Health*, 17 (1), 1–10.

Najmah queen. (t.th). Perhitungan Sensitivitas dan Spesifisitas [Internet]. Tersedia pada: <http://metopidfkmsnri.blogspot.co.id/2014/10/sensitivitas-dan-Spesifisitas.html>. [diakses 15 Nopember 2016].

Riki. 2012. Sport and Science. Tersedia pada: http://sportsscience7.blogspot.co.id/2012/09/mengukur-tingkat-kebugaran-jasmani_29.html [diakses 19 Jan 2018].

Ryadi ALS, 2011. Wijayanti T. Dasar-Dasar Epidemiologi. Jakarta, Salemba Medika.

Shin S. 2009. ROC analysis for the evaluation of continuous biomarkers : Existing tools and new features in SAS © 9 . 2. Analysis, 1–15.

Tully MA, et al. 2013. Physical activity and the rejuvenation of Connswater (PARC study): protocol for a natural experimnet investigating the impact of urban regeneration on public health. *BMC Public Health*, 13 (1), 1.

WHO. 2012. Global Physical Activity Questionnaire (GPAQ) Analysis Guide, 1–22. Geneva.