Tania Winzenberg
Kelly A Shaw

Screening for physical inactivity in general practice

A test of diagnostic accuracy

Background
It is unclear what is the best method of accurately identifying physically inactive patients in general practice. This study aimed to compare the performance of different methods of assessing patient physical activity levels in general practice.

Method
Thirteen general practitioners were randomly allocated to perform either their usual assessment, or this with a Lifescripts tool, on consecutive patients. The authors measured patients' physical activity by accelerometer over 1 week, including steps per day, then calculated agreement, kappa specificity, sensitivity, positive and negative predictive value (PV) and ROC characteristics for each assessment method (GPs' usual assessment, Lifescripts tool and steps per day) against the reference standard of accelerometer classification.

Results
Data from 29 patients was included. Agreement between subjective assessments was highest for GPs' usual assessment (agreement 73%; kappa 0.47; p=0.03), which also gave the highest area under the ROC curve (0.75, 95% CI: 0.52–0.98). However, this still had low specificity (67%) and positive PV (63%). Using a cut-off of 7500 steps/day maximised the area under the ROC curve at 0.91 (95% CI: 0.82–1.00), 19.2% greater than GPs' usual assessment.

Conclusion
Measuring steps per day may be a feasible and more effective way to screen for physically inactive patients than self report. A large scale study to confirm these results is necessary.

Keywords: physical activity; screening, sensitivity, specificity; general practice

It is recommended that healthcare providers[1–5] routinely assess their patients' physical activity (PA). However, there is limited evidence to guide general practitioners in their choice of assessment method. Currently, it is suggested that GPs assess PA from patient self report.[6] Previous research by the authors has shown that self report by history taking was the method of choice for Australian GPs.[7] While using self report is acceptable to GPs, it has limitations,[8] including the risk of over-reporting (social desirability bias) and patients finding it difficult to translate their activities into the appropriate intensity grade.[7] An alternative to history taking is the use of questionnaires, but these also rely on self report, and even in the research setting such instruments vary in their effectiveness[8] when compared against objective measures of physical activity such as pedometers or accelerometers.

Two variations on a brief assessment approach have been tested in general practice.[9,10] One which included a comparison with an objective physical activity measure[10] demonstrated poor agreement with accelerometer measures. No study has included a comparison with GPs' usual assessment approach. A potential objective method for measuring PA in general practice is steps per day but the utility of this for classifying people is underexplored and has not been investigated in general practice.[11–13]

This pilot study aimed to investigate the performance of different methods of classifying general practice patients by whether they meet recommended PA levels.

Method

In 2008 the authors recruited a random sample of 13 of the 313 southern Tasmanian GPs.[14] Each was randomly assigned to assess and record the physical activity of up to eight consecutive adult patients using either:

- assessment based on their usual technique alone, or
- assessment based on their usual technique and the Lifescripts assessment tool.[15]

The Lifescripts tool is a 3-item questionnaire that asks about the number of 20 minute bouts of vigorous PA, 30 minute bouts of walking and 30 minute bouts of other moderate PA usually performed in a week. Responders mark an answer on an 8-point scale from 0 to 7+ for each item and these are added to give a total score.

General practitioners gave each patient an invitation to participate in the study with those willing to do so contacting the research team directly. After obtaining written informed consent, we collected the participant's age, gender and demographic information, and measured their weight and height. We obtained written permission from patients to obtain the results of their GP's physical activity assessment.

After the GP's assessment, as the reference method, we measured each participant's physical activity using the well validated ActiGraph GT1M accelerometer.[2,3] The ActiGraph GT1M measures physical activity by detecting vertical acceleration registering a count with g-forces of over 0.03, and measuring the magnitude of the acceleration in increments of 0.03 up to 2.00; it also measures steps taken. Seven consecutive days of physical activity[4] were monitored, measuring counts per day, steps per day and time spent in sedentary (<1.5 metabolic equivalent of task [MET])[16] and low intensity (1.5–2.9 MET), moderate intensity (3.0–5.9 MET) and vigorous intensity

(>6 MET) activity.[17] Participants kept a diary of accelerometer wear. Days were considered valid if the accelerometer was worn for 10 hours or more. Weeks were considered valid if there were five or more valid days of data.

Meeting physical activity recommendations by accelerometer was defined as more than 30 minutes of combined moderate or vigorous physical activity per day for at least 5 days (reference method). A Lifescripts score of 5 or more defined meeting recommended levels using this tool. General practitioners were also asked a yes/no question: 'In your opinion, does this patient meet recommended physical activity levels?' In view of the variation of cut-offs described in the literature, we used an empirical approach to determine the best cut-off point to define meeting physical activity recommendations for steps per day.[11–13,18] The authors identified the lowest mean daily steps recorded in patients who met PA recommendation determined by accelerometer, and incrementally varied the cut-off point around this level, including the conventionally accepted 10 000 steps per day cut-off.[18] An hour of intentional walking is equivalent to about 7300 steps.

Analyses were performed in Stata version 10 (Stata Corporation USA) unless otherwise stated. Agreement between GP assessment (by usual assessment or by Lifescripts) and steps per day at varying cut-offs, and accelerometer determination of meeting of physical activity guidelines, was calculated using the kappa statistic.[6] The authors calculated specificity, sensitivity and positive and negative predictive value with 95% confidence intervals using the Vassar Clinical calculator (http://faculty.vassar.edu/lowry/VassarStats.html) then also compared the receiver operator curve (ROC) characteristics of the approaches. The ROC plots the true positive rate (sensitivity) against the false positive rate (1-specificity) for each value of a given variable, in this case GPs' assessment by each of the two methods. The area under the curve (AUC) is a measure of the variable's ability to discriminate between people with or without a particular condition (in this case physical inactivity). A test which does not discriminate will have an AUC of 0.5 (diagonal); a perfect test will have an AUC of 1.0.

Ethics approval was received from the Human Research Ethics Committee (Tasmania) Network (reference number H9459).

Results

Figure 1 describes the flow of study participants and the available data for comparisons. In total, 25 patients had GP assessment data and accelerometer counts, 11 had Lifescripts data and accelerometer counts, 10 had GP assessment with Lifescripts data and accelerometer counts, and 29 had accelerometer counts and steps per day. *Table 1* gives the characteristics of the study participants.

Table 2 gives the level of agreement and kappa statistic between each method of assessment and accelerometer assessment of whether recommended physical activity levels were met. The level of agreement between subjective assessments and objective accelerometer using the Lifescripts tool reduced agreement to 60% (kappa 0.09; p>0.05).

Steps per day were highly correlated with accelerometer activity measures (r=0.58, p=0.008; r=0.85, p<0.0001; and r=0.95, p<0.0001 for minutes of vigorous activity, moderate activity and counts respectively). *Figure 2* shows the AUC ROC curves with different step cut-offs. The cut-off maximising AUC ROC was 7500 steps per day (AUC ROC=0.91 [95% CI: 0.82–1.00]).

Table 3 gives the specificity, sensitivity and positive and negative predictive value for each method. For the subjective measures, the best performance characteristics were achieved by the use of GPs' usual assessment without Lifescripts. Specificity, sensitivity and positive and negative predictive value were high for both step cut-offs of 10 000 and 7500 steps per day.

This is reflected in the ROC curve comparisons in *Figure 3* and *4*. The ROC curve plots the true positive rate (sensitivity) against the false positive rate (1-specificity) for each value of a given variable. The AUC ROC for GPs' assessment if

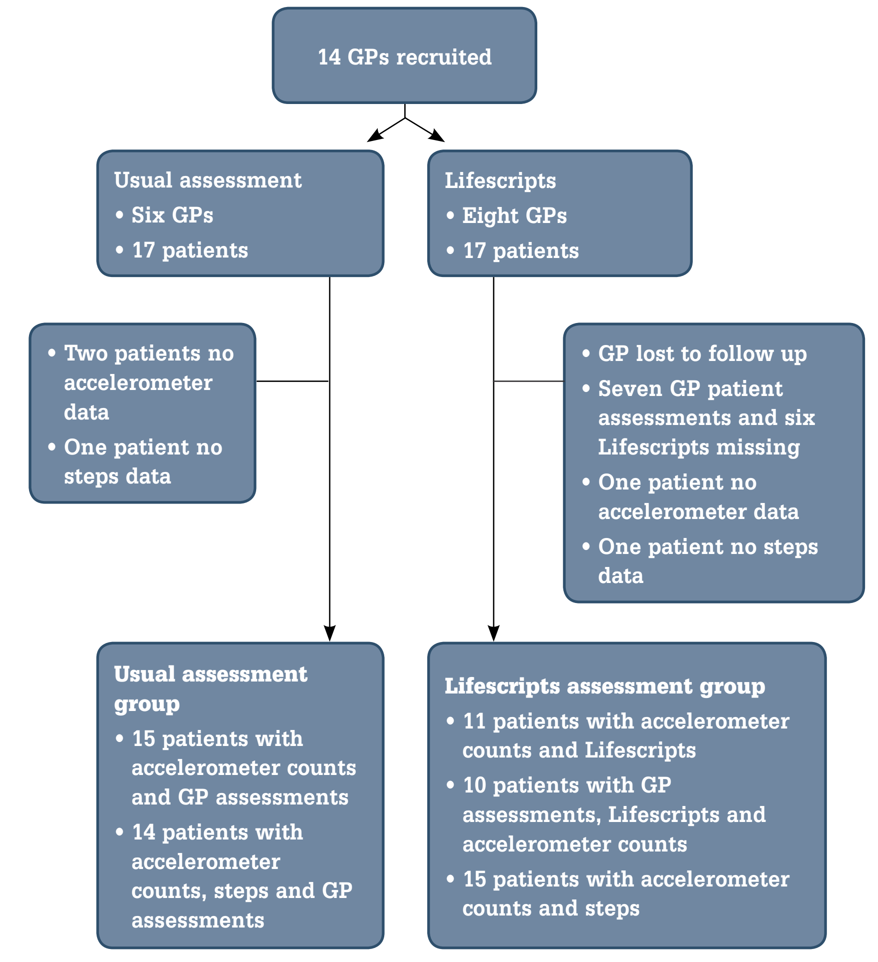

Figure 1. GP and patient participation showing the data available for each comparison in the usual assessment and Lifescripts assessment groups

Lifescripts was used was 0.56 (95% CI: 0.04–1.00) compared to 0.75 (95% CI: 0.52–0.98) for GPs using their assessment alone (*Figure 3*).

Figure 4a and *4b* compare the ROC curves of different assessment approaches compared to accelerometer assessment. *Figure 4a* compares the curves for 7500 steps per day and GP assessment. There was a trend toward statistical significance for the difference of 19.2% in favour of 7500 steps per day (p=0.069, n=23 in analysis). *Figure 4b* compares 7500 steps per day and 10 000 steps per day. The cut-off maximising AUC ROC was 7500 steps per day (AUC ROC=0.91 [95% CI: 0.82, 1.00]). However, the difference between the curves of 6.6% in favour of the 7500 cut-off was not statistically significant (n=29 for comparison).

Discussion

These findings have immediate implications for the assessment of physical activity in general practice and provide important preliminary evidence to guide GPs' decisions on what assessment approach might be most effective. This study is the first to determine the diagnostic accuracy of using steps per day to assess whether general practice patients meet recommended physical activity levels and shows that steps per day appears a potentially viable alternative to patient self report.

By contrast, subjective assessments have at best moderate agreement with objective accelerometer assessment. All three subjective methods performed relatively poorly as a screening test. Notably, the best agreement using a subjective approach was achieved by GPs using their own assessment technique, and this was not improved by the additional use of Lifescripts. In fact, the Lifescripts score alone showed the lowest level of agreement with accelerometer assessment, and also performed worst as a screening tool.

Steps per day performed well with a cut-off of 7500 steps per day providing the best overall performance. While this was not significantly different statistically from the cut-off of 10 000 steps frequently cited as a recommended level, the difference may be of clinical importance. If patients exceeded an average of 7500 steps per day, in 8 out of 10 cases they would be correctly classified as meeting recommendations and in only 2 out of 10 cases would patients and GPs be falsely reassured. With a cut-off of 10 000 steps per day, only 1 in 10 patients would be falsely reassured, but in 1 in 6 cases, patients would be falsely identified as not meeting recommended guidelines. This could result in GPs spending valuable consulting time on counselling when this was not needed. Concerns have also been raised that using a more stringent cut-off might create a perception of setting an unattainable goal at an individual intervention

Table 1. Characteristics of study participants

Characteristic*	All**	Lifescripts	Usual assessment
Age (years), mean (SD)+	56.2 (15.8)	59.0 (11.3)	53.4 (19.3)
Gender (male)	6 (18)	3 (17)	3 (17)
Height (m), mean (SD)	1.66 (0.08)	1.66 (0.07)	1.67 (0.08)
Weight (kg), mean (SD)	77.3 (16.2)	76.4 (14.3)	78.3 (18.3)
Employment			
≥20 hours/week	19 (56)	9 (53)	10 (59)
<20 hours/week	3 (9)	2 (12)	1 (6)
Unemployed	12 (35)	6 (35)	6 (35)
Main financial provider, unemployed or pension	7 (21)	5 (29)	2 (12)
Education			
High school or less	8 (24)	3 (18)	5 (29)
Year 11–12	6 (18)	4 (24)	2 (12)
Tertiary	20 (59)	10 (59)	10 (59)
Average accelerometer counts per day, mean (SD) [range] (n=31)	294 984 (175 496) [61 035–784 947]	295 008 (152 735) [61 035–674 433]	294 958 (202 492) [80 871–784 947]
Average steps per day, mean (SD) [range] (n=29)	8036 (4795) [1510–23 931]	7800 (3889) [1510–13137]	8290 (5752) [2675–23931]
Lifescripts score, mean (SD) [range]	NA	5.4 (3.4) [1–10]	NA

* n (%) unless otherwise stated; ** n=34 in total and 17 in Lifescripts and usual assessment groups unless otherwise stated; SD = standard deviation

Table 2. Agreement and kappa for comparisons between GPs' usual assessment, use of Lifescripts, steps per day (7500 and 10 000 cut-offs) and accelerometer measures of physical activity

	Lifescripts			Accelerometer		
	Agreement	Expected agreement	kappa	Agreement	Expected agreement	kappa
All GP assessments	–	–	–	68%	51%	**0.35***
Accelerometer	55%	48%	0.13	–	–	–
GP assessment alone	–	–	–	73%	49%	**0.47***
GP assessment with Lifescripts	–	–	–	60%	56%	0.09
Achieved 7500 steps/day	–	–	–	90%	50%	**0.79****
Achieved 10 000 steps/day	–	–	–	86%	53%	**0.71****

Note: Bold denotes statistical significance; *p=0.03; **p≤0.0001

Table 3. Specificity, sensitivity and positive and negative predictive values of steps per day compared to accelerometer assessment for detecting sufficient physical activity

	Specificity %	Sensitivity %	+ PV%	– PV%
GP assessment (all) (n=25)	65 (39, 85)	75 (36, 96)	50 (22, 78)	85 (54, 97)
Lifescripts (n=11)	50 (17, 83)	67 (13, 98)	33 (6, 76)	80 (30, 99)
GP assessment alone (n=15)	67 (31, 91)	83 (36, 99)	63 (26, 90)	86 (42, 99)
GP assessment with Lifescripts (n=10)	63 (26, 90)	50 (3, 90)	25 (1, 78)	83 (36, 99)
Achieved 7500 steps/day	82 (56, 95)	100 (70, 100)	80 (51, 95)	100 (73, 100)
Achieved 10 000 steps/day	94 (69, 99.7)	75 (43, 93)	90 (54, 99)	84 (60, 96)

Note: Figures given in brackets are 95% confidence intervals; PV = predictive value

level, potentially undermining the intent of setting public health guidelines.[18] A lower cut-off is supported by recent data that in older people and young women, achieving 5000 steps per day is associated with substantially lower prevalence of adverse cardiometabolic health indicators, with more modest reductions in prevalence with increases in steps above this level.[19] It is also supported by data in other populations.[11–13] However, the decision about which is the most appropriate cut-off is ultimately based on a value judgment of the relative harms of each approach. Commercially available pedometers could provide a cheaper (basic models cost from around $30) and more accessible way to measure steps per day to classify patients' activity levels in general practice than accelerometers, however further research is required to confirm this.

The more subjective screening approaches are simple and cheap but there is insufficient evidence that they classify patients well enough. All assessment approaches had good negative predictive value – over 80% of people who were identified by their GP as having insufficient physical activity had this confirmed by accelerometer. However, of greater concern are the low positive predictive values. For GP assessment alone this was 63%, which means that 1 in 3 patients who GPs assess as meeting guidelines in fact do not, and therefore remain at risk from diseases related to physical inactivity. For Lifescripts assessment, this applies to 2 in 3 such patients. Such false reassurance would not be acceptable in a diagnostic test for patients and should not be acceptable in physical activity screening processes.

The poor performance of Lifescripts in this pilot study brings into question the continued encouragement of GPs to use this or similar screening tools. By contrast, the performance of steps per day suggest that this should be further investigated as a method to use in general practice. Other literature supports this view. A single item screening question: 'As a rule, do you do at least half an hour of moderate or vigorous exercise (such as walking or sport) on five or more days of the week?' was compared to assessment by a lengthy validated questionnaire (the NZPAQ-LF9) using different criteria. Despite using a reference standard that was also self report, agreement was only moderate (kappa 0.46–0.56) and lower than in our study (kappa 0.79 for 7500 steps/day), as were negative predictive value and sensitivity (45–68% and 68–77% respectively). Another study examined using 2- or 3-item structured questioning to measure the number of bouts of vigorous intensity activity and of walking or moderate intensity activity in a usual week to determine whether patients were inactive.[10] Comparison measures were accelerometry and a validated self report measure, the Active Australia Questionnaire. Similarly to Rose's study,[9] even against a self report standard, agreement of a brief assessment was only moderate (kappa = 46.7% and 38.7% for 2-item and 3-item questioning respectively). Agreement was very poor with accelerometer measures (kappa 18.2% and 24.3% for 2-item and 3-item questioning respectively).

Study limitations

As a pilot, this study was not powered to find statistically significant differences between assessment methods. However, it found similar levels of agreement and kappa to a previous, larger study and the magnitude of the difference AUC ROC were substantial enough to be of clinical significance. Nonetheless, the results must be interpreted with caution and require confirmation in a larger sample of general practice patients before definitive recommendations are possible. Steps and counts were both measured by Actigraph GT1M accelerometer and were highly correlated. However, the correlation in our study is similar to that previously reported between pedometers and accelerometer counts measured with separate devices (median r=0.86),[20] therefore it is unlikely that this agreement is inflated to a significant extent. It seems a reasonable expectation that simpler and less expensive pedometers would be adequate for use in clinical practice, as they are in research settings,[20,21] but this nonetheless requires confirmation using pedometer brands that are readily and cheaply available in the community. The most appropriate cut-off of steps

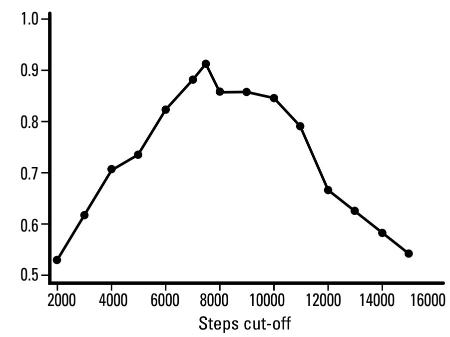

Figure 2. AUC ROC using different steps/day cut-offs as threshold for having met physical activity requirements

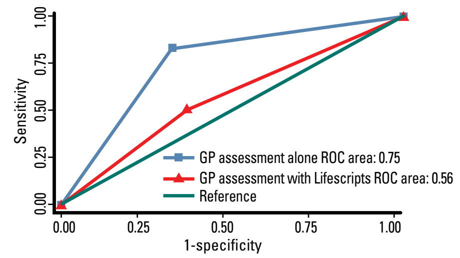

Figure 3. Comparison of AUC ROC curves for GPs using their assessment alone or using their own assessment and Lifescripts against accelerometer assessment

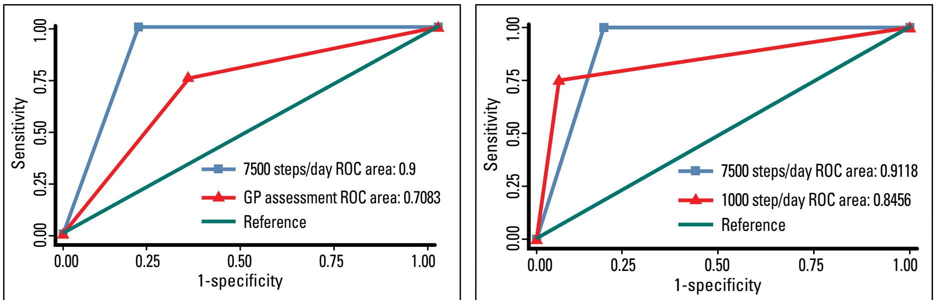

Figure 4a. A cut-off of 7500 steps/day and GP assessment compared to accelerometer assessment (reference)

Figure 4b. 7500 steps/day cut-off and 10 000 steps/day cut-off compared to accelerometer assessment

Figure 4. Comparison of ROC curves for different assessment methods compared to accelerometer assessment

per day for use in Australian general practice also needs to be confirmed in other general practice populations. The low level of agreement between GPs' assessment, Lifescripts and accelerometer measures may be in part due to the difference in time periods in which the measures were taken. However, there is substantial evidence demonstrating the limitations of subjective physical activity measures in situations where this does not apply[8] and our data still suggests that steps per day is likely to be superior to subjective methods. For privacy and logistical reasons, the response rate of patients could not be determined as the authors could not ascertain how many patients each GP asked to participate. However, the sample contained participants with a wide range of physical activity levels (*Table 1*), therefore the results should be generalisable to people of varying activity levels.

Conclusion

The results of this study suggest that pedometers may be a feasible and more effective way to screen for physically inactive patients in general practice than self report, particularly using short assessment tools such as Lifescripts. However, a large scale study to confirm these results and to further explore options for identifying patients at risk due to low physical activity levels in general practice is essential.

Authors

Tania Winzenberg MBBS, PhD, FRACGP, MMedSc(ClinEpi), is Senior Research Fellow, Menzies Research Institute, University of Tasmania. tania.winzenberg@utas.edu.au

Kelly A Shaw MBBS, MPH, PhD, FRACGP, FAFPHM, is Postdoctoral Research Fellow, Southern Cross University, Coff's Harbour, New South Wales.

Conflict of interest: none declared.

Acknowledgments

This work was supported by a Royal Australian College of General Practitioners Cardiovascular Research Grants in General Practice grant and a National Health and Medical Research Council General Practice Training Fellowship to Tania Winzenberg. We also wish to acknowledge the role of the participating GPs and their patients in making the study possible.

References

1. Fletcher GF, Blair SN, Blumenthal J, et al. Statement on exercise. Benefits and recommendations for physical activity programs for all Americans. A statement for health professionals by the Committee on Exercise and Cardiac Rehabilitation of the Council on Clinical Cardiology, American Heart Association. Circulation 1992;86:340–4.
2. Summary of the Second Report of the National Cholesterol Education Program Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel II). JAMA 1993;269:3015–23.
3. National Institutes of Health Consensus Development Panel on Physical Activity and Cardiovascular Health. Physical activity and cardiovascular health. JAMA 1996;276:241–6.
4. Medical aspects of exercise: benefits and risks. London: Royal College of Physicians, 1992.
5. The Royal Australian College of General Practitioners. Guidelines for preventive activities in general practice. South Melbourne: The RACGP, 2005.
6. Harris M. The Royal Australian College of General Practitioners National Standing Committee Quality Care. SNAP: smoking, nutrition, alcohol and physical activity: a population health guide to behavioural risk factors in general practice. Melbourne: The RACGP, 2004.
7. Winzenberg T, Reid P, Shaw K. Assessing physical activity in general practice: a disconnect between clinical practice and public health? Br J Gen Pract 2009;59:e359–67.
8. Sallis J, Salens B. Assessment of physical activity by self-report: status, limitations and future directions. Res Q Exerc Sport 2000;71:1–14.
9. Rose SB, Elley CR, Lawton BA, et al. A single question reliably identifies physically inactive women in primary care. N Z Med J 2008;121:U2897.
10. Smith BJ, Marshall AL, Huang N. Screening for physical activity in family practice: evaluation of two brief assessment tools. Am J Prev Med 2005;29:256–64.
11. Macfarlane DJ, Chan D, Chan KL, et al. Using three objective criteria to examine pedometer guidelines for free-living individuals. Eur J Appl Physiol 2008;104:435–44.
12. Ramirez-Marrero FA, Rivera-Brown AM, Nazario CM, et al. Self-reported physical activity in Hispanic adults living with HIV: comparison with accelerometer and pedometer. J Assoc Nurses AIDS Care 2008;19:283–94.
13. Rowe DA, Kemble CD, Robinson TS, et al. Daily walking in older adults: day-to-day variability and criterion-referenced validity of total daily step counts. J Phys Act Health 2007;4:434–46.
14. General Practice South. 2007 Directory of General Practitioners and Specialists in Private Practice. Hobart: General Practice South, 2007.
15. Kinect Australia for the Lifescripts consortium. Lifescripts physical activity assessment tool. Canberra: Commonwealth of Australia, 2005.
16. Matthews CE. Calibration of accelerometer output for adults. Med Sci Sports Exerc 2005;37:512–22.
17. Freedson PS, Melanson E, Sirard J. Calibration of the Computer Science and Applications, Inc. accelerometer. Med Sci Sports Exerc 1998;30:777–81.
18. Tudor-Locke C, Hatano Y, Pangrazi RP, et al. Revisiting 'how many steps are enough?' Med Sci Sports Exerc 2008;40:S537–43.
19. Schmidt MD, Cleland VJ, Shaw K, et al. Cardiometabolic risk in younger and older adults across an index of ambulatory activity. Am J Prev Med 2009;37:278–84.
20. Tudor-Locke C, Williams JE, Reis JP, et al. Utility of pedometers for assessing physical activity: convergent validity. Sports Med 2002;32:795–808.
21. Tudor-Locke C, Williams JE, Reis JP, et al. Utility of pedometers for assessing physical activity: construct validity. Sports Med 2004;34:281–91.

correspondence afp@racgp.org.au