

Prevention and Health

Wassenaarseweg 56

P.O. Box 2215

2301 CE Leiden

The Netherlands

www.tno.nl

T +31 71 518 18 18

F +31 71 518 19 01

info-zorg@tno.nl

TNO report

KvL/GB 2010.007

Development and evaluation of neural network models to identify types of activity using accelerometers, global positioning systems and heart rate monitors

Date	January 2010
Author(s)	F. Galindo Garre S.I. de Vries M. Engels L.H. Engbers
Assignor	Ministry of Health, Welfare and Sport
Project number	031.20097/01.13
Number of pages	25 (incl. appendices)
Number of appendices	1

All rights reserved. No part of this report may be reproduced and/or published in any form by print, photoprint, microfilm or any other means without the previous written permission from TNO.

All information which is classified according to Dutch regulations shall be treated by the recipient in the same way as classified information of corresponding value in his own country. No part of this information will be disclosed to any third party.

In case this report was drafted on instructions, the rights and obligations of contracting parties are subject to either the Standard Conditions for Research Instructions given to TNO, or the relevant agreement concluded between the contracting parties. Submitting the report for inspection to parties who have a direct interest is permitted.

Samenvatting

Achtergrond

Er wordt de laatste jaren in toenemende mate gebruik gemaakt van objectieve methoden om lichamelijke activiteit te meten. Deze trend wordt mede ingegeven door de huidige technologische ontwikkelingen waardoor het mogelijk is geworden om tegen steeds lagere prijzen steeds nauwkeuriger en langer te meten. Een veelgebruikte objectieve methode om lichamelijke activiteit te meten zijn versnellingsmeters. Versnellingsmeters leveren objectieve informatie over de frequentie, intensiteit en de tijdsduur van lichamelijke activiteit. De meeste versnellingsmeters geven echter geen informatie over het type lichamelijke activiteit. Tevens onderschatten zij de intensiteit van een aantal lichamelijke activiteiten, waaronder fietsen.

In 2008-2009 heeft TNO Kwaliteit van Leven de mogelijkheden verkend om met behulp van geavanceerde statistische modellen verschillende typen lichamelijke activiteit van volwassenen te classificeren op basis van versnellingsmeterdata. Uit dit onderzoek kwam naar voren dat artificiële neurale netwerk (ANN) modellen redelijk goed in staat zijn het type lichamelijke activiteit van volwassenen te classificeren op basis van versnellingsmeterdata van de heup of de enkel. Met deze modellen werd ruim 60% van de activiteiten zitten, staan, lopen, fietsen en traplopen correct geclassificeerd. De modellen waren echter niet in staat om op basis van versnellingsmeterdata onderscheid te maken tussen activiteiten die op verschillende intensiteiten (snelheden) werden uitgevoerd. Zo waren de ANN modellen niet in staat onderscheid te maken tussen rustig lopen of fietsen en stevig doorlopen en -fietsen. Stevig doorlopen en -fietsen werd hierbij veelal geclassificeerd als rustig lopen of fietsen. Het was daarnaast niet goed mogelijk om met deze modellen de activiteit zitten te onderscheiden van de activiteit staan. Mogelijk kunnen de modellen verbeterd worden door gegevens over de snelheid of de intensiteit van een activiteit gemeten met Global Positioning Systems (GPS) of hartslagmeters aan de modellen toe te voegen.

Doel

De doelstelling van het huidige onderzoek is om meer inzicht te krijgen in de toegevoegde waarde van GPS data (snelheid) en hartslagdata (intensiteit) in artificiële neurale netwerk modellen gebaseerd op versnellingsmeterdata voor het classificeren van het type lichamelijke activiteit van volwassenen. Het onderzoek beoogt tevens inzicht te verschaffen in welke combinatie van sensoren idealiter gebruikt zou moeten worden in onderzoek naar de lichamelijke activiteit van personen in termen van frequentie, intensiteit, tijdsduur en type activiteit.

Methode

Zesentwintig proefpersonen in de leeftijd van 22-61 jaar (8 mannen; 18 vrouwen) hebben een gestandaardiseerd beweegprotocol afgelegd waarbij zij een één-assige ActiGraph versnellingsmeter en GPS op de heup hebben gedragen, een één-assige ActiGraph versnellingsmeter om de enkel en een Polar hartslagmeter om de borst. Het beweegprotocol (50 minuten) bestond uit de volgende activiteiten: zitten, staan, lopen, fietsen en traplopen. Alle activiteiten, met uitzondering van traplopen, zijn zowel binnen als buiten uitgevoerd. Alle proefpersonen hebben op twee zelfgekozen snelheden buiten gelopen en gefietst, eenmaal op een rustig tempo en eenmaal op een steviger tempo. Vervolgens zijn op basis van de versnellingsmeterdata twee ANN modellen ontwikkeld. Het eerste ANN model is gebaseerd op versnellingsmeterdata van de heup, het tweede model is gebaseerd op versnellingsmeterdata van de enkel. In deze

modellen zijn de volgende kenmerken van het versnellingsmetersignaal opgenomen: 10de, 25ste, 75ste, and 90ste percentiel waarde, absolute afwijking, variatiecoëfficiënt, en lag-one autocorrelatie. Vervolgens is bekeken in hoeverre de prestatie van de modellen in het correct classificeren van verschillende activiteiten verbeterde door het toevoegen van GPS data en hartslagdata aan de modellen.

Resultaten

Het model gebaseerd op versnellingsmeterdata van de heup en het model gebaseerd op versnellingsmeterdata van de enkel classificeerden 61,8%, respectievelijk 56,5% van de activiteiten correct. Het eerste model presteerde beter als het ging om het correct classificeren van de activiteiten lopen (binnen), stevig doorlopen (buiten) en staan, terwijl het model gebaseerd op versnellingsmeterdata van de enkel beter in staat was om stevig doorfietsen te classificeren. Beide modellen waren goed in staat om zitten te classificeren (> 90% correct geclassificeerd).

Na het toevoegen van GPS data en hartslagdata aan de modellen, classificeerde het model gebaseerd op versnellingsmeterdata van de heup 69,7% van de activiteiten correct en het model gebaseerd op versnellingsmeterdata van de enkel 60,7%. De activiteiten lopen en fietsen, met name stevig doorlopen en –fietsen, werden met deze modellen vaker correct geclassificeerd dan met de modellen die alleen gebaseerd waren op versnellingsmeterdata. Echter voor de activiteiten zitten en staan daalde het percentage correct geclassificeerde activiteiten aanzienlijk na het toevoegen van GPS data en hartslagdata aan de modellen.

Conclusie

Op basis van de resultaten kan geconcludeerd worden dat de ontwikkelde ANN modellen redelijk goed in staat zijn om op basis van versnellingsmeterdata, al dan niet aangevuld met GPS data en hartslagdata, een aantal lichamelijke activiteiten van volwassenen te classificeren. Het aantal keer dat de lichamelijke activiteiten niet correct geclassificeerd worden, is echter ook nog redelijk hoog. Uit de resultaten kwam verder naar voren dat het includeren van GPS data en hartslagdata in ANN modellen gebaseerd op versnellingsmeterdata een beperkte toegevoegde waarde hebben voor het correct classificeren van het type lichamelijke activiteit van volwassenen. Het wordt dan ook niet aanbevolen proefpersonen in grootschalig onderzoek naast een versnellingsmeter ook een aparte GPS of een hartslagmeter te laten dragen als het gaat om het meten van de lichamelijke activiteit van personen in termen van frequentie, tijdsduur en type activiteit.

Aanbevelingen

Het verdient aanbeveling te onderzoeken in hoeverre de ontwikkelde modellen verbeterd kunnen worden door andere kenmerken van het versnellingsmetersignaal als input variabelen van het ANN het model te hanteren, zoals kenmerken die de overgang tussen verschillende activiteiten kunnen weergeven of kenmerken die het cyclische karakter van bepaalde activiteiten zoals fietsen kunnen weergeven. Het wordt tevens aanbevolen te onderzoeken of andere lichamelijke activiteiten, zoals huishoudelijke activiteiten en diverse sporten of bepaalde activiteitscategorieën (bijvoorbeeld de categorieën sedentaire, licht inspannende en matig tot zwaar inspannende activiteiten) kunnen worden onderscheiden op basis van versnellingsmeterdata. De ontwikkelde ANN modellen zouden ook getest kunnen worden voor het classificeren van verschillende typen lichamelijke activiteit van andere leeftijdsgroepen zoals kinderen en ouderen. Tot slot wordt aanbevolen ook de mogelijkheden van andere geavanceerde

statistische patroonherkenningmethodieken zoals boommodellen en wavelets¹ analyse te verkennen voor het classificeren van verschillende typen lichamelijke activiteiten op basis van versnellingsmeterdata.

¹ Wavelets, letterlijk kleine golven, vormen basisfuncties voor een functietransformatie. Zij zijn een alternatief voor de klassieke Fourier-analyse en ze zijn dan ook bijzonder geschikt in signaal- of beeldverwerking. Een wavelet decompositie laat toe het signaal te ontbinden in deelcomponenten met verschillende resolutie. Een belangrijke toepassing van wavelets is het verwijderen van ruis.

Summary

Background

Due to the limitations (e.g. social desirability) that are inherent to subjective methods that measure physical activity, currently more and more objective methods are being used for this purpose. Technological developments make it easier to measure physical activity more accurately with objective methods. Moreover, many of these new technologies are getting cheaper. Accelerometers have, in recent times, become the method of choice in physical activity research. These lightweight, unobtrusive devices provide objective information about the frequency, intensity, and duration of physical activity. However, most accelerometers cannot register, or they underestimate, the intensity of activities other than walking. Furthermore, accelerometers do not provide information about the type of activity people engage in. Recently, an alternative strategy for coping with some of the weaknesses of accelerometers has been suggested: the use of more sophisticated statistical techniques for analyzing accelerometer data, such as artificial neural network (ANN) models.

A previous study of TNO Quality of Life in 2008-2009 has shown that relatively simple ANN models perform reasonably well in identifying the type, but not the speed of the activity of adults from accelerometer data. The accuracy of these models may improve by including more information about the intensity of activities in the models, for example by adding heart rate data or by adding information about the velocity of activities from global positioning systems (GPS). These features may discriminate better between low and high speeds for the same activities, such as regular walking and brisk walking.

Purpose

The purpose of the present study was to examine whether the accuracy of the developed ANN models can be improved by including more data on the intensity of activities in the models, i.e., by adding heart rate data or velocity data to the models.

Methods

Twenty-six healthy subjects (8 males, 18 females) performed a controlled sequence of activities: sitting, standing, climbing stairs, and walking and cycling at two self-paced speeds. All subjects wore a uni-axial ActiGraph accelerometer on the hip and the ankle, a GPS on the hip, and a Polar heart rate monitor around the chest. First, two ANN models were fitted for the hip and the ankle accelerometer data respectively. Next, new ANN models were developed by adding heart rate or velocity data to the input variables. In the ANN models the following accelerometer signal characteristics were used: 10th, 25th, 75th, and 90th percentiles, absolute deviation, coefficient of variability, and lag-one autocorrelation.

Results

The ANN model based on hip accelerometer data correctly classified activity type 61.8% of the time, while the model based on ankle accelerometer data attained a percentage of 56.5%. The first model was better able to classify the activities walking (indoors), brisk walking (outdoors) and standing still, while the latter model was better able to classify brisk cycling. Both models performed well when classifying sitting (> 90% correctly classified).

After the inclusion of GPS (velocity) data and heart rate data the model based on hip accelerometer data correctly classified the type of activity 69.7% of the time, while the

model based on ankle accelerometer data correctly classified the activities 60.7% of the time. Walking and cycling, especially brisk walking and cycling, were better classified with these models than with models based on accelerometer data alone. However, the models performed worse when classifying sitting and standing after including GPS data and heart rate data to the models.

Conclusion

It can be concluded that the developed ANN models are able to predict a number of physical activities of adults relatively well based on accelerometer data. Nevertheless, the number of times that an activity is classified incorrectly is still quite high. Moreover, the addition of GPS and/or heart rate data did not improve the percentage of correctly classified activities of the models. Based on these data, it is not recommended to use a combination of single sensors in large scale epidemiological studies on physical activity with the purpose to improve objective measurement and classification of physical activity.

Recommendations

Future research is needed on:

- The improvement of the developed ANN models by including other features of the accelerometer signal, such as features that mark the transition between activities or features representing the cyclic characteristics of movements (i.e. cycling);
- The classification of household activities and sports with the developed ANN models;
- The use of the developed ANN models to classify physical activities of other age groups such as children and elderly;
- The usability of other advanced pattern-recognition-based methods in classifying activities based on accelerometer data such as tree-models and wavelet analyses.

Contents

	Samenvatting.....	2
	Summary	5
1	Introduction.....	8
2	Methods	11
2.1	Subjects and protocol.....	11
2.2	Statistical analyses	11
3	Results.....	14
3.1	General results	14
3.2	Activity classification	16
4	Conclusion and recommendations.....	19
4.1	Discussion and conclusion.....	19
4.2	Recommendations.....	20
5	References.....	21
	Appendices	
	A Contingency tables	

1 Introduction

Physical activity is a very complex behavior. It is usually defined as “any bodily movement produced by skeletal muscles that result in energy expenditure” (American College of Sport Medicine, 2000). The accurate assessment of physical activity is essential for the examination of trends in physical activity over time, the improvement of our understanding of the dose-response relationship between physical activity and health, the identification of determinants of physical activity, the detection of people at risk, and the evaluation of intervention strategies designed to increase physical activity (Figure 1)(De Vries, 2009; Welk, 2002).



Figure 1. Conceptual model of links between physical activity assessment and different domains of physical activity research (De Vries, 2009)

There are numerous methods available to assess physical activity, such as double-labeled water, direct observation, calorimetry, heart rate monitors, pedometers, accelerometers, and self-reports. Each method assesses different aspects of physical activity. Physical activity can be expressed in terms of energy expenditure (kcal), external workload (Watt), units of movement (counts), frequency (days per week), intensity (metabolic equivalents), duration (minutes), and type of activity (De Vries, 2009). Physical activity has traditionally been measured with self-reports. Self-reports are easily-administered, low-cost methods which provide information about the self-perceived frequency, intensity, duration, and types of activity people engage in during specific periods of time within specific settings. However, self-reports tend to overestimate time spent in vigorous physical activities and to underestimate the time spent in unstructured daily physical activities, such as walking (Armstrong & Welsman, 2006; Richardson et al., 2001; Tudor-Locke & Myers, 2001). By contrast with self-reports, accelerometers are not influenced by recall bias or social desirability. These lightweight, unobtrusive devices provide objective information about the frequency,

intensity, and duration of physical activity. Accelerometers have therefore, in recent times, become the method of choice in physical activity research. However, most accelerometers are not waterproof (De Vries et al., 2009), hence certain water related activities, such as swimming are not registered. In addition, accelerometers perform best in registering ambulatory activities; they cannot register, or they underestimate, the intensity of certain cyclic activities such as climbing stairs, lifting or carrying weights, cycling, and rowing (Levine et al., 2001; Rowlands et al., 2004; Tudor-Locke & Myers, 2001). Furthermore, accelerometers do not provide information about the type of activity people engage in.

Recently, an alternative strategy for coping with some of the weaknesses of accelerometers has been suggested (Esliger et al., 2005): the use of more sophisticated statistical techniques for analyzing accelerometer data, such as approaches based on pattern recognition such as quadratic discriminant analysis (Pober et al., 2006), decision trees (Bonomi et al., 2009) or artificial neural network models (Rothney et al., 2007; Staudenmayer et al., 2009). By contrast with more traditional approaches to handling accelerometer data, such as reporting the average daily activity level or the time spent at different intensity levels, these more advanced statistical techniques aim to detect different types of activity at each time point. Approaches based on pattern recognition are used to distinguish between activities that produce similar total acceleration over time but different energy expenditure, or between activities that have similar energy expenditure but different total acceleration over time. In this way, time spent at different intensity levels can be estimated more accurately. In addition, information of this kind provides an insight into the contribution of different types of activity to total physical activity or total energy expenditure. Furthermore, pattern-recognition-based approaches can be useful in distinguishing between periods of sedentary activities, periods of sleeping, and periods when the accelerometer is not worn.

To date, most of the pattern-recognition-based algorithms and models are based on accelerometer data from a limited number of laboratory activities (Pober et al., 2006; Rothney et al., 2007). It is questionable whether laboratory-derived algorithms and models can be applied to free-living activities. Furthermore, in most studies to date, a single device was placed on subjects' hips (Pober et al., 2006; Rothney et al., 2007; Staudenmayer et al., 2009). A model based on hip accelerometer data may not correctly classify certain physical activity types (Staudenmayer et al., 2009). This was confirmed in a recent study of De Vries et al. (2009; submitted 2009). In this study three relatively simple artificial neural network (ANN) models were developed, compared and evaluated for the purposes of classifying nine physical activity types based on data from the hip accelerometer, ankle accelerometer or both. The results showed that the hip model produced a better classification of the activities brisk walking and going down the stairs, whereas the ankle model was better able to correctly classify the activities regular walking, brisk cycling and going up the stairs. In general, the model based on the hip accelerometer data and the model based on the ankle accelerometer data correctly classified the type of activity 60,3% and 64,2% of the time respectively, while the model based on the data from both sensors correctly classified the activities 69,1% of the time. All three models performed reasonably well (> 80% correctly classified) when classifying walking, cycling, and sitting. However, the models performed worse when classifying climbing stairs and standing still and when discriminating between two self-paced speeds of walking and cycling. The purpose of the present study was to examine whether the accuracy of the developed ANN models can be improved by including more data on the intensity of activities in the models, i.e., by adding heart rate

data or by adding information about the velocity of activities using global positioning systems (GPS)

2 Methods

2.1 Subjects and protocol

A convenience sample of 26 healthy subjects (8 males, 18 females) between the age of 22 and 61 years participated in the study. The characteristics of the sample are shown in Table 1. Each subject was observed by a research assistant while performing a controlled sequence of 50 minutes comprising the following activities: sitting, standing still, climbing stairs, walking, and cycling. In order to imitate free-living activities, all activities were performed at a self-paced speed. Walking and cycling were performed at two self-paced speeds: 'regular' and 'brisk'. Each subject walked indoors as well as outdoors. Cycling was outdoors on a standard bicycle. All activities were conducted between May and July 2009 in similar weather conditions (i.e., no rain, mild wind). The subjects wore various measurement instruments: a heart rate receiver unit (Polar Electro S610i, Finland) on the wrist with the transmitter (Polar T61 Coded Transmitter, Finland) worn on the chest, a uni-axial ActiGraph accelerometer (ActiGraph GT1M, Pensicola, FL) and a GPS (QSTARZ travel recorder V4.3, Taipei, Taiwan) on the right hip, and a uni-axial ActiGraph accelerometer on the right ankle. The ActiGraph GT1M is the most widely used uni-axial motion sensor. It measures changes in acceleration 30 times each second (30 Hz). The accumulated value is then stored in the memory of the device at the end of the epoch period. The ActiGraph has good reproducibility, validity, and feasibility when used to assess physical activity patterns or to estimate energy expenditure (De Vries et al., 2009; Welk et al., 2004). Accelerometer data (counts) were collected in one second epochs (1 Hz). The default sample-rate frequency setting of the GPS is also 1 Hz and the Polar heart rate monitor sampled at 0.2 Hz (once in every 5 seconds). Body height and body weight were measured with a portable stadiometer (Seca 225, Vogel & Halke GmbH & Co, Germany) and a digital scale (Soehnle 62882, Leifheit AG, Germany).

After following the stepwise procedure of The Central Committee on Research involving Human Subjects (Dutch CCMO) a review of the study protocol by an independent medical ethics review board was not deemed necessary. Written informed consent was obtained from all subjects.

Table 1 Sample characteristics (M $\pm$ SD)

	Males (n = 8)	Females (n = 18)	All (n = 26)
Age (yr)	40 $\pm$ 14	38 $\pm$ 11	39 $\pm$ 11
Height (m)	1.81 $\pm$ 0.05	1.70 $\pm$ 0.07	1.74 $\pm$ 0.08
Weight (kg)	85.2 $\pm$ 11.8	65.2 $\pm$ 7.2	71.4 $\pm$ 12.8
BMI (kg/m ²)	25.8 $\pm$ 2.9	22.5 $\pm$ 1.8	23.5 $\pm$ 2.6

Note: BMI: body mass index

2.2 Statistical analyses

When the data collection was complete, the accelerometer data was downloaded to a personal computer and processed using the ActiLifeGT1M 2.2.3 software program (ActiGraph GT1M, Pensicola, FL). GPS data was also downloaded to the computer and processed using Qstarz Travel Recorder PC Utility V4 software (Taiwan). Polar

Precision Performance software (Almere, The Netherlands) was used to read the Polar Electro S610i receiver.

First, descriptive statistics were used to characterize the sample and to describe the output variables. Next, correlations between accelerometer counts and heart rate data, and between accelerometer counts and velocity data were computed to study the relationship between these variables. The Spearman correlation coefficient was chosen because the count and velocity variables were non-normally distributed.

In order to classify the activity type, feed-forward ANN models with a single hidden layer, five hidden units, and weight decay equal to 0.01 (see Figure 2) were used. Figure 2 shows a feed-forward ANN with five hidden units. An ANN model is an advanced regression model in which the variable physical activity is the dependent variable. The independent variables are statistical summaries chosen to describe the signal characteristics. In order to select suitable signal features, signal features used by Rothney et al. (2007) and Staudenmayer et al. (2009) were studied. In total, sixteen signal characteristics were computed over ten seconds of accelerometer data, and the correlations between all signal features were analyzed. Finally, the following summaries were computed for each ten seconds of accelerometer data: 10th, 25th, 75th, and 90th percentiles, absolute deviation, coefficient of variability and lag-one autocorrelation. The dependent variable had the following $K=9$ categories: sitting, standing, going up the stairs, going down the stairs, walking indoors, regular walking outdoors, brisk walking outdoors, regular cycling, and brisk cycling. First, two ANN models were fitted for the hip and the ankle accelerometer data respectively. Next, new ANN models were developed by adding heart rate or velocity data to the input variables. The statistical summaries chosen to describe heart rate and velocity signals were means over ten seconds of data. The model fit was evaluated by leave-one-subject-out cross-validation (Venables, 2002).

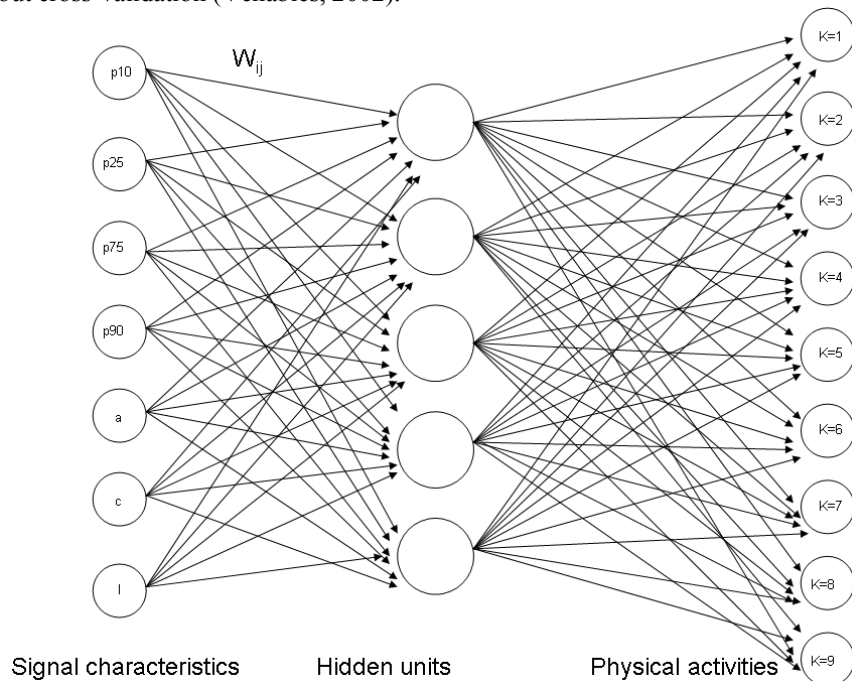


Figure 2 Feed-forward neural network model for $K=9$ activities

Note: The inputs, which represent the characteristics of the acceleration signal, are connected to the outputs, which represent physical activities by hidden units which are latent variables representing the neurons; p_{10} = 10th percentile; p_{25} = 25th percentile; p_{75} = 75th percentile; p_{90} = 90th percentile; a = absolute deviation; c = coefficient of variability; and l = lag-one autocorrelation.

All descriptive statistical analyses were performed using SPSS 17.0 (SPSS Inc. Chicago, IL). The classification models were developed with the function `nnet` (Venables, 2002) in the software package R version 2.8.0 (R Development Core Team, 2008). Both R and `nnet` are freely available.

3 Results

3.1 General results

Figure 3a-c shows the main output per measurement instrument. Figure 3b illustrates the differences in velocity between the activities. There were significant differences between the outdoors activities ($F(3,8047)=6838,29$, $p=0,000$). Figure 3c illustrates the differences in heart rate between the activities. Comparing to sitting, which was used as reference category; significant differences were found for all the activities ($F(8,13689)=1844,82$, $p=0,000$).

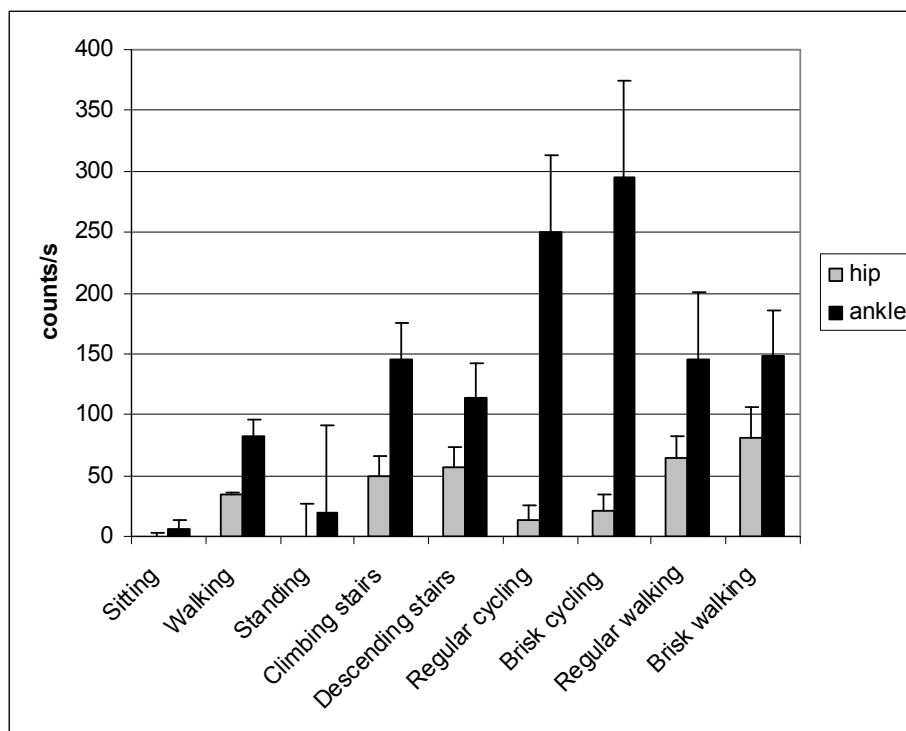


Figure 3a Accelerometer data: mean counts per second for the nine activities

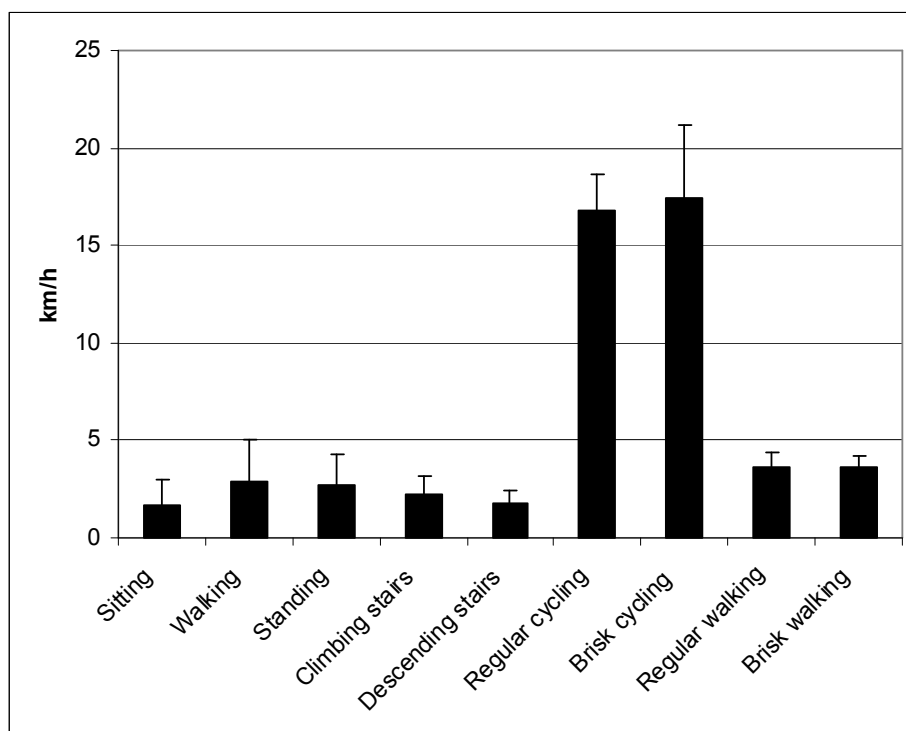


Figure 3b GPS data: mean kilometers per hour for the nine activities

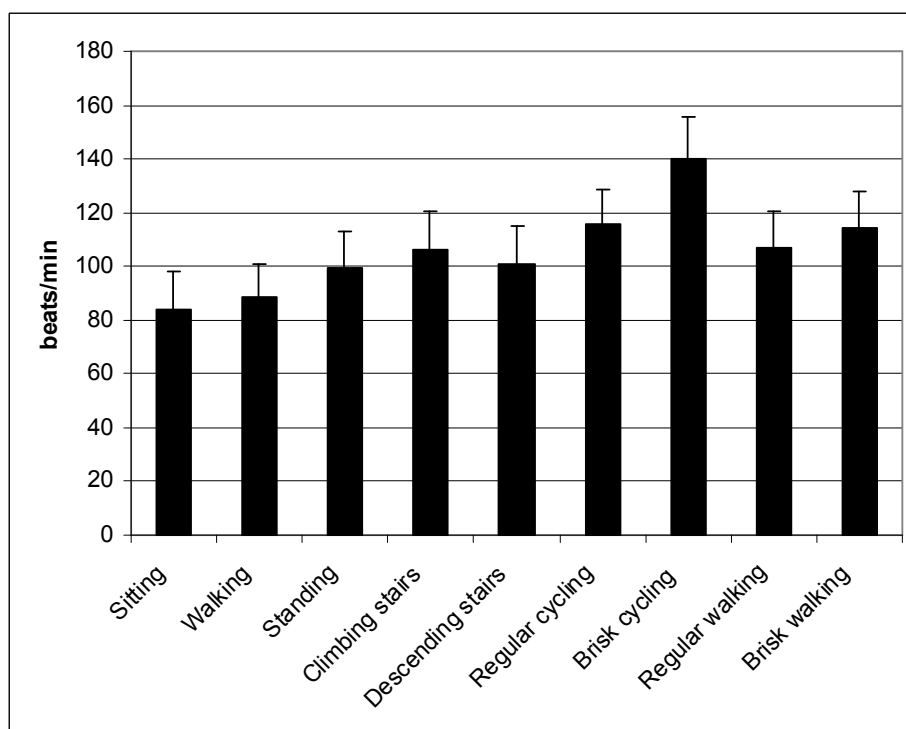


Figure 3c Heart rate data: mean beats per minute for the nine activities

Table 2 illustrates that there was a positive relationship between accelerometer counts (hip and ankle), heart rate and velocity. For velocity, activities performed indoors were excluded. The correlation between heart rate and velocity for the outdoors activities was 0,345. The correlations were in all cases significant ($p < 0.01$).

Table 2 Correlations between hip and ankle counts, velocity and heart rate data

	Hip counts	Ankle counts
Velocity (GPS)	-0.582**	0.655**
Heart rate	0.345**	0.522**

3.2 Activity classification

Table 3 reports the sensitivity of the cross-validated results for the ANN models with hip accelerometer data in terms of correctly classifying the nine activity types. Table 4 reports the sensitivity for the ANN model with ankle accelerometer data. In general, both models performed reasonably well in classifying the activities walking, cycling, and sitting. However, the models performed worse in correctly classifying climbing stairs and standing still. Overall, the ANN model based on hip accelerometer data correctly classified activity type 61.8% of the time, while the model based on ankle accelerometer data attained a percentage of 56.5%. The latter percentage is lower than the percentage of correctly classified activities in the previous study (De Vries et al., 2009; De Vries et al., submitted).

Table 3 shows that the ANN model based on hip accelerometer data was better able to classify the activities walking (indoors), brisk walking (outdoors) and standing still, while the model based on ankle accelerometer data (Table 4) was better able to classify brisk cycling. Both models performed well when classifying sitting ($> 90\%$ correctly classified). In the appendix (A) the classification errors of the models are presented in more detail in four contingency tables representing the relationship between the observed and the predicted physical activities. These tables were built with the cross-validated results for each model.

Table 3 Percentage of physical activities correctly classified by ANN models for hip accelerometer data

Model	Hip	Hip + GPS	Hip + HR	Hip + GPS + HR
Walking	87.0	92.0	90.4	93.6
Walking indoors	29.4	31.9	46.1	47.1
Regular walking	72.2	71.1	77.0	78.6
Brisk walking	63.6	67.1	58.8	56.6
Cycling	85.8	93.6	87.8	95.1
Regular cycling	76.1	73.1	73.4	81.2
Brisk cycling	23.4	59.5	54.8	65.6
Climbing stairs	5.5	10.5	25.6	18.8
Going up	4.0	52.1	29.4	16.6
Going down	10.6	37.5	13.6	9.8
Standing still	49.2	43.6	33.5	31.7
Sitting	92.6	73.3	84.6	89.4
Total	61.8	68.1	65.7	69.7

Table 4 Percentage of physical activities correctly classified by ANN models for ankle accelerometer data.

Model	Ankle	Ankle + GPS	Ankle + HR	Ankle + GPS + HR
Walking	74.6	83.1	77.3	84.9
Walking indoors	19.2	4.1	32.3	9.0
Regular walking	75.7	81.0	70.7	69.7
Brisk walking	0.0	0.0	9.4	20.7
Cycling	83.4	89.3	84.3	89.9
Regular cycling	71.8	67.4	71.7	69.4
Brisk cycling	55.0	36.7	64.8	59.54
Climbing stairs	9.1	6.2	20.2	17.0
Going up	5.8	1.1	18.2	6.1
Going down	0.0	0.0	1.0	0.5
Standing still	3.9	0.0	2.9	0
Sitting	93.3	64.2	83.6	61.9
Total	56.5	56.6	59.4	60.7

Both the hip model as well as the ankle model improved by adding velocity and/ or heart rate data to the models. After the inclusion of GPS and heart rate data the model based on hip accelerometer data correctly classified the type of activity 69.7% of the time, while the model based on ankle accelerometer data correctly classified the activities 60.7% of the time. Especially walking and cycling were better classified with the models that use GPS and/or heart rate data compared to the models based on accelerometer data alone. On the other hand, the combined models performed worse when classifying sitting and standing. This might be explained by the inaccuracy of the GPS measuring velocity within three meters radius. In addition, the activities standing still and sitting were often performed after more intense activities leading to a raised heart rate instead of a rest heart rate.

When looking at both models in more detail, Table 3 shows that the model based on hip accelerometer data could better classify the activities regular and brisk cycling after the inclusion of GPS and heart rate data. Another remarkable improvement was achieved for the activities going up and down the stairs. However, these activities were better classified with the models including data from one extra device (accelerometer and GPS or heart rate monitor) than with the complete model. Appendix A illustrates that after the inclusion of GPS data to the model based on hip accelerometer data there was no longer misclassification errors between cycling and climbing stairs. However, the misclassification error returned after the inclusion of heart rate data.

Table 4 shows that the percentage of correctly classified activities of the ankle model slightly improved from 56.5% to 59.4% with the inclusion of heart rate and to 60.7 with the inclusion of GPS and heart data. In general, including GPS data to the ankle model did not lead to an improvement. Although some activities could be better classified after the inclusion of data from one extra device to the model based on ankle accelerometer data, the improvements disappeared when both GPS and heart rate data were added to the model.

In general, it can be concluded that including GPS or heart rate data to an ANN model based on single sensor accelerometer data leads to a small improvement in the percentage of correctly classified activities with the model based on hip accelerometer data and to hardly any improvement with the model based on ankle accelerometer data.

4 Conclusion and recommendations

4.1 Discussion and conclusion

This study developed, compared and evaluated relatively simple artificial neural network (ANN) models for the purposes of classifying nine physical activity types based on accelerometer data from the hip or ankle, GPS data, and heart rate data of 26 healthy adults. Our results showed that all models performed reasonably well (> 75% correctly classified) when classifying walking and cycling (Poher et al., 2006; Staudenmayer et al., 2009). However, the models performed worse when classifying climbing stairs and standing still or discriminating between two self-paced speeds of walking and cycling. The inclusion of GPS data and/or heart rate data to the models did not result in a large improvement of the model based on hip accelerometer data, and the model did not improve using ankle accelerometer data alone.

In contrast to our previous study (De Vries et al., 2009; De Vries et al., submitted), the model based on hip accelerometer data performed better than the model based on accelerometer data from the ankle. The model based on hip accelerometer data correctly classified activity type 61.8% of the time, while the model based on ankle accelerometer data attained a percentage of 56.5%. Therefore, in line with international standards we recommend to wear the accelerometer on the hip in studies to assess physical activity levels. Although adding a GPS or heart rate monitor to the measurement instruments lead to a better classification of some of the activities (i.e., brisk cycling), we believe that at this moment this advantage is disproportional to the costs in terms of equipment, increased amount of time for the researcher for analyzing the data, and subject burden.

For future studies, the combination of GPS and an accelerometer on the hip seems most hopeful. This combination improved the classification of outdoor activities. However, the GPS did not perform very well indoors, in particular for the activity climbing stairs. A disadvantage of GPS is that they usually are not able to connect to any or enough satellites when they are worn inside buildings. This problem might be solved by adding an additional antenna and transmitters inside buildings or by developing a filter to distinguish false or unrealistic data from true data. Adding a heart rate monitor to a hip-mounted accelerometer was of no further significance in this study, most likely because of the delayed response of a person's heart after a change in activity or intensity.

It can be concluded that the developed ANN models are able to predict a number of physical activities of adults relatively well based on accelerometer data. Nevertheless, the number of times that an activity is classified incorrectly is still quite high. Moreover, the addition of GPS and/or heart rate data did not improve the percentage of correctly classified activities of the models. Based on these data, it is not recommended to use a combination of single sensors in large scale epidemiological studies on physical activity with the purpose to improve objective measurement and classification of physical activity.

4.2 Recommendations

Future studies should determine whether the accuracy of the ANN models can be improved by including alternative signal characteristics into the models based on wavelets transformation or signal characteristics that add information about the cyclic nature of activities. Data from an accelerometer with three axes instead of one axis may also improve the model. ActiGraph has recently launched the GT3X accelerometer. This version also contains an inclinometer. Future studies should test ANN models based on 3D accelerometer data and inclinometer data for classification of physical activity patterns. Furthermore, this study examined nine free-living activities. It would be interesting to assess whether other activities, such as household activities, gardening, and different sports can also be classified using ANN models. In addition, we recommend examining whether ANN models can improve the accuracy of accelerometer measurements in children and elderly. Finally, further research should be performed to investigate if alternative pattern-recognition-based methods, such as tree models or mixed models with wavelets perform better than the developed ANN model in correctly classifying physical activities.

5 References

ARMSTRONG N, WELSMAN JR. The physical activity patterns of European youth with reference to methods of assessment. *Sports Med* 2006; 36 (12): 1067-86.

AMERICAN COLLEGE OF SPORTS MEDICINE. ACSM's guidelines for exercise testing and prescription. 6th ed. Philadelphia: Lippincott Williams & Wilkins; 2000.

BONOMI AG, PLASQUI G, GORIS AH, WESTERTERP KR. Improving assessment of daily energy expenditure by identifying types of physical activity with a single accelerometer. *J Appl Physiol* 2009 September;107(3):655-61.

ESLIGER DW, COPELAND JL, BARNES JD, TREMBLAY MS. Standardizing and optimizing the use of accelerometer data for free-living physical activity monitoring. *J Phys Act Health* 2005; (3): 366-83.

LEVINE JA, BAUKOL PA, WESTERTERP KR. Validation of the Tracmor triaxial accelerometer system for walking. *Med Sci Sports Exerc* 2001 September; 33 (9): 1593-7.

POBER DM, STAUDENMAYER J, RAPHAEL C, FREEDSON PS. Development of novel techniques to classify physical activity mode using accelerometers. *Med Sci Sports Exerc* 2006; 38 (9): 1626-34.

RICHARDSON MT, AINSWORTH BE, JACOBS DR, LEON AS. Validation of the Stanford 7-day recall to assess habitual physical activity. *Ann Epidemiol* 2001 February; 11 (2): 145-53.

ROTHNEY MP, NEUMANN M, BEZIAT A, CHEN KY. An artificial neural network model of energy expenditure using nonintegrated acceleration signals. *J Appl Physiol* 2007 October; 103 (4): 1419-27.

ROWLANDS AV, THOMAS PW, ESTON RG, TOPPING R. Validation of the RT3 triaxial accelerometer for the assessment of physical activity. *Med Sci Sports Exerc* 2004 March; 36 (3): 518-24.

STAUDENMAYER J, POBER D, CROUTER SE, BASSETT DR, FREEDSON P. An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer. *J Appl Physiol* 2009 Oct; 107 (4): 1300-7.

TUDOR-LOCKE CE, MYERS AM. Challenges and opportunities for measuring physical activity in sedentary adults. *Sports Med* 2001 February; 31 (2): 91-100.

VENABLES WN, RIPLEY BD. *Modern Applied Statistics with S*. New York: Springer-Verlag; 2002: 7.

VRIES SI DE. Activity-friendly neighborhoods for children: measurement of physical activity and environmental correlates. Proefschrift Vrije Universiteit Amsterdam. Leiden: De Bink, 2009. ISBN 978-90-8659-392-7.

VRIES SI DE, ENGBERS LH, HILDEBRANDT VH. Objectief meten bewegen; Resultaten en conclusies fase I 'state of the art' en fase II 'patroonherkenning'. Leiden: TNO Kwaliteit van Leven. TNO-rapport KvL/GL 2009.050.

VRIES SI DE, GALINDO GARRE F, ENGBERS LH, HILDEBRANDT VH, BUUREN S VAN. Evaluation of neural networks to identify types of activity using accelerometers. Submitted for publication.

VRIES SI DE, HIRTUM WJEM VAN, BAKKER I, HOPMAN-ROCK M, HIRASING RA, MECHELEN W VAN. Validity and reproducibility of motion sensors in youth: a systematic update. Med Sci Sports Exerc 2009; 41 (4): 817-27.

WELK GJ. Introduction to physical activity research. In: Welk GJ, editor. Physical activity assessments for health-related research. Champaign IL: Human Kinetics; 2002. p. 3-18.

WELK GJ, SCHABEN JA, MORROW JR. Jr. Reliability of accelerometry-based activity monitors: a generalizability study. Med Sci Sports Exerc 2004 September; 36 (9): 1637-45.

A Contingency tables

Model 1. Cross-validation results for the classification of nine physical activities with hip accelerometer data

Test set	Observed activities								
Predicted activities	Walking indoors	Regular walking	Brisk walking	Regular cycling	Brisk cycling	Going up the stairs	Going down the stairs	Standing still	Sitting
Walking indoors	29.39	1.93	0.52	1.12	5.38	0.95	8.54	6.60	0.76
Regular walking	9.39	72.15	30.53	0.66	2.88	0.00	6.03	6.09	0.00
Brisk walking	0.91	21.06	63.65	0.07	0.00	0.00	5.53	9.64	0.00
Regular cycling	20.61	0.50	0.00	76.13	60.13	12.48	8.54	1.52	4.75
Brisk cycling	18.79	1.17	0.00	10.85	23.38	0.32	22.11	11.17	0.41
Going up the stairs	0.30	0.00	0.00	1.25	0.38	3.95	0.00	0.00	1.00
Going down the stairs	3.64	0.59	0.52	0.72	2.38	0.00	10.55	9.64	0.00
Standing still	7.58	2.43	4.40	0.26	1.00	0.00	34.17	49.24	0.00
Sitting	9.09	0.00	0.00	8.68	3.88	81.36	1.01	2.03	92.55

Model 2. Cross-validation results for the classification of nine physical activities with hip accelerometer data and GPS data

Test set	Observed activities								
Predicted activities	Walking indoors	Regular walking	Brisk walking	Regular cycling	Brisk cycling	Going up the stairs	Going down the stairs	Standing still	Sitting
Walking indoors	31.88	10.48	3.49	4.80	20.09	1.75	15.28	5.24	6.99
Regular walking	2.84	71.09	21.40	0.15	0.92	0.00	2.15	1.46	0.00
Brisk walking	0.15	29.74	67.06	0.00	0.00	0.00	1.31	1.75	0.00
Regular cycling	1.58	0.19	0.00	73.12	21.70	1.83	0.19	0.00	1.39
Brisk cycling	6.67	1.00	0.00	27.17	59.50	2.00	0.67	0.33	2.67
Going up the stairs	2.74	0.00	0.00	5.48	10.96	52.05	0.00	0.00	28.77
Going down the stairs	7.14	10.71	5.36	3.57	7.14	0.00	37.50	26.79	1.79
Standing still	10.48	16.13	13.71	1.61	2.42	0.00	12.10	43.55	0.00
Sitting	1.97	0.00	0.00	0.95	1.33	22.34	0.06	0.06	73.29

Model 3. Cross-validation results for the classification of nine physical activities with hip accelerometer data and heart rate data

Test set	Observed activities								
Predicted activities	Walking indoors	Regular walking	Brisk walking	Regular cycling	Brisk cycling	Going up the stairs	Going down the stairs	Standing still	Sitting
Walking indoors	46.06	1.59	1.30	2.04	1.13	0.32	17.59	14.21	0.35
Regular walking	10.00	77.01	36.92	0.85	1.75	0.00	13.57	12.69	0.00
Brisk walking	0.91	16.69	58.81	0.00	0.13	0.00	7.04	7.61	0.00
Regular cycling	25.45	0.84	0.00	73.44	36.38	14.38	17.09	7.11	3.87
Brisk cycling	1.52	0.84	0.00	12.62	54.75	1.26	10.05	6.09	1.47
Going up the stairs	1.52	0.00	0.00	5.92	0.50	29.38	0.00	1.02	9.15
Going down the stairs	1.82	0.76	0.39	0.85	0.75	0.00	13.57	13.20	0.00
Standing still	5.76	2.10	2.20	0.20	0.88	0.00	17.09	33.50	0.00
Sitting	6.67	0.00	0.00	3.81	3.13	53.71	0.50	0.51	84.63

Model 4. Cross-validation results for the classification of nine physical activities with hip accelerometer data, GPS data and heart rate data

Test set	Observed activities								
Predicted activities	Walking indoors	Regular walking	Brisk walking	Regular cycling	Brisk cycling	Going up the stairs	Going down the stairs	Standing still	Sitting
Walking indoors	47.14	1.68	0.91	1.10	2.88	0.68	20.33	9.76	0.73
Regular walking	13.22	78.61	40.70	0.29	1.00	0.00	28.46	26.83	0.00
Brisk walking	0.44	16.61	56.57	0.00	0.13	0.00	8.13	9.76	0.00
Regular cycling	12.78	0.08	0.00	81.24	27.78	8.16	2.44	0.00	1.05
Brisk cycling	4.41	0.84	0.00	14.86	65.58	3.85	13.01	8.13	1.94
Going up the stairs	2.20	0.00	0.00	0.88	0.25	16.55	0.00	0.00	6.21
Going down the stairs	1.32	0.34	0.00	0.07	1.00	0.00	9.76	7.32	0.00
Standing still	4.85	1.68	1.43	0.22	0.13	0.00	11.38	31.71	0.00
Sitting	13.22	0.00	0.00	1.03	0.63	69.39	0.81	0.00	89.35

Model 5. Cross-validation results for the classification of nine physical activities with ankle accelerometer data

Test set	Observed activities								
Predicted activities	Walking indoors	Regular walking	Brisk walking	Regular cycling	Brisk cycling	Going up the stairs	Going down the stairs	Standing still	Sitting
Walking indoors	19.19	3.24	3.43	2.66	2.43	2.88	13.66	19.51	1.58
Regular walking	17.44	75.65	80.76	5.77	0.85	3.64	23.90	37.07	0.56
Brisk walking	0.00	0.08	0.00	0.00	0.00	0.00	0.00	0.49	0.00
Regular cycling	6.69	7.20	3.80	71.78	30.26	3.03	25.37	2.93	0.73
Brisk cycling	17.15	5.42	2.33	10.65	55.04	2.73	9.27	4.88	1.75
Going up the stairs	3.49	0.32	0.25	0.32	1.09	5.77	1.95	3.90	1.41
Going down the stairs	0.00	0.24	0.00	0.06	0.00	0.15	0.00	0.00	0.00
Standing still	1.45	0.16	0.86	0.00	0.12	0.15	1.46	3.90	0.11
Sitting	34.30	7.52	8.21	8.50	9.60	80.73	20.98	23.41	93.34

Model 6. Cross-validation results for the classification of nine physical activities with ankle accelerometer data and GPS data

Test set	Observed activities								
Predicted activities	Walking indoors	Regular walking	Brisk walking	Regular cycling	Brisk cycling	Going up the stairs	Going down the stairs	Standing still	Sitting
Walking indoors	4.07	2.91	2.82	0.13	0.97	1.21	2.44	2.93	1.07
Regular walking	20.06	80.99	82.84	1.84	6.20	3.95	26.34	29.27	1.64
Brisk walking	0.00	0.00	0.00	0.13	0.00	0.00	0.00	0.00	0.00
Regular cycling	7.56	0.32	0.12	67.41	44.47	5.92	1.46	0.49	1.18
Brisk cycling	7.56	3.16	1.84	12.43	36.70	2.43	4.88	3.41	0.90
Going up the stairs	0.00	0.24	0.37	1.84	0.36	1.06	0.98	1.95	0.34
Going down the stairs	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Standing still	0.58	0.16	0.12	0.00	0.24	0.00	0.00	0.00	0.06
Sitting	25.87	8.50	5.88	2.16	7.53	51.44	20.49	18.05	64.24

Model 7. Cross-validation results for the classification of nine physical activities with ankle accelerometer data and heart rate data

Test set	Observed activities								
Predicted activities	Walking indoors	Regular walking	Brisk walking	Regular cycling	Brisk cycling	Going up the stairs	Going down the stairs	Standing still	Sitting
Walking indoors	32.27	1.54	1.96	1.33	1.34	4.40	8.29	12.20	2.03
Regular walking	16.86	70.71	69.00	8.12	1.34	2.58	24.88	38.05	0.85
Brisk walking	0.00	4.53	9.44	0.25	0.24	0.30	1.46	0.49	0.00
Regular cycling	8.72	7.04	6.99	71.66	21.75	7.44	32.68	8.29	1.64
Brisk cycling	2.03	5.26	1.72	7.36	64.76	3.64	3.41	6.83	1.35
Going up the stairs	8.14	1.70	0.49	2.28	2.79	18.21	8.29	7.32	6.15
Going down the stairs	0.58	0.00	0.00	0.00	0.00	0.30	0.98	0.00	0.06
Standing still	1.16	0.00	0.00	0.06	0.00	0.15	0.49	2.93	0.00
Sitting	25.87	5.50	4.66	5.14	4.37	58.12	13.17	16.10	83.59

Model 8. Cross-validation results for the classification of nine physical activities with ankle accelerometer data, GPS data and heart rate data

Test set	Observed activities								
Predicted activities	Walking indoors	Regular walking	Brisk walking	Regular cycling	Brisk cycling	Going up the stairs	Going down the stairs	Standing still	Sitting
Walking indoors	9.01	3.96	1.72	0.38	0.49	1.67	4.39	3.90	0.90
Regular walking	20.35	69.66	62.87	1.27	2.67	4.40	27.32	32.20	1.58
Brisk walking	1.16	11.81	20.71	0.44	0.73	0.46	0.98	0.49	0.17
Regular cycling	5.81	0.73	1.23	69.37	24.91	4.70	1.95	0.00	0.62
Brisk cycling	3.20	3.88	1.84	9.58	59.54	5.61	3.90	3.90	2.09
Going up the stairs	6.10	0.32	0.49	3.68	2.67	6.07	5.37	1.95	2.14
Going down the stairs	0.29	0.00	0.00	0.13	0.00	0.00	0.49	0.00	0.06
Standing still	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Sitting	19.77	5.91	5.02	1.08	5.47	43.10	12.20	13.66	61.87