This is a repository copy of *A comparison of the World Health Organisation's HEAT model results using a non-linear physical activity dose response function with results from the existing tool [under peer review]*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/183404/

Version: Submitted Version

Article:

RESEARCH ARTICLE

A comparison of the World Health Organisation's HEAT model results using a non-linear physical activity dose response function with results from the existing tool [version 1; peer review: 1 approved with reservations]

Robert Smith 🆔, Chloe Thomas, Hazel Squires 🆔, Elizabeth Goyder 🆔

School of Health and Related Research, University of Sheffield, Sheffield, South Yorkshire, S1 4DA, UK

Abstract

IntroductionThe WHO-Europe's Health Economic Assessment Tool is a tool used to estimatethe costs and benefits of changes in walking and cycling. Due to data limitationsthe tool's physical activity module assumes a linear dose response relationship be-tween physical activity and mortality.MethodsThis study estimates baseline population physical activity distributions for 44 coun-tries included in the HEAT. It then compares, for three different scenarios, the re-sults generated by the current method, using a linear dose-response relationship,with results generated using a non-linear dose-response relationship.ResultsThe study finds that estimated deaths averted are relatively higher (lower) using thenon-linear effect in countries with less (more) active populations. This difference islargest for interventions which affect the activity levels of the least active the most.Since more active populations, e.g. in Eastern Europe, also tend to have lowerValue of a Statistical Life estimates the net monetary benefit estimated by the sce-narios are much higher in western-Europe than eastern-Europe.ConclusionsUsing a non-linear dose response function results in materially different estimateswhere populations are particularly inactive or particularly active. Estimating base-line distributions is possible with limited additional data requirements, although themethod has yet to be validated. Given the significant role of the physical activitymodule within the HEAT tool it is likely that in the evaluation of many interventionsthe monetary benefit estimates will be sensitive to the choice of the physical activitydose response function.

Keywords

Physical Activity, Dose Response Function, HEAT, Walking, Cycling

Corresponding author: Robert Smith (rasmith3@sheffield.ac.uk)

Introduction

There is a growing recognition of the importance of considering health in all policies[1–3]. One example of successful integration of health impact in another policy domain is the World Health Organization's Health Economic Assessment Tool (HEAT), which has been widely used, primarily by transport planners, to estimate the health benefits associated with increased walking and cycling[4]. The success of the HEAT is in part due to its simplicity, requiring relatively few user inputs compared to other health economic models[5].

However, a limitation of the HEAT is that despite broad consensus that the relationship between physical activity and all-cause mortality is non-linear, such that the greatest health benefits from an extra unit of physical activity accrue in those who are least active[6–8], the HEAT assumes a linear relationship between physical activity and mortality. The HEAT methods and user guide states that "a linear relationship was chosen to avoid additional data requirements on baseline activity levels (which would be needed using a non-linear dose–response function)"[4]. There is however a recognition that improvements in data availability could allow for a non-linear relationship to be used in the future. The same report states that "An approach based on a non-linear relationship could be adopted as part of future updates of HEAT, when suitable data on the baseline level of physical activity in different populations are available to provide default values for HEAT" (p.9).

This study uses a method developed by Hafner *et al.*[9] to estimate the distribution of physical activity in 44 countries in the WHO European Region for which the HEAT applies. It then compares, for three hypothetical scenarios, the number of deaths averted and the monetary benefit when assuming a linear relationship, as done by the current HEAT model, and a non-linear relationship between physical activity and all-cause mortality. Although previous analysis has shown the importance of estimating changes in the distribution of physical activity, rather than categorizing activity levels[10], this is the first time that the effect of changing the shape of the dose response relationship has been analyzed by replicating the outputs of a widely used model.

Material and methods

Data and measures

This study uses data on the prevalence of insufficient physical activity in 44 HEAT countries from a publication by [11], the self-reported physical activity levels of a representative sample of the English population from the Health Survey for England 2015[12], country specific mortality rates for those aged 20–74 from the European Mortality Database[13] and value of a statistical life estimates from a systematic review[14]. It uses the linear dose-response relationship between physical activity and mortality from [7] as described in the HEAT methodology paper[4], and a non-linear dose-response relationship as described in [15]. A summary of data including sources can be found in Table 1.

Analysis

We estimate the number of deaths averted per 100,000 and the net monetary benefit using both the non-linear dose-response method and the linear dose-response currently used by HEAT for 44 European countries in three scenarios:

1. Scenario 1: An extra 10 minutes of daily walking for every person in the population.

2. Scenario 2: Every adult meets WHO Guidelines. Every adult in the country who doesn't already meet WHO guidelines of 600 MET-mins per week (equivalent to around 150 minutes of brisk walking per week) increase their activity to that level. Those meeting guidelines are unchanged.

3. Scenario 3: A 10% increase in physical activity levels of the population aged 20–74, such that those who are the most active have the largest absolute activity increase, and those who are least active have the smallest absolute activity increase.

Table 1. Variable names, description and source of data used in analysis.

Variable	Description	Source
From the HEAT methodology		
MR	Country specific mortality rates (for ages 20–74)	European Mortality Database (2017)[13]
RR_lit	Relative risk in literature	7; 4
mins_ref	Reference physical activity duration	7; 4
RR_min	Minimum relative risk (max effect)	7; 4
VSL	Value of a Statistical Life for each country	14
Other sources		
piap	% of population inactive	11 Appendix 5
mets	Distribution of met mins in English population	HSE 2015[12]
t	Log-linear dose response function power (t)	15

This analysis is not an attempt to estimate the probability, feasibility or costs of achieving the scenarios. For each scenario we assume that the outcome is achieved, and we estimate the benefits in terms of deaths averted per 100,000 and monetize these benefits using the VSL.

The current HEAT method using a linear dose response relationship

The current HEAT method requires the user to input pre-intervention and post-intervention physical activity levels, in terms of minutes of walking and cycling[4]. It estimates the relative risk associated with each activity level using Equation 1 below.

$$R = max\{1 - (1 - RR_{lit}) \times \frac{Mins_{local}}{Mins_{ref}}, RR_{min}\} \qquad (1)$$

For a walking intervention the relative risk RR_{lit} is 0.89, the reference minutes of activity from the literature $Mins_{ref}$ is 168mins per week and the risk reduction cap RR_{min} is 0.7, such that every additional 10 minutes of weekly walking ($Mins_{local}$ = 10) reduces relative risk by 0.65 percentage points, to a limit of 30 percentage points. Number of deaths averted DA is then calculated by multiplying the absolute difference in relative risk between intervention and comparator (RR_i–RR_c) by the country specific mortality rate of the population aged 20–74 MR_c and the population affected, pop (Equation 2 below). This is then monetized in terms of monetary benefit (MB) in Equation 3 by multiplying the number of deaths averted by the country specific value of a statistical life VSL_c:

$$DA = (RR_i - RR_c) \times MR_c \times pop \qquad (2)$$

$$MB = DA \times VSL_c \qquad (3)$$

The adapted method using a non-linear dose response relationship

The non-linear dose response method requires a baseline distribution of physical activity. We use weekly metabolic equivalent of task minutes (MET-minutes) from moderate and vigorous physical activity to summarize an individual's physical activity level in one number[16]. A distribution of weekly MET-mins for each country was imputed using a method from Hafner *et al.* (2019)[9]. This method combines estimates of prevalence of physical inactivity for each of the 44 countries with the distribution of physical activity in a generic distribution (we use the distribution derived from the Health Survey for England). Each percentile, *n*, of physical activity in the target country, *c*, distribution is calculated separately using the equation below.

$$p_c^n = p_c^{n-1} + (p_g^n - p_g^{n-1}) \frac{x_c}{x_g} \qquad (4)$$

The weekly MET-mins, *p*, for each country *c*, at each percentile, n, is based on the prevalence of sufficient physical activity, x_c, in the country, *c*, compared to the prevalence of sufficient activity in the generic distribution, x_g. The values for each percentile then form the estimated physical activity distribution for each country. More detail on this method, as well

as comparisons of country distributions, can be found in Hafner *et al.* (Appendix C)[9]. The estimates derived from these equations are available in the supplementary material.

The population relative risk is calculated as the simple arithmetic mean of relative risk for each percentile of the physical activity distribution, as shown in Equation 5 below. For each percentile relative risk is estimated using a log-linear relationship, calculated using the relative risk from the literature (RR_{lit} = 0.89), percentile MET-mins ($mets_p$), reference MET-mins ($mets_{ref}$) which is simply 4 (METs associated with moderate physical activity) × $mins_{ref}$ (from Kelly *et al.*, 2014[7]), and a power transformation *t*. The power transformation is 0.375 in the main analysis, and varied from 0.25 to 0.75 in sensitivity analysis, as recommended in Woodcock *et al.*[15].

$$RR = \frac{\sum_{(p=1)}^{(N=100)} RR_{lit}^{(\frac{mets_p}{mets_{ref}})^t}}{100} \qquad (5)$$

Figure 3 shows the dose response relationship between physical activity and all cause mortality risk for the linear model and the non-linear models with different values of t is shown in the supplementary material.

Once relative risk is calculated, the deaths averted and monetary benefit are calculated using Equation 2 and Equation 3.

Comparison

In each of the three scenarios, the number of deaths averted per 100,000 persons aged 20–74 and monetary benefit was estimated for 44 HEAT countries using the linear (HEAT) and non-linear dose response functions. The results for the number of deaths averted are displayed using simple scatter plots with a 45-degree line of equality and monetary benefit estimates are shown on choropleth maps of Europe.

Results
The estimated distributions of physical activity for each of the 44 countries in the analysis are provided in the supplementary material which can be found on a GitHub repository: https://github.com/RobertASmith/HEAT_DRF.

A comparison of the number of premature annual deaths averted per 100,000 people using the two different methods in each of the three scenarios for the 44 WHO European Region countries is shown in Figure 1 below. The estimates derived using the linear dose response method are shown on the x-axis and the non-linear dose response on the y axis. A 45-degree line of equality is plotted to aid comparison. The country points are labelled with ISO3 codes and shaded from black for low insufficient physical activity prevalence (IPAP) to blue for those with a high IPAP.

The figure shows that for the first scenario, an additional 10 minutes of daily walking, countries with particularly inactive (active) populations tend to have higher (lower) estimated deaths averted using the non-linear function compared to the linear function.
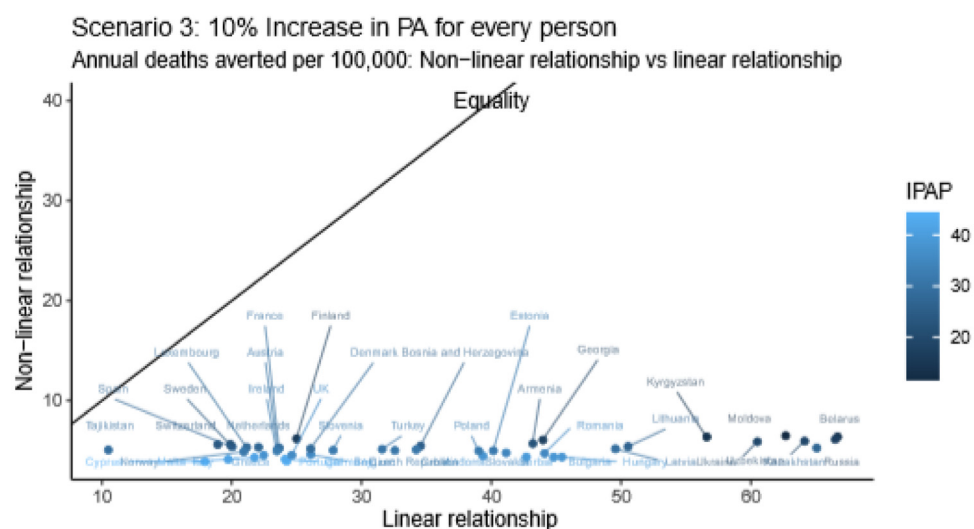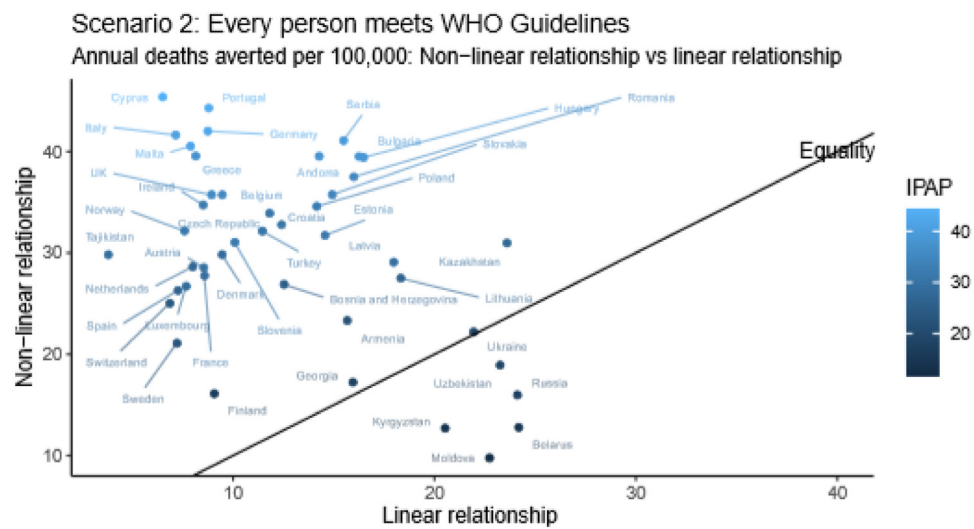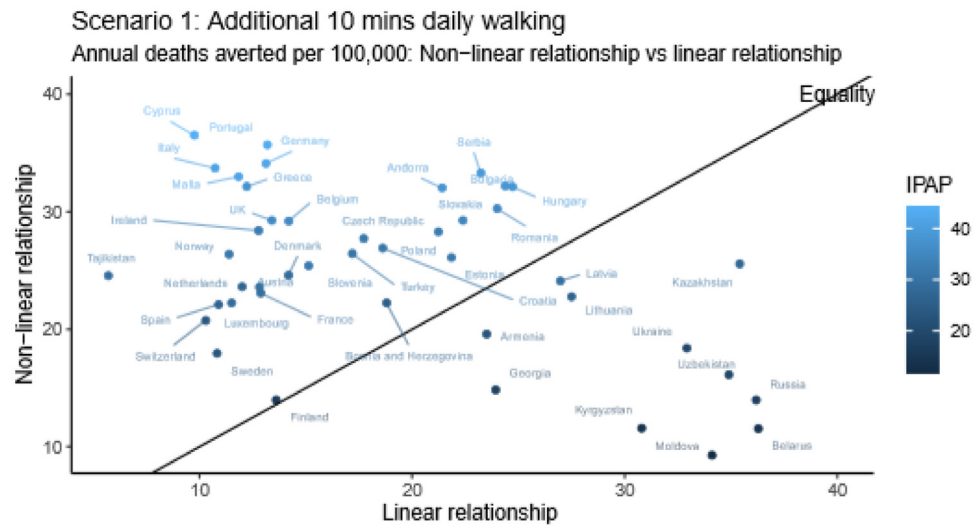
Figure 1. Deaths averted per 100,000 for three scenarios using the non-linear and the current (linear) relationship.

In the second scenario all individuals with activity levels below WHO physical activity guidelines of 600 MET-mins per week increase activity to meet guidelines. Here, the non-linear function results in higher deaths averted than the linear function in most countries, except for some with especially low prevalence of insufficient physical activity (e.g. Moldova and Belarus).

In the third scenario, in which all individuals increase their physical activity level by 10%, estimates derived using a non-linear function are much lower than using a linear function for all countries, regardless of the prevalence of insufficient physical activity. This is because those with low physical activity levels, who would benefit the most from increased physical activity according to a non-linear model, have low increases in MET-mins, while those who are highly active have high absolute increases in MET-mins but benefit little in terms of premature mortality reduction when using a non-linear model.

In order to allow for trade-offs in decision making between health and non-health outcomes, the HEAT tool monetises the deaths averted using the Value of a Statistical Life (VSL)[17], giving an estimate in terms of monetary benefit. Figure 2 below shows the monetary benefit associated with Scenario 1, using a log-linear dose response function with a power transformation

of 0.375. The monetary benefits tend to be higher in countries with higher insufficient physical activity prevalence and higher VSL (e.g. Ireland, the UK and Luxemburg) and markedly lower in countries with lower VSL and/or lower physical inactivity prevalence such as Ukraine and Moldova, this results in marked differences between the West and East Europe.

Discussion
Increasing population physical activity is likely to yield large benefits in health, wellbeing & productivity worldwide[9]. However, trade-offs often exist between increasing population physical activity and achieving other health and non-health outcomes. It is therefore important to have a robust method to consider whether interventions that improve activity levels provide good value for money. The HEAT is an example of a tool, often used by transport planners, which allows users to estimate, and monetize, the benefits of increased walking and cycling[3]. In general, the estimates derived from the physical activity module of the tool have been shown to contribute the most to total monetary benefit (Mueller et al., 2015).

We describe an adaption to the current HEAT physical activity module which applies a non-linear dose response relationship between physical activity and mortality risk to estimated country specific baseline distributions of physical activity. The method
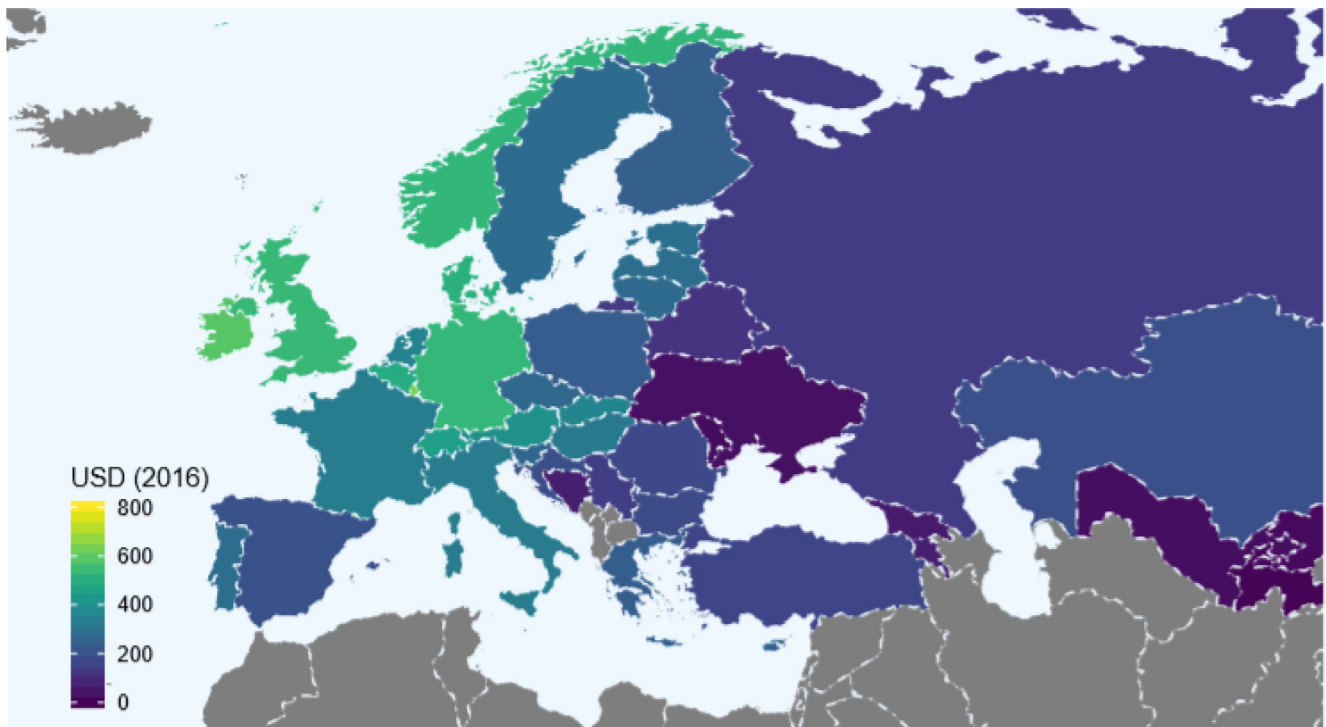
Figure 2. Annual Monetary Benefit of an additional 10 minutes daily walking for 44 European Countries, in 2016 USD.

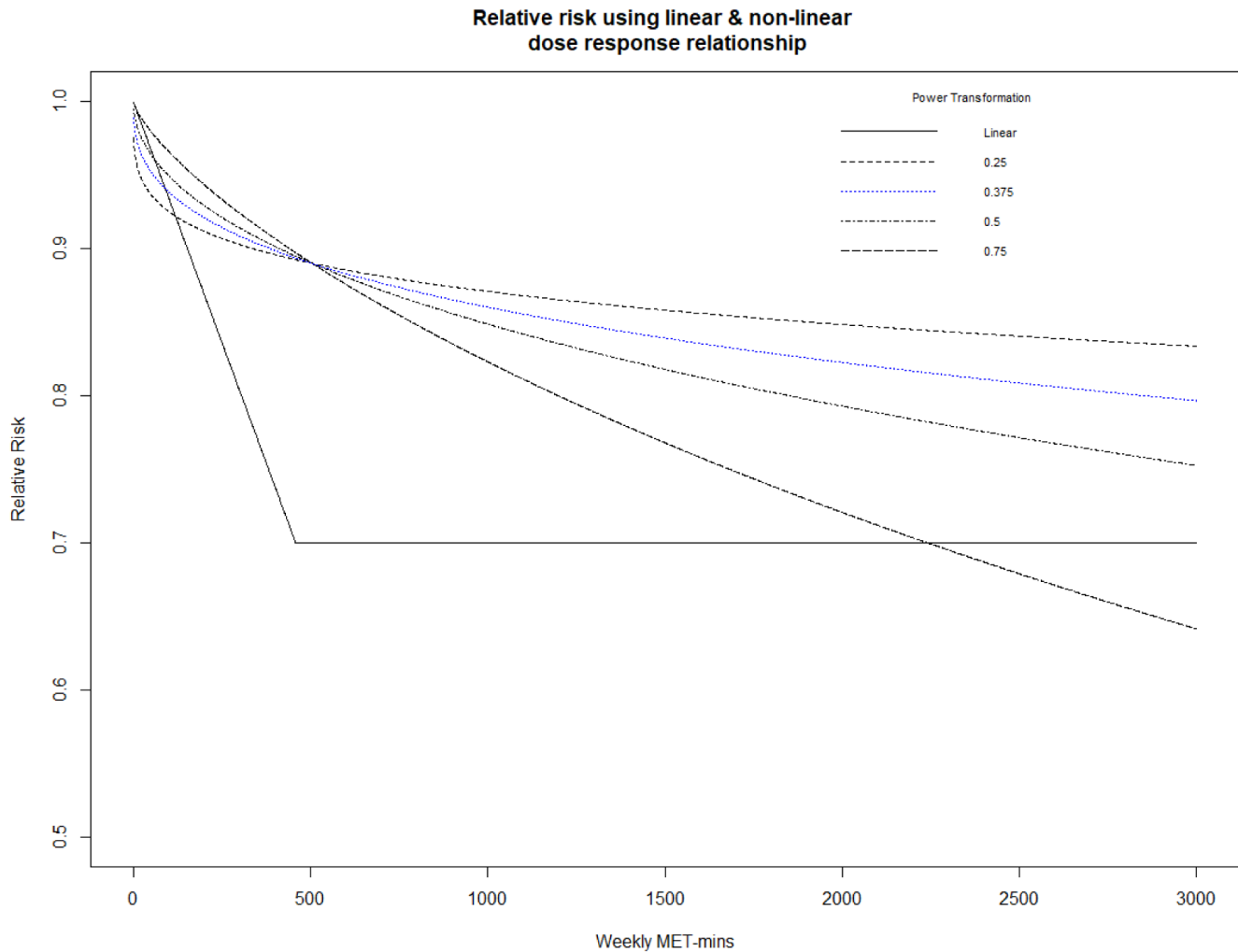## Relative risk using linear & non-linear
dose response relationship

Figure 3. Relative risk using linear & non-linear dose response functions with different power transformations:

is more sensitive to interventions which increase the activity levels of the least active, and less sensitive to interventions which increase the activity levels of the most active. This means that similar scenarios may yield less health benefit in more active countries. Since countries with higher GDP tend to have a higher Value of a Statistical Life[14] and higher prevalence of insufficient physical activity[11], the estimated net monetary benefit tends to be higher in western Europe than eastern Europe.

There are numerous limitations of this analysis. Firstly, the method used to estimate the baseline distributions of physical activity in each of the HEAT countries (from 9) assumes that the shape of the physical activity distribution is relatively similar in every country. Comparing the distributions estimated by this method, and provided in the supplementary material, with more detailed datasets would help to validate the estimates of population physical activity distributions. It is likely that the method is reliable for similar countries (e.g. the UK and Germany) but may not be reliable where culture differs

(e.g. the UK and Chad). However, it is unlikely that this would affect the main finding of this study, since large differences in the linear and non-linear functions exist when using the UK distribution which is based upon survey data.

A further limitation of this study is that we do not consider the usability of the tool, only show that a more conceptually valid method is possible. Since the tool is designed to be used by users with little to no public health, epidemiology, statistics and programming ability it is also important that the methods behind the tool are easy to explain, and the tool is simple to use. Increased complexity, in terms of more, or more detailed inputs, and a more difficult to explain model structure may make the tool less 'use-able', used here as a loose term which encompasses both technical feasibility of use and user understanding & confidence, and therefore less valuable. Further work to determine whether stakeholders understand the use of a non-linear dose response relationship on baseline and intervention distributions, and whether users can obtain intervention group physical activity distributions, will likely be a

determining factor as to the feasibility of adapting the HEAT tool. Nevertheless, this paper demonstrates that the two approaches do result in substantial differences at the population level, and therefore where possible the non-linear dose response function should be used by researchers.

The trade-off between the 'usability' and 'accuracy' of health impact assessment tools (and public health economic models more generally) is one that needs further attention in the academic literature. Models and tools tend to be either high accuracy but low usability - for example models created in high level programming languages with high computational demands and long runtimes - or low accuracy but high usability - including the HEAT physical activity modules. Understanding how to utilize new tools from data-science to make models which are very accurate and usable would be a useful avenue of future research. Likewise, understanding how to incrementally improve the accuracy of highly usable models (like HEAT) without compromising usability would be a valuable endeavor.

Conclusions

We show that for the WHO European Region countries included in the HEAT tool, the estimates of deaths averted, and therefore monetary benefit, differs substantially depending on the dose response function used. The nonlinear dose response function results in greater estimated benefits, relative to the linear

dose response function, where increased physical activity accrues to those who are relatively inactive. It therefore results in greater benefits in countries with higher prevalence of physical inactivity, or interventions which are targeted toward the least active. Developing tools which are both usable, in terms of data requirements and ease of explanation to users, and highly accurate is an important avenue for future research in health impact assessment and public health economics more widely.

References

1. World Health Organization *et al.*: **Health in all policies: Helsinki statement. framework for country action.** 2014.
   Reference Source

2. Koivusalo M: **The state of health in all policies (hiap) in the european union: potential and pitfalls.** *J Epidemiol Community Health.* 2010; **64**(6): 500–503.
   PubMed Abstract | **Publisher Full Text**

3. Kahlmeier S, Racioppi F, Cavill N, *et al.*: **"Health in all policies" in practice: guidance and tools to quantifying the health effects of cycling and walking.** *J Phys Act Health.* 2010; **7 Suppl 1**(s1): S120–S125.
   PubMed Abstract | **Publisher Full Text**

4. Kahlmeier S, Götschi T, Cavill N, *et al.*: **Health economic assessment tool (heat) for walking and for cycling. methods and user guide on physical activity, air pollution, injuries and carbon impact assessments.** 2017.
   Reference Source

5. Smith R, Thomas C, Squires H, *et al.*: **The price of precision: trade-offs between usability and validity in the world health organization health economic assessment tool for walking and cycling.** *Public Health.* 2021; **194**: 263–269.
   PubMed Abstract | **Publisher Full Text**

6. Arem H, Moore SC, Patel A, *et al.*: **Leisure time physical activity and mortality: a detailed pooled analysis of the dose-response relationship.** *JAMA Intern Med.* 2015; **175**(6): 959–967.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

7. Kelly P, Kahlmeier S, Götschi T, *et al.*: **Systematic review and meta-analysis of reduction in all-cause mortality from walking and cycling and shape of dose response relationship.** *Int J Behav Nutr Phys Act.* 2014; **11**(1): 132.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

8. Ekelund U, Tarp J, Steene-Johannessen J, *et al.*: **Dose-response associations between accelerometry measured physical activity and sedentary time and all cause mortality: systematic review and harmonised meta-analysis.** *BMJ.* 2019; **366**: l4570.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

9. Hafner M, Yerushalmi E, Phillips WD, *et al.*: **The economic benefits of a more physically active population: an international analysis.** 2019.
   Publisher Full Text

10. Minton J, Dimairo M, Everson-Hock E, *et al.*: **Exploring the relationship between baseline physical activity levels and mortality reduction associated with increases in physical activity: a modelling study.** *BMJ Open.* 2013; **3**(10): e003509.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

11. Guthold R, Stevens GA, Riley LM, *et al.*: **Worldwide trends in insufficient physical activity from 2001 to 2016: a pooled analysis of 358 population-based surveys with 1·9 million participants.** *Lancet Glob Health.* 2018; **6**(10): e1077–e1086.
    PubMed Abstract | **Publisher Full Text**

12. NHS Digital: **Health Survey for England.** London, UK. 2016.
    Reference Source

13. MDB (European detailed mortality database): **Copenhagen: WHO Regional Office for Europe.** 2017; accessed 18 October 2017.

14. **Mortality Risk Valuation in Environment OECD: Health and Transport Policies.** OECD Publishing, Paris. 2012.
    Publisher Full Text

15. Woodcock J, Franco OH, Orsini N, *et al.*: **Non-vigorous physical activity and all-cause mortality: systematic review and meta-analysis of cohort studies.** *Int J Epidemiol.* 2011; **40**(1): 121–138.
    PubMed Abstract | **Publisher Full Text**

16. Ainsworth BE, Haskell WL, Whitt MC, *et al.*: **Compendium of physical activities: an update of activity codes and met intensities.** *Med Sci Sports Exerc.* 2000; **32**(9 Suppl): S498–S504.
    PubMed Abstract | **Publisher Full Text**

17. Viscusi WK, Aldy JE: **The value of a statistical life: a critical review of market estimates throughout the world.** *J Risk Uncertainty.* 2003; **27**(1): 5–76.
    Publisher Full Text

Open Peer Review

Current Peer Review Status: ❓

Version 1

Reviewer Report 25 January 2022

https://doi.org/10.21956/wellcomeopenres.19250.r47911

❓ **James Woodcock** iD

MRC Epidemiology Unit, University of Cambridge, Cambridge, UK

This is a good paper and addresses an important issue.

While there has been insufficient attention to this topic I am not sure that it is quite right to say "*this is the first time that the effect of changing the shape of the dose response relationship has been analyzed by replicating the outputs of a widely used model*", but I would refer them to my article in 2013 that included comparisons of the dose response relationship using HEAT assumptions against ITHIM (which uses non-linear relations).[1]

I said "*To summarise the comparison, if using both tools with the recommended values (and combining results from changes to cycling and walking in HEAT) there were higher numbers of premature deaths averted with HEAT than with ITHIM if summing individual diseases in ITHIM but typically a smaller number of deaths averted with HEAT than with ITHIM if directly modelling all-cause mortality with ITHIM. However, due to the exclusion of impacts on older age groups with HEAT the premature deaths averted would be at an older average age in ITHIM compared with HEAT, and hence would tend to correspond to fewer years of life lost.*"

The authors compare a linear versus a power-transformed exposure. The estimation of the power transformation is taken from my paper (Woodcock *et al.* 2011). However, meta-analytic methods have moved on and I would now recommend cubic spline based approaches where avaiable e.g. Smith *et al.* 2016,[2] n.b. this paper is only for diabetes) The advantage of this approach is it allows a more flexible shaped curve. The transformed curves produce implausibly large benefits at lowest exposure levels (below 1 MET h/wk) and correspondingly produce smaller benefits beyond that. This means the transformed approach is overly sensitive to how many people are assumed to be doing zero activity. Given that PA survey design can lead to notably different results, including considerable variation in the proportion of the population who are active, care is needed.

I support the authors approach to modelling PA distributions. However, careful consideration is needed to harmonise responses to different questionnaires. Such harmonisation should also consider the questionnaires used as part of generating the dose response meta-analyses.

Walking and cycling vs broader measures of PA:
- When generating the PA distributions if I understand correctly these are not just walking or

cycling, but a broader measure (probably non-occupational PA)?

○ The RRs from HEAT and from Woodcock *et al.* 2011 are for walking alone. These should not be used in combination with prevalence estimates based on broader measures of PA.

○ Moving to a broader measure of PA has advantages as there are more studies to inform the DRF and thus it can be estimated for a wider range of diseases and not just all-cause mortality (which has problems about transferability).

○ I would like to see more information in the main text on the PA distributions, and the data used to parametrise them. While having the supplementary tables on Github has advantages, it is not clear where to look for what.

Usability:
○ While I support the idea of modelling PA distributions the user of a tool like HEAT would then need to know who's PA is being changed, the active or the inactive people, alternatively the tool would need to be able to estimate this in some way based on the type of intervention. In ITHIM we use stochastic matching based on demographics and other variables between PA and travel survey distributions, but additional approaches would be needed depending on the use case.

Figure 3:
○ In Figure 3 the linear dose response function looks wrong, 168 minutes per week of walking should equate to a RR of 0.89 but in the figure it seems to equate to RR 0.7

References
1. Woodcock J, Givoni M, Morgan AS: Health impact modelling of active travel visions for England and Wales using an Integrated Transport and Health Impact Modelling Tool (ITHIM).*PLoS One*. 2013; **8** (1): e51462 PubMed Abstract | Publisher Full Text
2. Smith AD, Crippa A, Woodcock J, Brage S: Physical activity and incident type 2 diabetes mellitus: a systematic review and dose-response meta-analysis of prospective cohort studies.*Diabetologia*. **59** (12): 2527-2545 PubMed Abstract | Publisher Full Text

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Yes

Are all the source data underlying the results available to ensure full reproducibility?

Partly

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: I sit on the core expert group of the WHO HEAT, and I developed the alternative ITHIM model.

Reviewer Expertise: Transport and health modelling

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.