APPENDIX 1: The National Propensity to Cycle Tool for England (PCT-England), version 1: Description of methodology for scenario-building

A1.1 Datasets for model building and model parameterisation

To estimate cycling potential, the Propensity to Cycle Tool (PCT) was designed to use the best available geographically disaggregated data sources on travel patterns. Currently for England and Wales this is the 2011 Census data on main mode of travel to work. For this reason the commuting layer was the first layer added to PCT-England. The 2011 Census was conducted in England and Wales on 27th March 2011 and covered an estimated 94% of the population [1]. All individuals aged 16 or over with a current job were asked "How do you usually travel to work? (Tick one box only, for the longest part, by distance, of your usual journey to work)". The commuting layer of PCT-England is based on the 22,676,958 commuters living in England, with adults who reported that their home address was also their place of work being treated as non-commuters. We hope in due course to add a commuting layer of PCT-Wales based on the 1,226,591 commuters living in Wales.

The core input dataset contained origin-destination (OD) pairs that linked each commuter's usual place of residence to the workplace location of their main job, and disaggregated these OD pairs by commute mode (N= 2,339,535 OD pairs for commuters living in England, 92,206 OD pairs for commuters living in Wales, available as an open-access dataset from https://wicid.ukdataservice.ac.uk/). Usual place of residence was identified at the level of the middle layer super output area (MSOA); MSOAs are administrative regions designed to contain a population of around 7500 individuals (average 3300 commuters). Workplace location was likewise identified at the MSOA level for those with a fixed workplace within England or Wales (90.5% commuters), and OD pairs were also present to capture commuters with no fixed workplace (9.1% commuters), working outside England or Wales (0.3% commuters) or working on an offshore installations (0.2% commuters). These OD pairs are directional, with one OD pair for travel from origin A to destination B, and another for travel from origin B to destination A. We enhanced this OD dataset by merging in other Census and route characteristic data, including the number of male and female commuters and the number of male and female cyclists in each OD pair (see Section A1.3.4); the distance and gradient of the 'fastest' routes estimated by CycleStreets.net (see Section A1.2); and the background mortality rate for existing and new cyclists under different scenarios (further details below in Section A1.5).

In addition to these input datasets, some of our analysis decisions and model parameterisation drew on analyses of the National Travel Surveys (NTS) in England and Wales (2008-2014, although data for Wales only collected up to 2012, accessed from http://discover.ukdataservice.ac.uk/), the Netherlands (2010-2014, accessed from https://easy.dans.knaw.nl/ui/home) and Switzerland (2010, obtained from the Swiss Federal Statistical Office, Neuchâtel [2], with data processing by Thomas Götschi). All three are nationally-representative surveys that include a travel diary, of duration 1 week in England and 1 day in the Netherlands and Switzerland.

A1.2 Estimating route distance and hilliness gradient across OD pairs

To model propensity to cycle we assigned distance and hilliness values to OD pairs for four different types of OD pair, as summarised in Table 1. The most important of these categories were between-MSOA flows of <30km, which account for 70% of all commuters and almost 80% of commuter cyclists. For each of these OD pairs we estimated the fastest route cycling distance between the population-weighted centroids of the origin and destination MSOA. We selected a 30km upper limit because above this very few commute trips are cycled in the English and Welsh or Dutch NTS, even among ebike owners (e.g. the proportion of commute trips 30-40km cycled was 0.3% among English or Welsh adults, and 2.9% among Dutch ebike owners). Fastest route distance was assigned using a routing algorithm 'developed for cyclists by cyclists' by the not-for-profit organisation CycleStreets (www.CycleStreets.net). Their main product is the journey planner which estimates 'fastest', 'quietest' and 'balanced' routes along roads, cycle paths and other travel network features (see http://www.cyclestreets.net/help/overview/#journeyplanner for more information) using data from OpenStreetMap (http://www.openstreetmap.org). For each route, we also extracted the total change in gradient experienced along the course of the route, as estimated by CycleStreets using data from the Ordnance Survey's 'OS Terrain 50' open access dataset[1]. A gradient of 2% indicates that for every 100m travelled horizontally the route involves a total change in vertical distance of 2m. This change of 2m could potentially reflect a rise of 2m or a fall of 2m or, for example, a rise of 1m followed by a fall of 1m.

For commuters living and working within the same MSOA, we estimated the average commute distance as one third the average of the three shortest between-MSOA OD pairs of the origin MSOA. MSOA gradient was estimated as the average of the gradient of these three shortest OD pairs. These decisions generated associations between distance, hilliness and cycling prevalence that closely matched the associations observed in between-MSOA OD pairs.

For commuters with 'no fixed workplace' or working overseas we did not have access to data that would allow us to assign values for the average distance that commuters would have to travel to cycle to work, nor could we assign the average hilliness of their routes. Likewise for OD pairs longer than 30km, we did not think it meaningful to calculate the apparent fastest-route cycling distance between the origin and destination as we suspected that in such pairs a large majority of cycle commuters were in reality travelling a much shorter distance (e.g. travelling to their workplace from a second home in an unknown location). Yet although we could not estimate route-allocated cycling distances for commuters in these OD pairs, we did estimate the average distance travelled by cyclists, as described in the final column of Table 1. This allowed us to include the existing cyclists in these OD pairs when estimating the health and carbon impacts of the current levels of cycling compared to a 'no cyclists' counterfactual (see Section A1.4).

[1] see https://www.ordnancesurvey.co.uk/business-and-government/products/terrain-50.html

Table 1: Summary of parameter and propensity estimation across different types of OD pairs in the PCT model

Type of OD pair	% of OD pairs	% of comm-uters	% of cyclists at baseline	Included in count of cyclists at baseline?	Modelled as increasing in scenarios?	Fastest-route cycling distance for commuters	Gradient of fastest cycling route for commuters	Inputs to propensity to cycle calculation	Distance travelled by cyclists, for health and carbon impacts
Type 1: <30km, between MSOAs	44.1%	69.6%	78.0%	Yes	Yes	From CycleStreets	From CycleStreets	Distance + gradient	Equal to fastest-route distance in OD pair
Type 2: within MSOAs	0.3%	9.3%	13.3%	Yes	Yes	1/3 mean distance in shortest 3 between-MSOA pairs	Mean gradient of shortest 3 between-MSOA pairs	Distance + gradient	Equal to fastest-route distance in OD pair
Type 3: No fixed workplace	0.3%	9.1%	4.9%	Yes	Yes	Not estimated	Not estimated	MSOA mean propensity to cycle in type 1 and 2 OD pairs	Mean distance of cyclists in type 1 and 2 OD pairs <10km apart in the MSOA in question†
Type 4: >30km within England or Wales, or workplace outside England or Wales	55.3%	12.0%	3.9%	Yes	No	Not estimated	Not estimated	Not estimated	Mean distance of cyclists in type 1 and 2 OD pairs nationally‡

† Among individuals who said cycling was their usual main commute mode in the English and Welsh NTS, the average total cycle commute distance across the week was similar between those who had a fixed workplace <10km from their home (N=1101, 24.6km) versus those who worked at different places (N=136, 25.7km).
‡The English and Welsh NTS did not provide adequate data on workplace location to test this assumption, but what testing was possible indicated that this assumption may be somewhat conservative. Specifically, among individuals who said cycling was their usual main commute mode, the average total cycle commute distance across the week was lower between those who had a fixed workplace <30km from their home (N=1283, 31.6km) versus those who worked overseas or made at least one commute trip of >30km (N=31, 44.4km). Given the limitations and small sample size of this analysis, it seemed better to adopt this potentially conservative approach.

A1.3 Modelling propensity to cycle, and numbers of cyclists, across four scenarios

A1.3.1 Modelling baseline propensity to cycle in the 2011 Census

In order to generate 'what if' scenarios regarding possible future levels of cycling, we first sought to model current propensity to cycle – i.e. the current proportion of commuters who cycle to work. Again, we did this in different ways for different types of OD pairs, as set out in Table 1.

For all within-MSOA and between-MSOA OD pairs in England and Wales with a fastest-route distance of <30km, we modelled the relationship between the proportion of commuters cycling (the dependent variable) and the fastest route distance and route gradient (the two explanatory variables). We did this using an individual-level logit model, expanding the ~1.1 million OD pairs to their constituent ~19 million commuters. Distance decay was modelled using linear, square-root and square terms (Equation 1A). The 'gradient' variable was entered as the original gradient derived from CycleStreet.net minus 0.97%, which is the estimated average route gradient in the Netherlands (see Box 1). By centring our gradient measure on the estimated Dutch average in this way, we facilitated the subsequent addition of 'Go Dutch' parameters to the baseline equation (see Section A1.3.4).

Box 1: Estimating hilliness in England, Wales and the Netherlands

We calculated the average gradient of each lower super output area (LSOA) in England and Wales using elevation data from NASA's Shuttle Radar Topography Mission ('Version 4' dataset, available at http://srtm.csi.cgiar.org/). Across the UK, the average resolution of the raster is 56.5m east-west and 92.6 m north-south. We converted the elevation data into a gradient in degrees for each raster cell (using R's 'raster' package), and aggregated these to generate the average per LSOA. We likewise calculated average hilliness of 'Neighbourhoods' in the Netherlands. These Neighbourhoods are administrative geographical units created by Statistics Netherlands that contain an average of 1424 individuals. They are therefore of a similar average size to LSOAs in England and Wales, which are designed to contain approximately 1500 individuals each. The Statistics Netherlands neighbourhoods vary considerably in size, however, and so we weighted by population when making comparisons to the distribution of LSOAs in England and Wales.

A comparison of the gradient distribution of the English and Welsh LSOAs with the Dutch Neighbourhoods confirmed that the average level of hilliness in the Netherlands is much lower than in England and Wales: for example, the average population-weighted gradient of Neighbourhoods in the Netherlands is 0.69 degrees, but only 12% of the English and Welsh population lives in Lower Super Output Areas with an average gradient of 0.69 degrees or below. We assumed that the same relationship applies to commuting routes, i.e. that the average commuting route gradient in the Netherlands is equal to the gradient experienced by commuters on the 12th percentile for gradient in England and Wales. In the English and Welsh OD data, commuters on the 12th percentile for route gradient experienced a gradient of 0.97% (as estimated by CycleStreets). We therefore subtracted 0.97% from our measure of gradient when building our propensity to cycle model, and thereby sought to centre the model on the estimated average gradient for trips in the Netherlands. In addition, when pooling English, Welsh and Dutch NTS data to estimate the relative increase in propensity to cycle in a 'Go Dutch' scenario, we weighted the English and Welsh participants such that the hilliness profile of their home area corresponded to that of the Netherlands (weights presented in final column of Table 2). In combination, these measures allowed us to estimate 'Go Dutch' parameters in a way that did not overestimate the propensity to cycle in a 'Go Dutch' scenario by ignoring the fact that the Netherlands is flatter than England and Wales.

Table 2: Distribution of hilliness in ones area of residence, in England and Wales versus the Netherlands

Twentieth of home-area gradient (England & Wales)	% English & Welsh population in twentieth (A)	% Dutch population with comparable home-area gradient (B)	Weight given to English and Welsh NTS participants, according to their home-area gradient (B/A)
1 (flattest)	5%	32.5%	6.49
2	5%	25.6%	5.12
3	5%	15.7%	3.15
4	5%	10.1%	2.01
5	5%	5.3%	1.05
6	5%	3.3%	0.66
7	5%	2.1%	0.42
8	5%	1.3%	0.26
9	5%	0.9%	0.18
10	5%	0.6%	0.12
11	5%	0.6%	0.12
12	5%	0.5%	0.09
13	5%	0.4%	0.08
14	5%	0.3%	0.05
15	5%	0.3%	0.06
16	5%	0.2%	0.04
17	5%	0.2%	0.05
18	5%	0.1%	0.02
19	5%	0.1%	0.01
20 (hilliest)	5%	0.0%	0.01

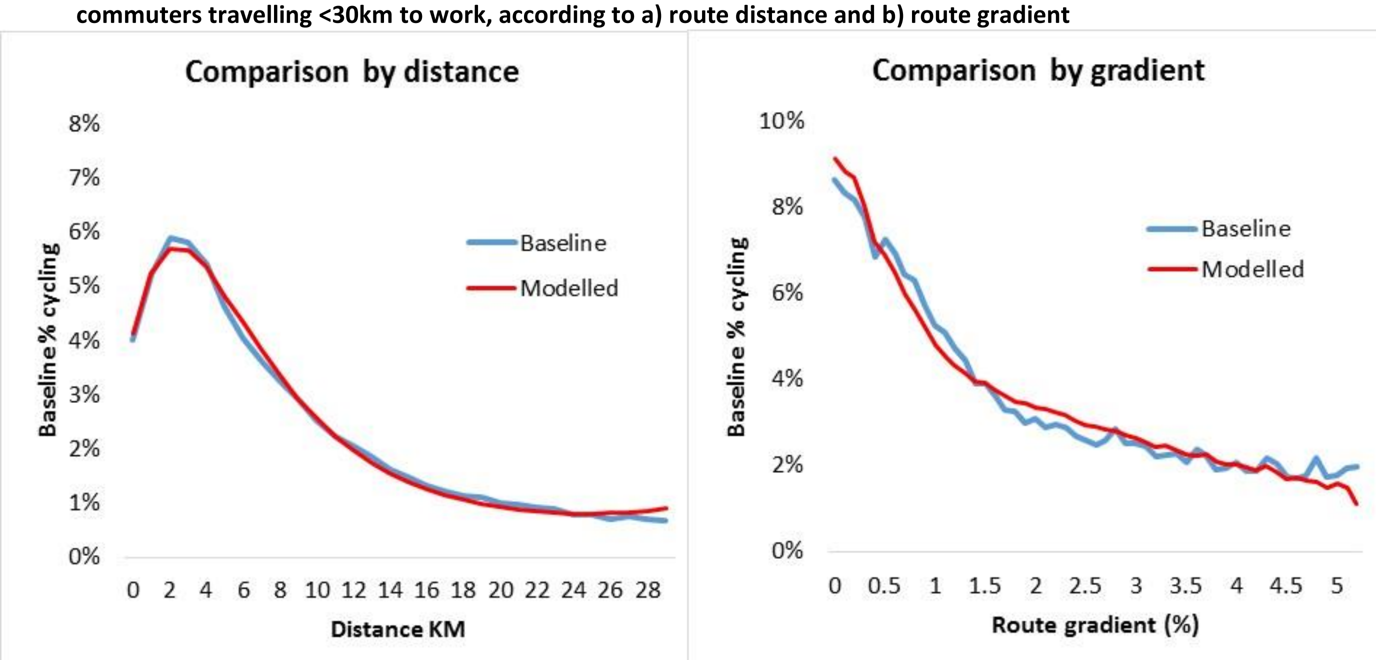
Interaction terms were included to capture the relationship between distance and gradient, on the basis of evidence that the deterrent effect of a steeper slope was somewhat stronger for individuals travelling intermediate distances (Coefficient for gradient (change per 1% increase) -0.32 (SD=0.001) for distances 0-4.9km; -0.35 (SD=0.003) for distances 5-9.9km; -0.37 (SD=0.005) for distances 10-14.9km; -0.30 (SD=0.008) for distances 15-19.9km; and -0.26 (SD=0.009) for distances 20-29.9km).

The resulting equation for baseline propensity to cycle was:

Equation 1A: logit (pcycle) = -3.959 + (-0.5963 * distance) + (1.866 * $distance_{sqrt}$) + (0.008050 * $distance_{sq}$) + (-0.2710 * gradient) + (0.009394 * distance*gradient) + (-0.05135 * $distance_{sqrt}$ *gradient)
pcycle = exp ([logit (pcycle)]) / (1 + (exp([logit(pcycle)])

where 'pcycle' is the proportion of cyclists expected; 'distance' is the fastest route distance in km, '$distance_{sqrt}$' and '$distance_{sq}$' are, respectively the square-root and square of distance; and 'gradient' is the fastest-route gradient (centred on 0.97%). Note that although this equation was derived at the individual level, it can be applied at the level of the OD pairs as distance and gradient are constant within OD pairs. Equation 1A showed good fit to the observed data with respect to both distance and hilliness (Figure 1).

Figure 1: Observed versus predicted prevalence of cycling to work among 18,871,463 English and Welsh commuters travelling <30km to work, according to a) route distance and b) route gradient

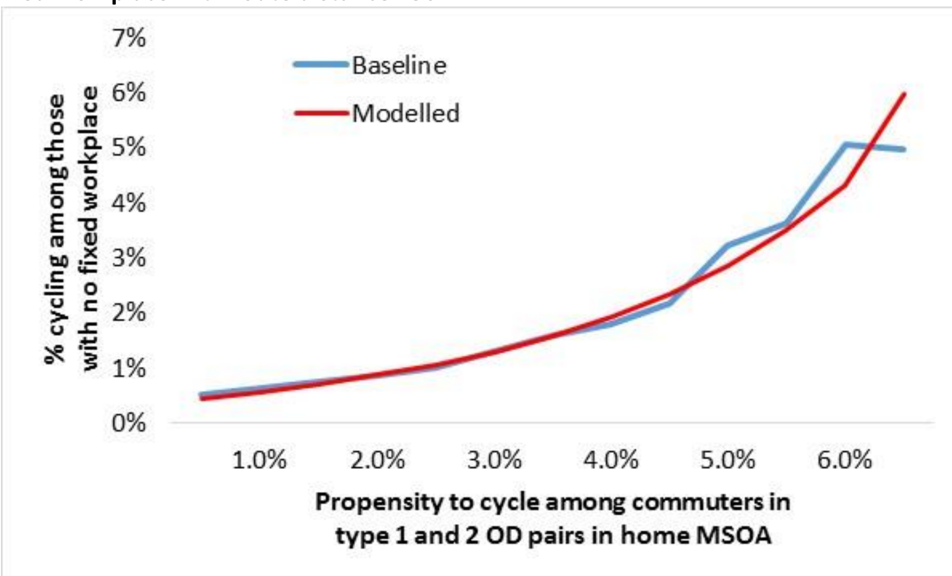


For commuters with no fixed workplace, we modelled propensity to cycle as a function of the average propensity to cycle among commuters living in the same MSOA and commuting <30km. The resulting equation for baseline propensity to cycle among those with no fixed workplace was:

Equation 2A: logit (pcycle) = -6.399 + (184.0 * $meanpropensity_{sq}$) + (10.36 * $meanpropensity_{sqrt}$)
pcycle = exp (logit(pcycle)) / (1 + (exp(logit(pcycle))

where ‘$\text{meanpropensity}_{sq}$’ is the square of the mean propensity to cycle among type 1 and type 2 OD pairs in the home MSOA in question, and ‘$\text{meanpropensity}_{sqrt}$’ is the square root term. This resulted in the model fit shown in Figure 2.

Figure 2: Observed versus predicted prevalence of cycling to work among 2,165,685 English and Welsh commuters with no fixed work place, according to the modelled propensity to cycle among commuters with a fixed workplace with route distance <30km



Finally, we did not model baseline propensity to cycle among individuals living more than 30km from their place of work or commuting outside England or Wales. Instead, given the considerable uncertainties about where the cycling reported by these individuals was taking place, we assumed no increase in cycling levels among these commuters in our scenarios.

A1.3.2 Government Target Scenario

The ‘Government Target’ scenario models a doubling of cycling nationally, corresponding to the proposed target in the UK government's draft Cycling Delivery Plan to double cycling between 2013 to 2025 [3]. To model the total number of cyclists in the Government Target scenario propensity to cycle (‘pcycle’) in each OD pair was estimated using the equations set out in Section A1.3.1. After multiplying this by the total number commuters in the OD pair this value was added to the recorded number of cyclists in the 2011 Census (see Table 3).

This is scenario is illustrated by the following example. Take an OD pair of 200 commuters containing 7 cyclists in the 2011 Census, and with a modelled propensity to cycle of 5.2% (from Equation 1A). The total number of cyclists in the Government Target scenario would be 7 + (200 * 0.052) = 7 + 10.4 = 17.4. Thus the Government Target scenario leads to a doubling of cyclists in England and Wales but not necessarily of each OD pair. Note the reported ‘baseline’ number of cyclists directly influences the total number of cyclists in the scenario (column B2 in Table 3), but does not influence the scenario increase in the number of cyclists (Column C).

Table 3: Summary of scenario generation rules

Scenario	Baseline no. cyclists (A)	Initial estimation of scenario no. cyclists (B1)	Additional processing of scenario no. cyclists (B2)	Scenario increase in no. cyclists (C)
Government Target	Recorded no. in Census 2011, OD pair types 1-4.	Column A + (Baseline propensity to cycle [Equations 1A+2A]† in OD pair types 1-3 * no. commuters)	• Cap Column B1 at 100%.	Column B2 – Column A
Go Dutch	Recorded no. in Census 2011, OD pair types 1-4.	'Go Dutch' propensity to cycle [Equations 1B+2B, with 'dutch'=1 and 'ebike'=0] in OD pair types 1-3 * no. commuters.	• Set Column B1 as equal to Column A if B1 is less than A.	Column B2 – Column A
Ebikes	Recorded no. in Census 2011, OD pair types 1-4.	'Ebikes' propensity to cycle [Equations 1B+2B, with 'dutch'=1 and 'ebike'=1] in OD pair types 1-3 * no. commuters.	• Set Column B1 as equal to Column A if B1 is less than A.	Column B2 – Column A
Gender Equity	Recorded no. in Census 2011, OD pair types 1-4.	Apply Equation 3 in OD pair types 1-3.	• Set Column B1 as equal to Column A if number of males in the OD pair is zero, or if B1 is less than A.	Column B2 – Column A

† Or, equivalently, using equations 1B + 2B in Section A1.3.3, 'dutch'=0 and 'ebike'=0

A1.3.3 Go Dutch and Ebike scenarios

The 'Go Dutch' scenario models the level of cycling expected if English and Welsh people cycled as much as people in Netherlands, taking into account differences in the distribution of hilliness and trip distances. The 'Ebike' scenario is an extension of the Go Dutch scenario, but makes the further assumption that all cyclists in the Go Dutch scenario own an ebike. For these scenarios, our approach was to start from the Equations estimating baseline propensity to cycle (Section A1.3.1) and add additional parameters.

The Go Dutch scenario required us to model the increase in propensity to cycle that would be observed if English and Welsh commuters became as likely to cycle a given trip as Dutch commuters. We estimated this additional parameter using trip-level analysis of the English and Welsh and Dutch National Travel Surveys, restricting the analysis to commute trips of less than 30km (N = 264,912 trips among 35,390 adults in the 2008-2014 English and Welsh data; N = 82,274 trips among 42,223 adults in the 2010-2014 Dutch data). Trip-level analysis was used because the necessary individual-level data (e.g. usual main commute mode) was not available in the Dutch NTS. In the English and Welsh NTS, however, the distance-decay curve at the trip level is very similar to the curve at the individual level, which gives confidence in our approach.

In estimating the increased propensity to cycle among Dutch people, we included both a main effect term and an interaction term with distance (as a linear term). We introduced the interaction term to reflect the fact that Dutch propensities to cycle exceed English and Welsh propensities by a greater amount for short distances (e.g. Dutch people are 5.5 times more likely to cycle a trip of 0-4.9km versus 3.7 times more likely to cycle a trip 10-14.9km). As hilliness data was not available in the Dutch survey, we weighted the data so that the English and Welsh sample of commuters lived in areas with the same hilliness profile as the Dutch commuters (see Box 1), to allow comparisons that were not affected by differences in average hilliness between England and Wales versus the Netherlands. In the logit model, based on 347,186 trips among 77,613 English/Welsh and Dutch commuters, the coefficient for a main effect with Dutch (versus English) status was 2.523, while the interaction term between Dutch status and distances was -0.07626.

The Ebike scenario builds on the Go Dutch scenario and models the further increase in propensity to cycle that would be observed if all Dutch cyclists acquired an ebike. To generate the relevant parameters, we restricted our analysis to the Dutch NTS 2013-2014, the only years that measured ebikes as a separate mode. We further restricted our analysis to the 26,807 commute trips made by 13,693 adults who owned a bicycle. We then duplicated the subset of trips made by ebike owners (2175 trips by 1087 individuals) and used logit regression to compare propensity to cycle between the population of duplicated ebike-owner trips (N = 2175) with the full population of all bicycle-owner trips (N = 26,807). This analysis therefore takes into account the fact that some ebike owners are already present in the 'Go Dutch' scenario, and captures only the extra cycling that would occur if *everyone* with a traditional bicycle acquired an ebike.

In estimating the extent to which this would increase propensity to cycle in the Ebike scenario, we focussed on interaction terms with distance (as a linear and squared term). We did this to capture the fact that owning an ebike increases propensity to cycle more for long trips than for short trips (e.g. Dutch ebike owners are 1.1 times more likely than all Dutch bicycle owners to cycle a trip 0-4.9km versus 2.3 times more likely to cycle a trip 10-14.9km). We adjusted for age and sex to take account of the fact that at present ebike owners in the Netherlands are more likely to be female than bicycle owners in general (61% vs. 48%) and are also older on average (mean age 54 years vs. 43 years). The magnitude of the interaction term between ebike status and distance was 0.05710, while the magnitude of the interaction term between ebike status and distance squared was -0.0001087. Because we did not have data on hilliness in the Dutch National Travel Survey we could not estimate the magnitude of any interaction between ebike ownership and hilliness in this dataset. In addition, this might in any case not have been feasible as so little of the Netherlands contains hills.

We therefore instead estimated the interaction term between ebike use and average route gradient using data from the Swiss National Household Travel Survey 2010. In this nationally-represented household travel survey, a random sub-sample of participants were asked questions on the number of bicycles in their household in everyday use, and also on the number of ebikes in their household. In this subsample, 21,327 trips of <30km were reported among 5598 adult participants living in households that owned at least one bicycle in everyday use. As when analysing the Dutch data, we duplicated the subset of trips made by individuals living in households owning at least one ebike (798 trips <30km by 107 adults). Route distance on the road network, and the average route gradient, was estimated in ArcGIS based the on origin and destination co-ordinates of each stage, with route and gradient data drawn from www.swisstopo.admin.ch. We fitted a regression model including the same terms as used in the Dutch data (age, gender, distance, distance-squared, and interaction terms between being in the ebike population and distance) plus terms for gradient and the interaction between being in the ebike population and gradient. The main effect term of gradient in this model was -0.1313 ($p<0.001$) while the interaction term with ebikes was +0.08799 ($p=0.16$). Thus there was a trend for an interaction effect in the expected direction, with the deterrent effect of gradient being considerably smaller for ebike users, but this was not significant. Because we believe an interaction of ebike ownership with hilliness is plausible, and the non-significant result likely reflects the small number of trips in the ebike population, we decided nevertheless decided to include this parameter in our model. We did this by multiplying the main effect term for gradient in our model by the ratio of the Swiss interaction term / main effect term = +0.08799/-0.1313 = -0.67. We used this ratio rather than the absolute magnitude of the Swiss interaction term (+0.08799) because we believe the absolute magnitude of the Swiss term is plausibly diluted by the fact that ebike ownership is measured at the household-level not the individual level. We did not estimate any three-way interaction parameters for ebike population by gradient by distance, as we felt that we had too little power for this to be meaningful. We therefore decided only to apply this ratio of -0.67 to the main effect parameter for gradient, i.e. generating an ebike*gradient interaction term of -0.67 * -0.2710 = 0.1812. The results this generates are almost identical to the results generated after applying a ratio of -0.5 to the gradient main effect term plus also to the interaction terms – i.e. our parameterisation is effectively equivalent to assuming that owning an ebike halves the effect of hilliness.

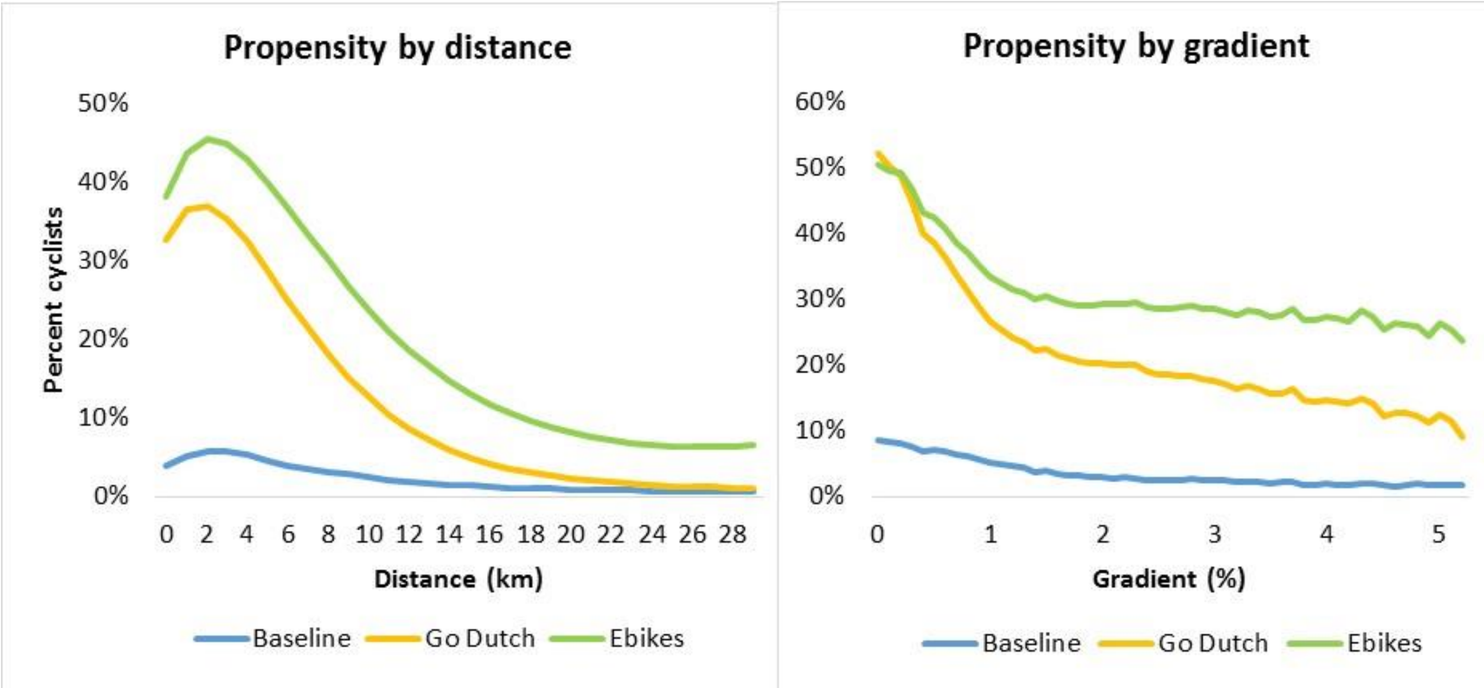
Adding all these 'Go Dutch' and 'Ebikes' parameters together, we derived the following propensity to cycle equation:

Equation 1B: logit(pcycle) = Equation 1A + Dutch parameters + Ebike parameters

logit(pcycle) = = -3.959 + (-0.5963 * distance) + (1.866 * $distance_{sqrt}$) + (0.008050 * $distance_{sq}$) + (-0.2710 * gradient) + (0.009394 * distance*gradient) + (-0.05135 * $distance_{sqrt}$ *gradient) + (2.523 * dutch) + (-0.07626 * dutch * distance) + (0.05710 * ebike * distance) + (-0.0001087 * ebike * $distance_{sq}$) + (0.1812 * ebike * gradient).

where 'pcycle' is the proportion of cyclists expected; 'distance' is the fastest route distance in km, '$distance_{sqrt}$' and '$distance_{sq}$' are, respectively the square-root and square of distance; 'gradient' is the fastest-route gradient (centred on 0.97%); 'Dutch' is a binary variable that takes the value '0' for the Government Target scenario and '1' for the Go Dutch or the Ebike scenario; and 'ebike' is a binary variable that takes the value '0' for the Government Target and Go Dutch scenario and '1' for the Ebike scenario. Figure 3 shows the distribution of cycling propensity generated by Equation 1B, according to distance and hilliness.

Figure 3: Prevalence of cycling to work at baseline among English and Welsh commuters travelling <30km to work, and modelled prevalence of cycling to work in Go Dutch and Ebike scenarios, according to a) route distance and b) route gradient



For commuters with no fixed workplace, we similarly started with Equation 2A, and extended this as follows.

Equation 2B: logit(pcycle) = Equation 2A + mean Dutch parameters + mean Ebike parameters
= -6.399 + (184.0 * meanpropensity$_{sq}$) + (10.36 * meanpropensity$_{sqrt}$) + (dutch * meandutch) + (ebike * meanebike)

where 'meanpropensity$_{sq}$' is the square of the mean propensity to cycle among type 1 and type 2 OD pairs in the home MSOA in question, and 'meanpropensity$_{sqrt}$' is the square root term; 'meandutch' is the average value of the Equation 1B Dutch parameters for commuters living in the same home MSOA; and 'meanebike' is the average value of the Equation 1B Ebike parameters for commuters living in the same home MSOA.

A1.3.4 Gender Equity Scenario

In the 2011 Census, women accounted for 48% of all English and Welsh commuters but only 27% of all cycle commuters. This gender disparity is seen across the country, with no local authority having a proportion of female cyclists greater than 50%. The 'Gender Equity' scenario seeks to capture a situation in which this gender disparity was eliminated. In this respect, it differs somewhat from the preceding three scenarios, as it does not use distance and hilliness data to model propensity to cycle. Instead it assumes that male propensity to cycle remains unchanged – i.e. there is no change in the number of male cycle commuters – and that female propensity to cycle rises to match male propensity. In other words, we sought to model a situation in which, in any given OD pair, the proportion of females cycling rises to match the proportion of males cycling:

$$SNcyclists_f \ / \ BNcommuters_f = BNcyclists_m \ / \ BNcommuters_m$$

where '$SNcyclists_f$' is number of female cycle commuters in the gender equality scenario, '$BNcyclists_m$' is the recorded number of male cycle commuters at baseline, and '$BNcommuters_f$' and '$BNcommuters_m$' are the total numbers of females and males in the OD pair respectively. This allows estimation of the total number scenario number of cyclists ('SNcyclists') as follows:

$$SNcyclists_f = BNcommuters_f * BNcyclists_m \ / \ BNcommuters_m$$
$$SNcyclists = BNcyclists_m + SNcyclists_f$$
$$SNcyclists = BNcyclists_m + (BNcommuters_f * BNcyclists_m \ / \ BNcommuters_m)$$

<u>Equation 3:</u>
$$SNcyclists = BNcyclists_m * (1 + (BNcommuters_f \ / \ BNcommuters_m))$$

To illustrate how this method works in practice, imagine an OD pair in which 50 from a total of 500 people commute by cycle, 35 males ($BNcyclists_m = 35$) and 15 females ($BNcyclists_m = 15$). 300 of the total trips in the OD pair are made by males ($BNcommuters_m = 200$) and 200 by females ($BNcommuters_f = 200$). Applying Equation 3:

$$SNcyclists = BNcyclists_m * (1 + (BNcommuters_f \ / \ BNcommuters_m))$$
$$SNcyclists = 35 * (1 + (200 / 300))$$
$$= 58.3$$

All of these extra 8.3 cyclists are female, giving a new total of 15 + 8.3 = 23.3 female cyclists (and still 35 male cyclists). Gender equality in cycling has been reached, such that 11.7% of commute trips are made by cycling among both men (35/300) and women (23.3 / 200). Equation 3 was applied to commuters with 'no fixed workplace' in the same way, and as in other scenarios we assumed no change among commuters travelling >30km or outside England and Wales.

A1.4: Modelling scenario mode shift in walking and car driving

To estimate the health impacts of our scenarios, we needed to estimate the number of new cyclists who had previously commuted on foot. Similarly, in order to estimate the carbon impacts of our scenarios, we needed to estimate the number of new cyclists who had previously commuted as car drivers (note that we specifically focus on car drivers, not car passengers, as the standard practice in estimating transport CO_2 emissions is to attribute all emissions to the car driver, to avoid double-counting). We assumed that within any given OD pair commuters were equally likely to shift to cycling from any baseline mode, and therefore the mode shift was proportional to mode share at baseline. For example, take an OD pair containing 220 commuters at baseline, of whom 20 cycle, 80 walk, 50 are car drivers and 70 use other modes of transport. If the 'Government Target' scenario number of cyclists rose to 50 in this OD pair, this would mean that the number of non-cyclists decreased to 170, giving a ratio of change among non-cyclists of 170 / 200 = 0.85. We assume this 0.85 applies to all modes, and therefore the number of pedestrians in the scenario 80 * 0.85 = 68; the number of car drivers in the scenario is 50 *0.85 = 42.5; and the number of commuters using other modes in the scenario is 70 * 0.85 = 59.5.

For the purposes of estimating health and carbon impacts of the current level of cycling relative to a ‘no cycling’ counterfactual, we made the same assumption. For example, again take the OD pair containing 220 commuters at baseline, of whom 20 cycle, 80 walk, 50 are car drivers and 70 use other modes of transport. In a ‘no cyclists’ counterfactual, the number of non-cyclists would increase to 220, giving a ratio of change among non-cyclists of 220 / 200 = 1.1. Thus in the ‘no cyclists’ counterfactual, the number of pedestrians would be 80 * 1.1 = 88, and so on. When estimating mode split in the ‘no cyclists’ counterfactual in the 5057 of OD pairs that at baseline consisted entirely of cyclists, we assumed a mode split of 31% walking, 35% car drivers and 34% other modes. These percentages correspond to the observed mode split among the 974 OD pairs in which 50-99% of individuals cycled in the 2011 Census.

A1.5: Estimating the physical activity health benefits

We used the World Health Organisation’s Health Economic Assessment Tool (HEAT, http://www.heatwalkingcycling.org/) to estimate the number of deaths avoided due to increased physical activity [4].

A1.5.1: Deaths avoided due to increased cycling

In line with the HEAT guidelines, our first step in estimating health benefits was to calculate the average weekly duration of cycling conducted by cycle commuters in a given OD pair. This average duration was calculated as equal to:

Weekly cycling duration among cycle commuters in a given OD pair
= (cycling commute distance * mean cycle commute trips per cyclist per week in a typical week) / mean cycling speed

Commute distance, commute trips per cycle commuter week in a typical week and cycling speed were estimated as described in Table 4 (parameters A, E and F). In the Go Dutch and Ebikes scenarios, we additionally drew on an estimate of mean Ebike cycling speed (parameter H) and multiplied the weekly duration of cycling by an activity intensity-related distance reduction factor for ebikes (parameter I). We applied these speed and activity-intensity parameters to the proportion of cycle trips in each OD pair that was estimated to be made by ebike (Table 4, parameters J and K). We used these adjustment factors to take account of the fact that cycling by ebikes is a) faster and b) less energy intensive than cycling on a normal bicycle, and so one needs to travel further to get the same physical activity benefits. For scenarios involving ebikes, our ‘weekly cycling duration’ values therefore do not capture the actual number of minutes spent cycling but instead capture a weekly duration of pedal cycling that would be expected to yield the same physical activity benefits.

After we had estimated the weekly duration of cycling among cycle commuters in each OD pair, we followed standard HEAT procedure and calculated a mortality protection effect for cyclists in each OD pair as being equal to:

Mortality protection for cycle commuters in a given OD pair
= (1-mortality reduction for reference cycling duration)*(weekly cycling duration / reference cycling duration)

The mortality reduction factor for cycling, and the reference duration of cycling are both fixed input parameters specified by HEAT (Table 4, parameters L and M). Note that the value of the mortality protection factor for cycling is capped at as a maximum of 0.45 for HEAT 2014. This cap is implemented in HEAT to account for the fact that additional physical activity confers diminishing health benefits if starting from a higher baseline activity level and to avoid implausibly large reductions in mortality risk from large increases in mortality. The estimated mortality protection factor was then used to estimate the number of deaths avoided per year as follows:

No. deaths avoided per year due to cycle commuting in each OD pair
= change in no. cycle commuters * annual mortality rate of cycle commuters * mortality protection for cycle commuters

The input parameters on the annual mortality rates of cyclists are summarised in Table 4 (parameters P, Q, R). For these, we first calculated age- and sex-specific mortality rates for each local authority in England and Wales using data published by the Office for National Statistics on deaths in 2014 [5] and the 2014 mid-year population [6]. We did so for all ages from 16-74 in five-year age bands (16-19, 20-24, 25-29 and so on up to 70-74). We then weighted these mortality rates by the age- and sex-profile of the relevant population of cyclists, i.e. to the population of cyclists whose numbers were changing in the scenario in question. When estimating the health impacts of current cycling levels against a 'no cycling' counterfactual, the relevant population was those cyclists recorded in the 2011 Census. The same was true for the Government Target scenario, as we assumed that the age- and sex- profile of new cyclists in this scenario would match that of existing cyclists in the 2011 Census. For the Gender Equity scenario we assumed that all new cyclists were female but with the age profile of existing cyclists in the 2011 Census. Finally, for the Go Dutch and Ebikes scenarios we assumed that the profile of cyclists matched those currently observed among cycle commuters in the Netherlands.

When seeking to compare the number of deaths avoided at baseline (in comparison to a 'no cycling' counterfactual), the change in the number of cycle commuters was equal to the recorded number of cyclists in the 2011 census. When estimating the number of deaths additionally avoided in the scenario in question as compared to baseline, the change in the number of cycle commuters was equal to the scenario increase in the number of cyclists (Column C in Table 3). When estimating the number of deaths avoided in a scenario versus a 'no cycling' counterfactual, the number of deaths avoided at baseline was added to the number of deaths additionally avoided in the scenario in question.

A1.5.2: Deaths previously avoided due to former walking among former pedestrians shifted to cycling

In calculating the total mortality impacts of a given scenario, one needs to offset the physical activity gained through increased cycling against the physical activity lost through any decrease in walking, since some new cyclists will formerly have walked to work. Among these former pedestrians, we calculated the weekly duration of walking that was displaced by cycling by substituting former walking distance and walking speed (Table 4, parameters B and G) for cycling speed, i.e.:

Displaced weekly walking duration among former pedestrians in a given OD pair
= (former walking commute distance * mean cycle commute trips per cyclist per week) / mean walking speed

We likewise substituted in the equivalent walking parameters for mortality reduction and reference duration (parameters N and O) to estimate the former mortality protection among former pedestrians as:

Former mortality protection for former pedestrian commuters in a given OD pair
= (1 - mortality reduction for reference walking duration)*(displaced weekly duration of walking / reference walking duration)

Note that the value of the mortality protection factor for walking is capped at as a maximum of 0.30 for HEAT 2014. We then estimated the number of deaths formerly avoided due to walking among former pedestrians as being equal to:

No. deaths formerly avoided due to walking among former pedestrians in each OD pair
= Change in no. pedestrians * annual mortality rate of cycle commuters * mortality protection for former pedestrians

with the change in the number of pedestrians (parameter T) being estimated using the mode shift calculations described in Section A1.4.

A1.5.3: Net change in deaths, and associated monetary value

Putting together our calculating relating to new cyclists and former pedestrians, the net effect on number of deaths was:

Net change in no. deaths avoided in each OD pair =
No. additional deaths due to reduced walking - No. deaths avoided due to cycle commuting

Note that this approach means that in some OD pairs the net change in the number of deaths avoided was a positive number, i.e. additional deaths were incurred. This could happen if the majority of new cyclists had formerly walked, as this may result in a decrease in the total amount of physical activity incurred by the new cyclists.

Finally, the monetary value of the mortality impact was calculated by multiplying drawing on the standard ‘value of a statistical life’ used by the Department for Transport (Table 4, parameter V), as follows:

Monetary value of the mortality impact in each OD pair =
Net change in no. deaths * -1 * value of a statistical life

Table 4: Input parameters for estimation of health impacts using HEAT, and for estimation of carbon impacts

ID	Parameter description	Used for health or carbon or both?	Parameter value	Parameter source	Comment
A	Cycling commute distance	Both	Variable by OD pair	CycleStreets fastest route or route average - see final column of Table 1.	
B	Former walking commute distance	Health	Variable by OD pair	Assumed equal to cycling commute distance (Input A).	We assumed former pedestrians previously used the same route, rather than walking a shorter distance to reach the same destination.
C	Former driving commute distance	Carbon	Variable by OD pair	Assumed equal to cycling commute distance (Input A).	We assumed former car drivers previously used the same route, rather than driving a longer distance to reach the same destination.
D	Mean cycle commute trips per cyclist per week	Carbon	5.24	English and Welsh NTS, 2008-2014.	This is the average number of cycle commute trips reported per week among people who say cycling is their usual main mode. It includes the 27% of NTS respondents who said cycling was their usual main commute mode, but reported no cycle commute trips in the past week.
E	Mean cycle commute trips per cyclist per week in a typical week	Health	7.17	English and Welsh NTS, 2008-2014.	This is the number of cycle commute trips reported per week among people who say cycling is their usual main mode*, and who reported at least one cycle commute trip in the past week.* The latter restriction is in place because the HEAT input data on mortality risk reduction is largely based on studies asking about a ‘typical week’ – which we assume will include at least one commute trip for those who say they use cycling as their usual main mode of travel to work.
F	Mean cycling speed	Health	14 km/hour	HEAT guidance 2014 [4, page 33].	Consistent with NTS 2008-2014, in which the mean speed was 13.7 km/hr for commute cycle trips among those for whom cycling is usual main mode and excluding trips with implausible speeds (defined as <2km/hr or >25km/hr).
G	Mean walking speed	Health	4.8 km/hour	HEAT guidance 2014 [4, page 16].	Consistent with NTS 2008-2014, in which the mean speed was 4.7 km/hr for commute walk trips among those for whom walking is usual main mode and excluding trips with implausible speeds (defined as >10km/hr).

ID	Parameter description	Used for health or carbon or both?	Parameter value	Parameter source	Comment
H	Mean ebike speed	Health	15.8 km/hour	Dutch NTS, 2013-2014.	In the Dutch NTS 2013-2014, mean cycling speed is 15.2km/hr for bicycle commute trips and 17.2 km/hr for ebike commute trips, i.e. the ebike speed was 17.2/15.2=1.13 times faster. We applied this to the HEAT 2014 assumed cycling speed of 14km/hour to get 14*1.13=15.8 km/hour.
I	Activity intensity-related distance reduction factor for ebikes	Health	0.648	Published literature on physical activity intensity [7, 8].	Physical activity intensity can be measured in Marginal Metabolic Equivalent Tasks (MMETs), namely the value MET rate minus 1. The estimated MMET value for ebiking is 3.5 [7], while for cycling for transport on a pedal bike it is 5.4 [8]. We scaled down the duration of ebiking by 3.5/5.4=0.648 in order to generate a duration value that captured the amount of physical activity benefit that would have been incurred by pedal bicycling. This approach of scaling by MMETs is compatible with the HEAT numbers because the current HEAT walking and cycling parameters equate to a very similar mortality benefit per MMET using our MMET rates. Specifically, HEAT assumes that cycling 100 minutes per week (input M) gives a relative risk reduction of 1-0.9 = 0.10 (input L). Since the MMET value of cycling is 5.4 [8], this indicates that 100*5.4=540 MMETS per week gives a relative risk reduction of 0.1, or 54 MMETS per week for a relative risk reduction of 1%. HEAT also assumes that walking 168 minutes per week (input O) gives a relative risk reduction of 1=0.89=0.11 (input N). Since the MMET value of walking is 3.6 [8], this indicates that 168*3.6=604.8 MMETS per week gives a relative risk reduction of 0.11, or 55.0 MMETS for a relative risk reduction of 1%.
J	Percent cycle trips made by ebikes in Go Dutch scenario	Health	Variable by OD pair, according to route distance	Dutch NTS, 2013-2014.	In the Go Dutch scenario, we assumed the percent of trips made by ebike corresponded to the recorded percentages among all cycle commute trips in the Dutch NTS 2013-2014. The values were 6% cycle trips by ebikes for trips <5km, 11% for trips 5-9.9km, 17% 10-19.9km, 23% for trips 20-30km.
K	Percent cycle trips made by ebikes in Ebikes scenario	Health	Variable by OD pair, according to route distance	Dutch NTS, 2013-2014.	In the Ebikes scenario, we assumed the percent of trips made by ebike corresponded to the recorded percentages among cycle commute trips made by ebike owners in the Dutch NTS 2013-2014. The values were 71% cycle trips by ebikes for trips <5km, 92% for trips 5-9.9km, 92% 10-19.9km, 100% for trips 20-30km.

ID	Parameter description	Used for health or carbon or both?	Parameter value	Parameter source	Comment
L	Mortality reduction for reference cycling	Health	0.9	HEAT guidance 2014 [4, page 14].	Reduced relative risk = 1-0.9 = 0.1 for the reference duration of cycling. After scaling for the actual duration of cycling this reduced relative risk was capped at 0.45
M	Reference cycling duration	Health	100 min/week	HEAT guidance 2014 [4, page 14].	
N	Mortality reduction for reference walking duration	Health	0.89	HEAT guidance 2014 [4, page 14].	Reduced relative risk = 1-0.89 = 0.11 for the reference duration of walking. After scaling for the actual duration of walking this reduced relative risk was capped at 0.30
O	Reference walking duration	Health	168 min/week	HEAT guidance 2014 [4, page 14].	
P	Background annual mortality rate for existing cyclists at baseline, and for new cyclists in Government Target scenario	Health	Variable by OD pair, according to local authority	Mortality rate for 16-74 year olds in England and Wales in 2014, weighting by the age/sex profile of current commuter cyclists.	For inputs P to R, we calculated mortality rates using data published by the Office for National Statistics on deaths [5] and the mid-year population [6] for each local authority in England in 2014. For each local authority, we took a) age/sex-specific 2014 mortality rates and b) the age/sex profile of cycle commuters in the 2011 census. We combined these to estimate the background mortality rate for existing cyclists, and for new cyclists in the Government Target scenario.
Q	Background annual mortality rate, Go Dutch and Ebikes scenarios	Health	Variable by OD pair, according to local authority	Mortality rate for 16-74 year olds in England and Wales in 2014, weighting by the age/sex profile of Dutch commuters cyclists.	For each local authority, we took a) age/sex-specific 2014 mortality rates and b) the age/sex profile of cycle commuters in the Dutch NTS 2010-2014. We combined these to estimate the background mortality rate for new cyclists in the Go Dutch and Ebike scenarios.
R	Background annual mortality rate for new cyclists, Gender equity scenario	Health	Variable by OD pair, according to local authority	Mortality rate for 16-74 year old women in England and Wales in 2014, weighting by the age profile of current female commuter cyclists.	For females in each local authority, we took a) age-specific 2014 mortality rates and b) the age profile of female cycle commuters in the 2011 census. We combined these to estimate the background mortality rate for new cyclists in the Gender Equity scenario.
S	Change in no. cycle commuters	Both	Variable by OD pair and by scenario	At baseline, equal to Census 2011. In scenarios, equal to the scenario-increase in cycling, see Table 3	
T	Change in no. former pedestrians	Health	Variable by OD pair and by scenario	Mode shift estimation described in Section 2.4	
U	Change in no. former car drivers	Carbon	Variable by OD pair and by scenario	Mode shift estimation described in Section 2.4	Note that we specifically focus on car drivers, not car passengers, as the standard practice in estimating transport CO_2 emissions is to attribute all emissions to the car driver, to avoid double-counting

ID	Parameter description	Used for health or carbon or both?	Parameter value	Parameter source	Comment
V	Value of a statistical life	Health	£1,855,315	DfT standard value of a statistical life, in 2014 money [10].	Calculated in 2014 money, drawing on published figures by the DfT's 'WebTAG'[10]. Note that to be consistent with other DfT 'value of a statistical life' calculations we used this same parameter in relation to HEAT 2014, even though the HEAT 2014 tool uses the a considerably higher value of £3,229,114 [11]
W	CO_2-equivalent emissions per km	Carbon	0.000186	DEFRA 2016 [12].	This is the 2015 value for an 'average' car of 'unknown' size in the UK government's carbon conversion factors [12].

CO_2= carbon dioxide; DEFRA=Department for the Environment, Food and Rural Affairs; DfT=Department for Transport; HEAT=Health Economic Assessment Tool; MET=Metabolic Equivalent Task; NTS=National Travel Survey; OD pair =origin-destination pair

A1.6: Estimating reductions in transport carbon dioxide emissions from car driving

When comparing each scenario to baseline, we estimated the reduction in transport carbon dioxide (CO_2) emissions as follows:

Change in CO_2-equivalent emissions (in kg) per year
= Change in no. car drivers * former distance travelled by former car drivers * mean cycle commute trips per cyclist per week * 52.2 * CO_2-equivalent emissions (in kg) per kilometre

The change in the number of car drivers was estimated using the mode shift calculations described in Section A1.4 (Table 4, parameter U). Their average former distance was assumed to be equal to the new 'fastest route' distance travelled by the cycle commuters (parameter C). The mean cycle commute trips per cyclist per week was estimated to be 5.24 (parameter D), meaning that the mean cycle commute trips per year was 5.24 * 52.2 = 273.5. The average CO_2-equivalent emission per kilometre car driving was taken as 0.186kg, which is the 2015 value for an 'average' car of 'unknown' size in the UK government's carbon conversion factors [12] (Table 4, parameter U).

A1.7: Aggregating results to the level of bidirectional OD pairs and areas

Sections A1.2 to A1.6 all involve estimating variables for directional OD pairs. To aggregate these values to correspond to bidirectional OD pairs, we added up the values in both directions between a given pair of locations (e.g. adding the values for the A-to-B OD pair with the values of the B-to-A OD pair). These bidirectional totals are what we present in our visualisation tool. In a similar way, we also aggregated the values for directional OD pairs to the area level. We initially did this by summing our outcome variables across all OD pairs with the same home MSOA. This gave us MSOA-level estimates of the total number of cycle, foot and car commuters living in each MSOA in each scenario, plus the total change in mortality and in CO_2 emissions resulting from behaviour change among residents of that MSOA. These MSOA-level values were then further aggregated to the local-authority, regional or national level, to provide estimates for a range of geographies.

References

1 Office for National Statistics. *2011 Census: Local authority response and return rates*: [Accessed 27/06/2013 from www.ons.gov.uk/ons/guide-method/census/2011/census-data/2011-census-data/2011-first-release/first-release--quality-assurance-and-methodology-papers/census-response-rates.xls] 2012.

2 Bundesamt für Statistik BfR. *Mobilität in der Schweiz: Ergebnisse des Mikrozensus Mobilität und Verkehr 2010.* Neuchâtel: BfS, ARE 2012.

3 Department for Transport. *Cycling Delivery Plan*. [Accessed 15/02/2016 from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/364791/141015_Cycling_Delivery_Plan.pdf] 2014.

4 Kahlmeier S, Kelly P, Foster C*, et al. Health economic assessment tools (HEAT) for walking and for cycling: economic assessment of transport infrastructure and policies. Methods and user guide, 2014 update*. Copenhagen: World Health Organisation 2014.

5 Office for National Statistics. *Mortality Statistics: Deaths Registered in UK by Area of Usual Residence, 2014*: [Accessed 18/2/2016 from http://www.ons.gov.uk/ons/rel/vsob1/deaths-registered-area-usual-residence/2014/index.html] 2016.

6 Office for National Statistics. *Population Estimates for UK, England and Wales, Scotland and Northern Ireland, Mid-2014*: [Accessed 18/2/2016 from http://www.ons.gov.uk/ons/rel/pop-estimate/population-estimates-for-uk--england-and-wales--scotland-and-northern-ireland/mid-2014/index.html] 2015.

7 Sperlich B, Zinner C, Hebert-Losier K*, et al.* Biomechanical, cardiorespiratory, metabolic and perceived responses to electrically assisted cycling. *European journal of applied physiology* 2012;**112**:4015-25.

8 Costa S, Ogilvie D, Dalton A*, et al.* Quantifying the physical activity energy expenditure of commuters using a combination of global positioning system and combined heart rate and movement sensors. *Prev Med* 2015;**81**:339-44.

9 Department for Transport. *TAG Unit A4.1: Social Impact Appraisal*. London: Department for Transport, Transport Analysis Guidance (TAG) 2014.

10 Department for Transport. *WebTAG databook, annual parameters. Autumn 2015 release v1.4*. London: Department for Transport 2015.

11 HEAT. *HEAT for cycling: user interface*. London: Accessed 18/2/2016 from http://heatwalkingcycling.org/index.php?pg=cycling&act=start 2016.

12 DEFRA. *DEFRA Carbon Factors: UK Government conversion factors for Company Reporting, 2015, V2.0* Department for the Environment, Food and Rural Affairs [Accessed 18/2/2016 from http://www.ukconversionfactorscarbonsmart.co.uk/] 2015.