Shifting towards healthier transport: carrots or sticks? Systematic review and meta-analysis of population-level interventions

Christina Xiao, Esther van Sluijs, David Ogilvie, Richard Patterson, Jenna Panter

Summary

Background Promoting active travel can be beneficial for both health and the environment. However, evidence about the most effective strategies is inconsistent. We aimed to compare the effectiveness of interventions with positive (ie, carrot), negative (ie, stick), or a combination of strategies on changing population-level travel behaviour. We also aimed to identify which intervention functions, or mechanisms of how interventions seek to alter behaviour (eg, by addressing safety or accessibility), affect transport outcomes.

Methods For this systematic review and meta-analysis, we searched eight online databases for studies published before March 28, 2022: Web of Science, MEDLINE, Scopus, Applied Social Sciences Index and Abstracts, Global Health, PsycINFO, CINAHL, and Transport Research International Documentation. We did not restrict searches by language or publication date. We included controlled before-and-after studies of population-level interventions and travel behaviours (ie, driving, public transport, walking, and cycling) from adults in the general population. We categorised interventions according to their function. Depending on whether gains or losses due to intervention function could occur, we classified interventions as carrot (eg, new bike-share programmes), stick (eg, congestion charging), or combined carrot-and-stick interventions (eg, pedestrianising areas by use of reallocated parking space).

Findings From 38 916 records screened, 102 reports describing 121 interventions met the inclusion criteria. 79 interventions were carrots, 22 were carrot-and-sticks, and 20 were sticks. Results for carrot interventions were less consistent than for stick or combined interventions. Findings from the meta-analysis (64 reports describing 67 interventions) agreed with those in the narrative synthesis; although effects were statistically non-significant, for driving outcomes, interventions with stick strategies (standardised mean difference [SMD] –0·17, 95% CI –0·36 to 0·02) and combined carrot-and-stick strategies (–0·13, –0·47 to 0·20) had point estimates of greater magnitude than those for interventions with carrot strategies (–0·10, –0·23 to 0·03). Likewise, for active travel outcomes, combined carrot-and-stick strategies had a higher point estimate (0·33, –0·01 to 0·68) compared with carrot interventions (0·08, –0·05 to 0·21). Functions thought to change behaviour using financial means were effective at decreasing driving behaviour, whereas those improving access, safety, and space were effective for increasing active travel outcomes.

Interpretation This Article found that, although transport interventions with only positive strategies are more commonly evaluated, interventions that combine both positive and negative strategies might be more effective at encouraging alternatives to driving at the population level. Further research is needed for interventions involving a stick strategy, which remain less widely implemented or well studied than those with only carrot strategies.

Introduction Active travel, or transport that includes walking or cycling, might increase population activity by giving opportunities to incorporate more activity into daily routines.7 Active commuters are more likely to meet physical activity guidelines and are at lower risk of cardiovascular disease, cancer, and mortality.7 Moreover, replacing some car trips with active modes can reduce vehicular emissions and respiratory health conditions, thus providing cobenefits for health and the environment.3,4 As such, interventions to encourage active travel and discourage private car use have been made use of as part of COVID-19 physical distancing (eg, temporary cycling lanes) and recovery strategies (eg, the UK Government’s “build back fairer” or “greener” schemes), creating opportunities to build resilience to future planetary health challenges.37 Despite the advantages of active travel, evidence on how best to increase it remains inconsistent, thus making it unclear which policies should be prioritised. One reason could be the challenge of synthesising...
Evidence before this study
A preliminary search of the literature was done on March 10, 2020, by searching MEDLINE and Google Scholar to identify existing English language systematic reviews of studies evaluating interventions to change transport behaviour, with no date restrictions. Search terms included those related to transport interventions ("built environment", "financial incentives OR disincentives", "road charging", "parking pricing", "car free zone", "speed limits", "traffic calming", "awareness", "campaign"), transport modes ("active travel", "walking", "cycling", "driving", "public transport", "modal share", "modal shift"), terms indicating studies evaluating interventions ("intervention", "evaluation", "policy", "experiment", "change", "trial"), and study type ("systematic review", "meta-analysis").

Evidence has suggested that policies encouraging active travel and discouraging motorised transport can provide numerous health and environmental benefits. However, previous reviews on promoting active travel have reported that the evidence on interventions is too diverse and inconsistent for conclusive statements about effectiveness to be made, although a few narrowly defined systematic reviews have explored the value of synthesising evidence on interventions by their function (eg, affecting behaviour through financial means or environmental changes affecting access, convenience, and safety) to allow for more generalisable comparisons. Two other reviews have also grouped transport interventions by whether they involve carrot strategies (ie, those which positively motivate behaviour change) or stick strategies (ie, those which negatively motivate behaviour change). These reviews have rarely, however, made comparisons between the two strategy types in terms of their effectiveness. Evidence suggests that packages of multiple transport interventions could be more effective at inducing shifts from driving to active forms of travel than their individual components, although no reviews have quantitatively assessed this claim.

Added value of this study
This Article develops a novel classification system that builds on the concept of carrots and sticks in which interventions are categorised based on how they seek to change health behaviours. Data from 121 interventions spanning five continents were synthesised using harvest and forest plots. These plots showed that interventions with a stick strategy, alone or combined with carrot strategies, might be more effective at reducing driving and increasing active travel behaviour than interventions with a carrot strategy alone. However, interventions with stick strategies were much less often evaluated than those with carrot strategies. Moreover, targeting of specific functions was found to be potentially effective for different travel outcomes; interventions that alter behaviour using financial means were most effective for discouraging driving behaviour, whereas interventions that increase access, safety, and space were most effective for changing active forms of travel.

Implications of all the available evidence
The results of our systematic review and meta-analysis suggest that interventions with stick strategies, both alone and in combination with carrot strategies, are not only less represented but potentially more effective than interventions with carrot-only strategies. Along with our findings identifying effective intervention functions, this information can be used to inform future research and support policies seeking to increase population-level active travel.

Methods
Search strategy and selection criteria
This systematic review and meta-analysis was reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.8–10 We...
registered the review protocol under the Open Science Framework.

We searched eight online databases for studies published before March 28, 2022: Web of Science, MEDLINE, Scopus, Applied Social Sciences Index and Abstracts, Global Health, PsycINFO, CINAHL, and Transport Research International Documentation. Although the searches were done in English, we did not restrict search results by date, context, or language. We also considered grey literature and searched the reference lists of relevant reviews for eligible studies. Search terms were informed by previous systematic reviews and related to specific transport interventions or policies (eg, building new active travel infrastructure) and transport behaviours (eg, driving, public transport, and active travel).20,21 Public transport was included as some walking or cycling is typically required to access it.22 The full search strategy can be found in the appendix (pp 5–7).

We included studies that evaluated a population-level intervention (defined as a non-targeted intervention affecting a group of individuals, community, or population), had a control group, presented before and after data, presented data on adults in the general population, and reported any measure of travel behaviour change. The following screening steps were done using Covidence software. After removal of duplicates, CX performed title and abstract screening according to the inclusion criteria and RP checked a 10% sample of exclusion decisions. The same two reviewers performed full-text screening and identified and categorised reasons for exclusion. Disagreements were resolved after discussions amongst the wider review team. Following full-text screening, CX did the data extraction using a pre-piloted data extraction form, with RP checking a 10% random sample of reports. Extracted data included information on the setting, intervention description, outcome characteristics, follow-up time, study design, and main study effects for each mode of transport.

Quality assessment
Pairs of reviewers (CX and JP or EvS) independently assessed the methodological quality of each report using a modified version of the Effective Public Health Practice Project (EPHPP) tool.23 The EPHPP tool was used instead of the Cochrane Risk of Bias in Non-randomized Studies-of Interventions (ROBINS-I) due to uncertainty over the suitability of ROBINS-I to assess natural experiments.24 Further descriptions of the quality assessment methodology can be found in the appendix (p 8). Discrepancies were resolved after discussion and all reports, irrespective of their quality score, were eligible to be included in the review.

Categorising interventions as carrot, stick, or combined
The definition of carrots and sticks has been subject to interpretation and existing classification schemes do not clearly distinguish between the two categories. For instance, Piatkowski and colleagues25 characterised carrot interventions from the point of view of active travellers, in which the intended proximal outcome of the intervention (ie, enabling active transport or deterring driving) is seen as either positive (ie, carrot) or negative (ie, stick). However, most individuals are not exclusively motorists or active travellers, as the mode of transport taken can depend on the destination and trip purpose. Moreover, further clarifications are needed to distinguish interventions that are combinations of carrots and sticks from those which are only one or the other. For instance, some interventions meant to encourage active travel that seem like carrots (eg, providing new pavements) might also involve a stick if they take road space away from traffic, thus discouraging driving.

We refined the categorisation of interventions aimed at encouraging active travel or discouraging driving by use of the concept of intervention function, based on how a given intervention is thought to influence behaviour (eg, by altering the cost of transport or changing accessibility or space for pedestrians, cyclists, motorists or active travellers, as the mode of transport required for a journey and providing information about routes).

Aesthetics: involves changing physical or natural features to make it more sensorially appealing or attractive to perform certain travel behaviours
Access: affects opportunities, barriers, and restrictions for certain travel behaviours
Awareness: notifying or informing individuals on the benefits of active travel or the harms of driving and providing information about routes
Convenience: involves increasing or decreasing the time, distance, or number of stages required for a journey
Financial: changes to whether an individual needs to pay more or less to perform certain travel behaviours
Safety: relates to the traffic safety of transport users, either by influencing the behaviour of drivers or by signalling to pedestrians and cyclists that spaces are safe to use Skills: influences the ability of individuals to do certain travel behaviours, such as by providing knowledge or instilling confidence
Space: changing physical boundaries or delineating areas to perform certain travel behaviours

Panel 1: Definition of intervention functions

For review registration see https://doi.org/10.17605/OSF.IO/8X3VU
See Online for appendix

For Covidence see www.covidence.org
interventions producing gains as carrots, those producing losses as sticks, and those producing both gains and losses as combined carrot-and-stick interventions (table 1).

### Table 1: Summary of intervention, function, potential effect, and carrot or stick classification

| Function                                      | Potential effect | Carrot or stick |
|-----------------------------------------------|------------------|-----------------|
| Cycle training programme                     | Skills           | Carrot          |
| Financial reward for active travel or reducing driving | Financial        | Carrot          |
| Subsidised or free bike                       |                  |                 |
| Function one                                  | Financial        | Carrot          |
| Function two                                  | Access           | Carrot          |
| Increased green space                         | Aesthetics       | Carrot          |
| New dedicated traffic signalling for pedestrians or cyclists | Safety           | Carrot          |
| New or extension of bike-share programme      | Access           | Carrot          |
| New or extension of e-scooter programme       | Access           | Carrot          |
| New, extension, or improvement of public transport line or stop | Access           | Carrot          |
| New or improved walking or cycling infrastructure | Access           | Carrot          |
| Provision of a park-and-ride facility         | Access           | Carrot          |
| Provision of information about active travel  | Awareness        | Carrot          |
| Signage related to active travel              | Awareness        | Carrot          |
| Simplifying public transit ticketing          | Convenience      | Carrot          |
| Subsidised carpooling                         | Financial        | Carrot          |
| Subsidised or free public transit ticket      | Financial        | Carrot          |
| Increased parking prices                      | Financial        | Loss            |
| Reduced parking spaces                        | Access           | Loss            |
| Reduced road space for cars                   | Access           | Loss            |
| Road user charging                            | Financial        | Loss            |
| Traffic calming                               | Convenience      | Loss            |
| Traffic rule restrictions                     | Access           | Loss            |
| Car purchasing quota                          | Access           | Loss            |

### Data analysis

In both the systematic review and meta-analysis, we included only the data points with the longest follow-up period within each study and among reports describing the same intervention. We also describe active travel as a separate outcome in this Article because some studies...
had collected or reported transport outcome data by use of a composite measure of walking and cycling. In both the harvest plot and meta-analysis, we report outcomes for each carrot and stick category and each transport mode (ie, driving, public transport, active travel, walking, and cycling). All analyses were conducted in R (version 4.0.2).

**Harvest plot**

We generated harvest plot matrices comparing results between carrot and stick categories according to which hypothesis their results most supported: no change, a decrease, or an increase in each travel behaviour in the intervention group compared with the control.26 We used the harvest plot matrices as part of the process of synthesis rather than for making conclusive statements about intervention effectiveness. A more detailed explanation of the study allocation process can be found in the appendix (p 9).

**Meta-analysis**

For studies that reported outcomes as means with a precision estimate, we converted effect estimates into standardised mean differences (SMD) with 95% CIs using the R package esc.27 To synthesise effect estimates, we did a random-effects meta-analysis between each carrot and stick category and outcome mode using the R package meta to account for uncertainty in heterogeneity estimates.28 In addition, we did a meta-analysis for each intervention function category. We calculated study heterogeneity using the $I^2$ statistic, which is the percentage of the variability in effect estimates due to

| Number of interventions | Carrot (65%) | Stick (18%) | Carrot and Stick (17%) | Total (100%) |
|-------------------------|-------------|-------------|------------------------|--------------|
| Continent of implementation |            |             |                        |              |
| North America | 40 (50%) | 11 (50%) | 11 (55%) | 62 (51%) |
| Europe | 25 (31%) | 6 (27%) | 6 (30%) | 37 (31%) |
| Oceania | 6 (8%) | 1 (5%) | 2 (10%) | 9 (7%) |
| Asia | 7 (9%) | 4 (18%) | 1 (5%) | 12 (10%) |
| South America | 1 (1%) | 0 (0%) | 0 (0%) | 1 (1%) |

| Intervention setting | Carrot (65%) | Stick (18%) | Carrot and Stick (17%) | Total (100%) |
|----------------------|-------------|-------------|------------------------|--------------|
| City | 47 (59%) | 7 (32%) | 10 (50%) | 64 (53%) |
| Neighbourhood | 15 (19%) | 0 (0%) | 7 (35%) | 22 (18%) |
| Workplace | 15 (19%) | 14 (64%) | 3 (15%) | 32 (26%) |
| University | 2 (3%) | 1 (5%) | 0 (0%) | 3 (2%) |

| Study design | Carrot (65%) | Stick (18%) | Carrot and Stick (17%) | Total (100%) |
|-------------|-------------|-------------|------------------------|--------------|
| Repeat cross-sectional | 48 (60%) | 8 (36%) | 13 (65%) | 69 (57%) |
| Longitudinal pre-post | 24 (30%) | 6 (27%) | 7 (35%) | 37 (31%) |
| Randomised controlled trial | 7 (9%) | 8 (36%) | 0 (0%) | 15 (12%) |

| Follow-up period | Carrot (65%) | Stick (18%) | Carrot and Stick (17%) | Total (100%) |
|------------------|-------------|-------------|------------------------|--------------|
| <12 months | 34 (43%) | 17 (77%) | 10 (50%) | 61 (50%) |
| ≥12 months | 45 (56%) | 5 (23%) | 10 (50%) | 60 (50%) |

| Sample size | Carrot (65%) | Stick (18%) | Carrot and Stick (17%) | Total (100%) |
|-------------|-------------|-------------|------------------------|--------------|
| <100 | 8 (10%) | 4 (18%) | 1 (5%) | 13 (11%) |
| 100-1000 | 25 (31%) | 10 (45%) | 9 (45%) | 44 (36%) |
| ≥1000 | 44 (55%) | 8 (36%) | 10 (50%) | 62 (51%) |

| Number of intervention components | Mean (SD) | Median (IQR) |
|-----------------------------------|-----------|--------------|
| Mean (SD) | 1.4 (0.8) | 1.0 (1.0–1.0) |
| Median (IQR) | 1.0 (1.0–2.0) | 2.5 (1.0–6.0) |

| Number of intervention functions | Mean (SD) | Median (IQR) |
|----------------------------------|-----------|--------------|
| Mean (SD) | 2.4 (1.2) | 1.5 (0.9) |
| Median (IQR) | 2.0 (2.0–3.0) | 4.0 (3.8–5.0) |

Data are n (%), mean (SD), or median (IQR).
Figure 2: Harvest plots of transport mode outcomes by carrot and stick category

This grid depicts all categories of carrot and stick interventions consisting of three rows and three columns (representing the direction of change hypothesised for each category of intervention). Each intervention is represented by a single bar. The quality of each study is represented by the height of its bar, with a height of three units denoting strong quality, two units denoting moderate, and one unit denoting weak. Each bar is annotated with the number of intervention components introduced within each intervention. NA=not applicable.
heterogeneity. For outcomes reported in ten or more interventions we also examined small study effects or potential publication bias using a funnel plot produced by the dmetar package in R and an Egger’s test for funnel plot asymmetry.29,30 To integrate findings, we then compared the meta-analysis results with the harvest plots. In addition, we conducted meta-regressions to evaluate whether meta-analysis estimates were modified by continent, setting (ie, city-level connecting multiple neighbourhoods, neighbourhood-level, workplace, or university), study design, follow-up period, sample size, publication year, number of intervention components, intervention functions, quality assessment score, and outcome measure (eg, frequency).

Role of the funding source
The funders of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report.

Results
Following identification of records through database searching (n=50830) and manually searching the reference lists of included publications (n=82), we did title and abstract screening of 38916 non-duplicated records (figure 1). We assessed 710 full texts for inclusion eligibility and included 102 reports describing 121 interventions (appendix pp 11–16). Of these, we included 64 reports describing 67 interventions that reported outcomes with precision estimates. Further details on how functions and carrot and stick classifications were derived from intervention descriptions can be found in the appendix (pp 17–50). Some interventions were described by more than one report (appendix pp 35–40).

Table 2 provides a summary of the 121 interventions included in the systematic review. 79 interventions were categorised as carrot interventions (65%), with 22 stick (18%) and 20 carrot-and-stick (17%) interventions. 62 interventions were implemented in North America (51%) and 37 in Europe (31%). 69 interventions were evaluated using a repeat cross-sectional study design (57%). 60 of the included studies had a follow-up of greater than 12 months (50%). Carrot and carrot-and-stick interventions were more commonly reported in city (78%) and neighbourhood (85%) settings, whereas stick interventions were more commonly reported in workplace settings (n=14; 64%). On average, carrot-and-stick interventions had more intervention components (median 2·5, IQR 1·0–6·0) and functions (4·0, 3·8–5·0) than either carrot or stick interventions. Most interventions, particularly those with a carrot component, entailed functions involving access, safety, space, and convenience. Stick-type interventions were found to generally target financial access, safety, space, and convenience. Stick interventions were more commonly reported in workplace settings (n=14; 64%) compared with city (78%) and neighbourhood (85%) settings.

For the 67 interventions included in the meta-analysis, pooled point estimates from the forest plots were generally in the expected direction (eg, negative for driving and positive for active travel modes; figure 3). However, across all transportation outcomes and carrot or stick categories, nearly all effects were non-significant since confidence intervals crossed the null. As confidence intervals overlapped between carrot and stick intervention categories for most transport modes, effects might not differ statistically between categories, although point estimates for decreasing driving were greater for stick interventions and increasing active travel modes for combined carrot-and-stick interventions (figure 3). This pattern was consistent across reports assessing carrot and carrot-and-stick interventions. Reports assessing stick interventions had the highest proportions of both strong and weak ratings. Most studies had a weak or moderate score for selection bias (99%), study design (92%), blinding (74%), and data collection methods (76%).

Harvest plot results can be interpreted using the analogy of a seesaw, with the null hypothesis column representing the fulcrum of the seesaw (figure 2). In general, results were found to be weighted towards either end, indicating that interventions decreased driving and increased active travel. However, carrot interventions had less consistent results, as shown by the relatively higher proportion of interventions which found no change or an increase in driving and decrease in active travel modes. Moreover, there were few stick interventions (n=2) among studies reporting walking or cycling outcomes. Higher quality studies tended to report results in favour of the intervention, as evidenced by the distribution of bar heights. There was no obvious difference between multicomponent and single interventions, as shown by the annotated numbers above each bar.

For the 67 interventions included in the meta-analysis, pooled point estimates from the forest plots were generally in the expected direction (eg, negative for driving and positive for active travel modes; figure 3). However, across all transportation outcomes and carrot or stick categories, nearly all effects were non-significant since confidence intervals crossed the null. As confidence intervals overlapped between carrot and stick intervention categories for most transport modes, effects might not differ statistically between categories, although point estimates for decreasing driving were greater for stick interventions and increasing active travel modes for combined carrot-and-stick interventions (figure 3). This pattern was consistent across reports assessing carrot and carrot-and-stick interventions. Reports assessing stick interventions had the highest proportions of both strong and weak ratings. Most studies had a weak or moderate score for selection bias (99%), study design (92%), blinding (74%), and data collection methods (76%).

Harvest plot results can be interpreted using the analogy of a seesaw, with the null hypothesis column representing the fulcrum of the seesaw (figure 2). In general, results were found to be weighted towards either end, indicating that interventions decreased driving and increased active travel. However, carrot interventions had less consistent results, as shown by the relatively higher proportion of interventions which found no change or an increase in driving and decrease in active travel modes. Moreover, there were few stick interventions (n=2) among studies reporting walking or cycling outcomes. Higher quality studies tended to report results in favour of the intervention, as evidenced by the distribution of bar heights. There was no obvious difference between multicomponent and single interventions, as shown by the annotated numbers above each bar.

For the 67 interventions included in the meta-analysis, pooled point estimates from the forest plots were generally in the expected direction (eg, negative for driving and positive for active travel modes; figure 3). However, across all transportation outcomes and carrot or stick categories, nearly all effects were non-significant since confidence intervals crossed the null. As confidence intervals overlapped between carrot and stick intervention categories for most transport modes, effects might not differ statistically between categories, although point estimates for decreasing driving were greater for stick interventions and increasing active travel modes for combined carrot-and-stick interventions (figure 3). This pattern was consistent across reports assessing carrot and carrot-and-stick interventions. Reports assessing stick interventions had the highest proportions of both strong and weak ratings. Most studies had a weak or moderate score for selection bias (99%), study design (92%), blinding (74%), and data collection methods (76%).

Harvest plot results can be interpreted using the analogy of a seesaw, with the null hypothesis column representing the fulcrum of the seesaw (figure 2). In general, results were found to be weighted towards either end, indicating that interventions decreased driving and increased active travel. However, carrot interventions had less consistent results, as shown by the relatively higher proportion of interventions which found no change or an increase in driving and decrease in active travel modes. Moreover, there were few stick interventions (n=2) among studies reporting walking or cycling outcomes. Higher quality studies tended to report results in favour of the intervention, as evidenced by the distribution of bar heights. There was no obvious difference between multicomponent and single interventions, as shown by the annotated numbers above each bar.

Figure 3: Pooled standardised mean difference estimates for transport outcomes by carrot or stick category

| Category          | n  | Standardised mean difference (95% CI)         |
|-------------------|----|---------------------------------------------|
| Driving           |    |                                              |
| Carrot            | 16 | -0.10 (-0.23 to 0.03)                        |
| Stick             | 5  | -0.17 (-0.36 to 0.02)                        |
| Carrot and stick  | 5  | -0.13 (-0.47 to 0.20)                        |
| Public transport  |    |                                              |
| Carrot            | 29 | 0.11 (-0.10 to 0.32)                         |
| Stick             | 6  | 0.02 (-0.09 to 0.13)                         |
| Active travel     |    |                                              |
| Carrot            | 8  | 0.08 (-0.05 to 0.21)                         |
| Stick             | 3  | 0.01 (0.00 to 0.03)                          |
| Carrot and stick  | 8  | 0.33 (-0.01 to 0.68)                         |
| Walking           |    |                                              |
| Carrot            | 18 | 0.06 (-0.12 to 0.24)                         |
| Carrot and stick  | 7  | 0.41 (-0.01 to 0.82)                         |
| Cycling           |    |                                              |
| Carrot            | 38 | 0.37 (-0.07 to 0.81)                         |
| Carrot and stick  | 14 | 0.54 (-0.03 to 1.11)                         |

www.thelancet.com/planetary-health  Vol 6  November 2022  e864
pattern was particularly evident for active travel, with carrot-and-stick interventions having a greater pooled point estimate for active travel (0.33, 95% CI –0.01 to 0.68, \( P = 0.93\% \)), walking (0.41, –0.01 to 0.82, \( P = 100\% \)), and cycling (0.54, –0.03 to 1.11, \( P = 98\% \)) than interventions with either carrot or stick strategies, which had point estimates ranging from 0.01 to 0.37. For studies reporting on driving outcomes stick interventions had a greater decrease in driving if they were made use of stick strategies (0.11, 95% CI –0.24 to 0.02, \( P = 0.20\% \)). For public transport, interventions with carrot strategies had a higher point estimate (0.11, 95% CI –0.39 to –0.08, \( P = 92\% \)) compared with interventions with either carrot or stick strategies, which had point estimates ranging from 0.01 to 0.11. For driving, stick intervention results clustered in the harvest plot column supporting the hypothesis of a decrease (a beneficial effect) and were congruently associated with larger negative (beneficial) pooled point estimates in the forest plot. For walking, cycling, and active travel, carrot-and-stick intervention results clustered in the harvest plot column in support of the hypothesis of an increase (a beneficial effect) and were congruently associated with larger positive (beneficial) pooled point estimates in the forest plot.

Based on visual inspection of the harvest plot and the pattern of results in the meta-analysis, findings from the two methods are consistent with, but not proof of, an overall beneficial effect of interventions on both driving and active travel modes. For carrot interventions, the variability of results in the harvest plot is supported by evidence of smaller point estimates in the forest plot. For driving, stick intervention results clustered in the harvest plot column supporting the hypothesis of a decrease (a beneficial effect) and were congruently associated with larger negative (beneficial) pooled point estimates in the forest plot. For walking, cycling, and active travel, carrot-and-stick intervention results clustered in the harvest plot column in support of the hypothesis of an increase (a beneficial effect) and were congruently associated with larger positive (beneficial) pooled point estimates in the forest plot.

Figure 4 presents the results of the pooled effect sizes on each mode of transport by function. For driving, interventions that altered financial functions had a significant negative (beneficial) effect (SMD –0.23, 95% CI –0.39 to –0.08, \( P = 92\% \)), whereas for active travel, walking, and cycling, interventions that targeted access, safety, and space functions generally had significant positive effects. Additional functions which had a significant positive effect on cycling included aesthetics (0.10, 95% CI 0.01 to 0.29, \( P = 97\% \)) and financial (0.35, 0.03 to 0.68, \( P = 89\% \)). None of the functions were significantly associated with public transport outcomes.

Most meta-regressions revealed no substantial or significant variations in effect estimates for each transportation outcome (appendix pp 63–67). There were some exceptions; for instance, carrot interventions in Asia were associated with an increase in driving compared with those in Europe. Furthermore, for every 1-month increase in follow-up length for carrot interventions, the point estimates for public transport increased by 0.02 (95% CI 0.00–0.04). For stick interventions, studies published more recently (ie, for every 1 year increase in publication date) were associated with a more positive point estimate (0.08, 95% CI 0.01–0.15) than older studies. Combined carrot-and-stick interventions were associated with a greater decrease in driving if they were done in the workplace rather than at a city-wide level (–0.61, 95% CI –0.78 to –0.43) or had a longer follow-up period (–0.02, –0.03 to 0.00). Moreover, additional function was associated with an increase in cycling outcomes for combined carrot-and-stick interventions (1.57, 95% CI 0.63–2.51).

Discussion

This systematic review and meta-analysis provides empirical evidence on the effectiveness of groups of interventions seeking to either discourage driving or promote active travel. Results from the harvest plots and meta-analysis were generally consistent: although

| Function          | n  | Standardised mean difference (95% CI) |
|-------------------|----|-------------------------------------|
| Driving           |    |                                     |
| Access            | 18 | –0.11 (–0.24 to 0.02)               |
| Financial         | 9  | –0.23 (–0.39 to 0.08)               |
| Space             | 8  | –0.18 (–0.36 to 0.00)               |
| Convenience       | 15 | –0.11 (–0.26 to 0.05)               |
| Public transport  |    |                                     |
| Aesthetics        | 9  | 0.14 (–0.02 to 0.29)                |
| Access            | 18 | 0.07 (–0.27 to 0.42)                |
| Awareness         | 6  | 0.12 (–0.13 to 0.37)                |
| Convenience       | 22 | 0.10 (–0.19 to 0.38)                |
| Financial         | 8  | 0.05 (–0.03 to 0.13)                |
| Safety            | 2  | 0.06 (–1.35 to 1.27)                |
| Space             | 4  | 0.05 (–0.19 to 0.29)                |
| Active travel     |    |                                     |
| Aesthetics        | 9  | 0.27 (–0.07 to 0.62)                |
| Access            | 16 | 0.20 (0.02 to 0.38)                 |
| Awareness         | 5  | 0.11 (–0.07 to 0.29)                |
| Convenience       | 11 | 0.27 (0.02 to 0.52)                 |
| Financial         | 2  | 0.05 (–0.19 to 0.29)                |
| Safety            | 15 | 0.24 (0.01 to 0.46)                 |
| Space             | 15 | 0.21 (0.02 to 0.41)                 |
| Walking           |    |                                     |
| Aesthetics        | 7  | 0.28 (–0.08 to 0.65)                |
| Access            | 22 | 0.15 (–0.05 to 0.34)                |
| Awareness         | 4  | 0.27 (–0.09 to 0.61)                |
| Convenience       | 18 | 0.21 (–0.02 to 0.44)                |
| Safety            | 15 | 0.23 (0.02 to 0.44)                 |
| Space             | 15 | 0.23 (0.02 to 0.44)                 |
| Cycling           |    |                                     |
| Aesthetics        | 20 | 0.41 (0.01 to 0.81)                 |
| Access            | 46 | 0.48 (0.09 to 0.87)                 |
| Awareness         | 8  | 0.91 (–0.09 to 1.96)                |
| Convenience       | 30 | 0.25 (–0.03 to 0.53)                |
| Financial         | 5  | 0.35 (0.03 to 0.68)                 |
| Safety            | 33 | 0.54 (0.00 to 1.07)                 |
| Space             | 34 | 0.56 (0.04 to 1.08)                 |
| Skills            | 4  | 0.52 (–0.16 to 1.22)                |
interventions understood to lead to gains in functions (carrots) were the most widely reported in the literature, they tended to report less consistent findings and smaller point estimates than those understood to lead to losses (sticks) or both (carrots-and-sticks). For driving, stick interventions (eg, workplace parking restrictions to discourage driving) and carrot-and-stick interventions for active travel (eg, creating cycle lane infrastructure at the expense of road space for motor vehicles) were less widely reported than carrot interventions but had some of the highest pooled effect sizes. However, caution should be exercised when interpreting these results, as there were few stick or carrot-and-stick interventions and most estimates were nonsignificant. Together, these findings suggest that interventions with carrot strategies alone might not be as effective as interventions that include stick strategies, though further evidence on interventions using stick strategies is necessary for more generalisable conclusions.

Meta-analyses based on how the intervention functions sought to encourage or discourage shifts in travel identified different functions, which were associated with more effective interventions. Interventions seeking to alter behaviour through financial means might be more effective for changing driving behaviour, whereas those that improve access, safety, and space might help to promote active travel. However, these findings could also be driven by them being the most commonly targeted functions for these transport outcomes.

Consistent with our findings, narrative reviews have suggested that comprehensive policy packages might be more effective than single interventions. Our finding that combined carrot-and-stick interventions had larger and more consistent point estimates than carrot interventions largely agrees with research on a range of other behavioural interventions, including smoking cessation programmes and workplace travel plans, which generally found that stick interventions (either alone or in combination with carrot strategies) were more effective than carrot interventions alone. This difference might be because stick interventions often depend on lower levels of individual agency for their effectiveness, or because people are generally more motivated to avoid losses than to pursue gains; it could also be due to carrot-and-stick interventions generally having a greater number of functions, which was found to be associated with greater effects on cycling outcomes in our sensitivity analysis.

Our finding that stick and combined carrot-and-stick interventions were much less extensively evaluated in the literature could be due to a variety of reasons, including that these types of interventions might be more challenging to implement, suffer from low rates of uptake, or are less politically feasible due to more pushback from various stakeholders (eg, motorists and the automotive industry). There is, however, a growing variety of interventions implemented with stick strategies, either alone (eg, licence plate restriction policies) or combined with carrot interventions (eg, low traffic neighbourhoods), which could provide further opportunities for natural experimental evaluation. To encourage political and social support of interventions with stick strategies, these perceived negative policies can be framed as beneficial by highlighting potential health gains or be combined with carrot interventions. Furthermore, although policies with a stick strategy are often met with resistance, public acceptability can change after they are implemented.

Reviews on this topic have typically examined interventions based on their surface form (eg, new cycling infrastructure, promotional activities, or reduced speed limits), frequently finding the available evidence to be too diverse, inconsistent, and flawed to make a conclusive statement about the effectiveness of specific interventions. A few reviews, however, have started to explore the value of analysing interventions in this field by function, such as those thought to affect behaviour through financial means or environmental changes affecting access, convenience, and safety. This Article builds upon these earlier efforts by assessing a greater variety of interventions and their corresponding functions.

By grouping interventions in terms of their common core function regardless of the form (ie, specific components or strategies) they might present as, we developed a preliminary approach to quantitatively compare the effectiveness of interventions and produce generalisable findings despite a diversity of intervention forms, study designs, and contexts. This approach embodies the dry stone wall metaphor proposed by Ogilvie and colleagues, which encourages flexibility when combining dissimilar stones in making sense of diverse or complex sets of evidence. Although we found this approach to be applicable to many diverse transport interventions, we acknowledge challenges in classifying diverse interventions into functional or carrot-or-stick categories. Our reliance on the review team’s interpretation of study authors’ descriptions within reports could have led to some misclassification. Intervention function is too complex to be entirely captured in an Article of this kind, as it includes intended and perceived impacts, unintended implications, and heterogeneity of effects across populations. Hence, there is scope for further development of classifications with various stakeholders, including quantitative and qualitative analysis. Improved reporting of mechanisms or theories underpinning their interventions would enable further development of function-oriented analysis.

Studies included in this analysis represent a global evidence base; however, there could be some limitations to the generalisability of the findings to different contexts. For instance, we found some differences in intervention effectiveness for certain characteristics (eg, Europe vs Asia).
in the sensitivity analysis, and we did not find many studies from low-income and middle-income countries such as those in Africa and South America.

In addition, we did not assess differences in intervention effectiveness for different population groups (eg, by sex, age, and socioeconomic position) because too few studies reported this information. However, this paucity of evidence suggests that future research should focus on how certain groups could benefit from different approaches or intervention functions. For instance, some policies such as congestion charging are just starting to be evaluated in terms of their effects on children’s lung development and long-term respiratory health by discouraging driving and decreasing associated vehicular emissions.39

As studies without control groups were excluded, some interventions, particularly stick interventions, might have been overlooked.40 However, by including only controlled studies with data from before and after the intervention, we have removed a potential source of bias, increased comparability between studies, and strengthened the basis for causal inference. Our systematic review and meta-analysis was also limited by the high level of heterogeneity between studies, which was only partly explained by the meta-regression analysis and the low quality of evidence.

Although carrot and stick terminology has been used to describe policies in previous studies, little is known about the use or generalisability of inference about intervention effectiveness based on the concept of

| Panel 2: Real world examples of intervention functions |
|------------------------------------------------------|
| The interventions below are examples that show how some of the functions described in this Article have been implemented in practice. |

**Example 1: Carrot intervention with access and convenience functions**

International Bikeshare Impacts on Cycling and Collisions41

Intervention description: public bike-share programmes were introduced in major North American cities across both the USA and Canada.

**Functions affected:**

- **Access (gain):** the intervention aimed to increase accessibility for cyclists by providing access to public bikes to rent and use
- **Convenience (gain):** the intervention aimed to increase convenience for cyclists by reducing steps and distance needed to rent a bike

Outcome: there was about a two-fold greater likelihood (odds ratio 1.84; 95% CI 1.00–3.39) of past-week cycling for individuals living near a new bike-share station in cities that implemented a bike-share programme compared with those that did not.

**Example 2: Stick intervention with a financial function**

Gothenburg Congestion Charge41

Intervention description: a time-of-day cordon-based congestion charge system, in which a fee was introduced during peak hours, applied to the entire Gothenburg, Sweden city centre and a connecting highway.

**Functions affected:**

- **Financial (loss):** the intervention aimed to change travel behaviour by making it more expensive to drive in certain areas within the city

Outcome: car commuting decreased and public transport use increased in the group of subjects who were exposed to the intervention, with no significant travel behaviour changes in the comparison group.

**Example 3: Carrot-and-stick intervention with aesthetics, access, convenience, safety, and space functions**

Eixo Central project44

Intervention description: two streets and one pedestrian plaza were retrofitted in Lisbon, Portugal, which involved constructing new cycling infrastructure (protected, physically separated from traffic, and bi-directional cycling lanes along with painted cycle lanes), widening pavements, creating new crosswalks, installing benches, lighting, and greenery and planting trees. To create space for pedestrian and cyclist amenities, traffic lanes and parking spaces were removed.

**Functions affected:**

- **Aesthetics (gain):** the intervention aimed to create a pleasant environment for pedestrians and cyclists by increasing the amount of green space and areas to sit and relax
- **Access (gain and loss):** the intervention aimed to increase accessibility for pedestrians and cyclists by increasing the amount of areas dedicated to walking and cycling; accessibility for driving was decreased via removal of street space accessible to cars
- **Convenience (gain and loss):** the intervention aimed to increase connectivity and thus convenience between walking and cycling infrastructure by increasing the length of dedicated pedestrian and cycling infrastructure; convenience was decreased for motorists by removing opportunities to park or drive on certain streets, thereby potentially increasing the time taken to find parking or reach destinations
- **Safety (gain):** the intervention aimed to increase pedestrian and cyclist safety by installing dedicated, protected spaces to walk and cycle separate from traffic
- **Space (gain and loss):** the intervention aimed to increase the space dedicated to pedestrians and cyclists by constructing dedicated cycling lanes and widening pavements; space was decreased for driving by removing traffic lanes and parking spaces

Outcome: significant increases in pedestrian volumes were observed on the streets that received the intervention, particularly on streets where larger-scale built-environment changes occurred, whereas there was no change on the comparison street.
function. However, conceptualising interventions in this way could help to inform decision makers about the general principles that might reasonably be targeted (eg, safety), even if there are multiple ways of achieving them (eg, reducing speed limits or creating dedicated cycling paths). The principle is already implicitly made use of in public health guidance: in offering recommendations on how to improve the physical environment to promote physical activity, the National Institute for Health and Care Excellence has sets of specific options from which options might be selected that are appropriate to each local context.\(^{[10]}\) Given the versatility of using function to characterise interventions, the approach might also be applicable to synthesising evidence of intervention effectiveness in other topics.

To further illustrate how functions might be used in practice, they can be envisaged as common underlying mechanisms that are triggered when various distinct policy levers, representing either carrot or stick interventions, are pulled. Policy levers can be pulled singly or in combination; some might be pulled at the expense of others. For example, fees collected from congestion charging might be used to fund public transport investment, or space might be taken from motor vehicles to widen pavements. Deciding which levers to pull could also depend on economic, sociopolitical, or physical constraints. Moreover, equity should be taken into consideration when designing interventions to ensure that they do not unfairly disadvantage population groups who are already socioeconomically disadvantaged or less able to be mobile (eg, people with disabilities). Altogether, this framework could allow decision makers to adapt the principles of effective intervention to their particular aims and contexts while encouraging creativity when designing new combinations of interventions.

Our findings suggest that interventions should include combined carrot-and-stick strategies, or a mix of carrot and stick policy levers. This Article also identified that policies triggering specific functions were effective at either decreasing driving (eg, financial) or increasing active travel (eg, access, safety, and space). A few examples of interventions involving these functions are presented (panel 2).\(^{[8]}\)

In general, evidence suggested that although stick and combined carrot-and-stick interventions had been less commonly implemented or evaluated, on average they had more consistently positive evidence of effectiveness, with greater point estimates than those for carrot interventions. In addition, different intervention functions were identified as associated with a greater likelihood of effectiveness, such as changing costs for driving and changing access and safety for active travel. These findings can guide decision makers in choosing and designing policies to better address challenges related to health and the environment.

**Contributors**

All authors contributed to the design of the review protocol. CX and RP conducted the initial screening, selected studies for inclusion, and extracted and checked the data from the included studies. CX, JP, and EvS performed the study quality assessment. All steps from screening to quality assessment were done in consultation with the wider review team. CX analysed the data and drafted the manuscript. JP is the guarantor. All authors contributed to the critical revision and approved the final version of the manuscript. All authors had access to all the data in the study and had final responsibility for the decision to submit for publication. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

**Declaration of interests**

We declare no competing interests.

**Data sharing**

All datasets generated and analysed, including the study protocol, search strategy, list of the included and excluded studies, data extracted, analysis plans, quality assessment, and assessment of the publication bias, are available in the Article and upon request from the corresponding author.

**Acknowledgments**

CX is funded by the Cambridge Trust. EvS, DO, RP, and JP are supported by the Medical Research Council (Unit Programme number MC_UU_00006/7, MC_UU_12015/6 and MC_UU_12015/7).

**References**

1. Sahlyapit S, Goodman A, Cooper AR, Ogilvie D. Change in active travel and changes in recreational and total physical activity in adults: longitudinal findings from the iConnect study. *Int J Behav Nutr Phys Act* 2013; 10: 28.
2. Celia-Morales CA, Lyall DM, Welsh P, et al. Association between active commuting and incident cardiovascular disease, cancer, and mortality: prospective cohort study. *BMJ* 2017; 357: j456.
3. Haines A. Health co-benefits of climate action. *Lancet Planet Health* 2017; 1:e4–5.
4. Brand C, Götsch T, Dons E, et al. The climate change mitigation impacts of active travel: Evidence from a longitudinal panel study in seven European cities. *Glob Environ Change* 2021; 67: 102224.
5. Woodcock J, Edwards P, Tonne C, et al. Public health benefits of strategies to reduce greenhouse-gas emissions: urban land transport. *Lancet* 2009; 374: 1930–43.
6. UK Department for Transport. £75 million more for cycling and walking as research shows public support. 2020. https://www.gov.uk/government/news/75-million-more-for-cycling-and-walking-as-research-shows-public-support (accessed June 1, 2021).
7. Marmot M, Allen J, Goldblatt P, Herd E, Morrison J. Build back fairer: the COVID-19 Marmot Review. London: The Pandemic, Socioeconomic and Health Inequalities in England, 2020.
8. Ogilvie D, Egan M, Hamilton V, Petticrew M. Promoting walking and cycling as an alternative to using cars: systematic review. *BMJ* 2006; 332: 763–66.
9. Stewart G, Anokye NK, Pokhrel S. What interventions increase commuter cycling? A systematic review. *BMJ Open* 2015; 5: e007945.
10. Ogilvie D, Adams J, Bauman A, et al. Using natural experimental studies to guide public health action: turning the evidence-based medicine paradigm on its head. *J Epidemiol Community Health* 2020; 74: 203–08.
11. Piatkowski DP, Marshall WE, Krizek RJ. Carrots versus sticks: assessing intervention effectiveness and implementation challenges for active transport. *J Plann Educ Res* 2019; 39: 50–64.
12. Petrunoff N, Rissel C, Wen LM, Martin J. Carrots and sticks vs carrots: comparing approaches to workplace travel plans using disincentives for driving and incentives for active travel. *J Transp Health* 2015; 2: 563–67.
13. Podczervinski S, Siedrich Z, Helbert L, et al. Employee influenza vaccination in a large cancer center with high baseline compliance rates: comparison of carrot versus stick approaches. *Am J Infect Control* 2015; 43: 228–33.
14. Halpern SD, French B, Small DS, et al. Randomized trial of four financial-incentive programs for smoking cessation. *N Engl J Med* 2015; 372: 2108–17.
22 Page MJ, McKenzie JE, Bossszyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021; 372: n71.

23 Scheepers CE, Wendel-Vos GCW, den Broeder JM, van Kempen EEM, van Wesemael PJV, Schuit AJ. Shifting from car to active transport: a systematic review of the effectiveness of interventions. Transp Res Part A Policy Pract 2014; 70: 264–80.

24 Rissel C, Curac N, Greensaway M, Bauman A. Physical activity associated with public transport use—a review and modelling of potential benefits. Int J Environ Res Public Health 2012; 9: 2454–78.

25 Effective Public Health Panaeza Project. Quality assessment tool for quantitative studies. 2022. https://www.ephipp.ca/quality-assessment-tool-for-quantitative-studies/ (accessed July 29, 2022).

26 Ogilvie D, Fayter D, Petticrew M, et al. The harvest plot: a method for synthesising evidence about the differential effects of interventions. BMC Med Res Methodol 2008; 8: 8.

27 Lüdecke D, esc: Effect size computation for meta analysis. 2018. https://zenodo.org/record/1249218#.YyM8LaTMKUk (accessed Jan 26, 2022).

28 Balduzzi S, Rücker G, Schwarzer G. How to perform a meta-analysis with R: a practical tutorial. Evid Based Ment Health 2019; 22: 153–60.

29 Sterne JAC, Sutton AJ, Ioannidis JPA, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ 2011; 343: d4002.