Transport-Health Equity Outcomes from ICT enabled incentives schemes – the Newcastle case study

Susan Grant-Muller[1], Gillian Harrison[1], Frances Hodgson[1] , Nicholas Malleson and Tom Redfern[2] , Robert Snowball[3]

[1]Institute for Transport Studies, University of Leeds, Leeds, LS2 9JT, UK
[2]School of Geography, University of Leeds, Leeds, LS2 9JT, UK
[3]Newcastle City Council

Abstract

Improved understanding of the spatial distribution of health impacts arising from the introduction of new travel initiatives will support more targeted and efficient policy development across both the transport and health sectors. Typical health impacts include those arising from changes in levels of personal activity with alternative mode choices. With a sectoral approach to policy development, positive impacts for one sector (i.e. improved transport services) may be negated by dis-benefits in another (e.g. low levels of active travel choice and increased obesity related disease burden). The goal of this research is to extend beyond an improved understanding of the overall consequential health impacts (positive and negative) for the travelling public, however, and to identify the distribution of impacts. To research this, the horizontal notion of equity (Thomopoulos, Grant-Muller and Tight, 2009) is applied using a range of transport-related health outcomes including cancer, heart disease and depression. The aim is to identify those segments of the population (or city sub-areas) where the travelling public are particularly advantaged or disadvantaged in terms of health related outcomes from new transport schemes. The research methodology interfaces new generation 'Track and Trace' information on individuals location and mode choices (detected as mobile phone app-based sensor data) with a new integrated transport and health model (IHITM), finally calculating an equity indicator based on distributional impacts.

The work presented here represents on-going research to harness the potential of new generation data ('Big Data') arising from innovative policies (based on persuasive technologies) in the transport sector. The case study data arises from the large scale implementation of positive incentives to encourage changes to travel choices using mobile phone apps in Newcastle (UK), which is part of a large project involving related studies in 10 other cities across Europe.

Methodology

A better understanding of the interfaces between travel choices and the chain of consequential impacts is needed to improve liveability in cities and communities, in Europe and globally. In particular, interfaces relating to health and particularly the disease burden from exposure to pollutants, obesity consequences from inactivity. Globally, exposure to ambient air pollution was estimated to have caused 3.7 million deaths in 2012 (WHO, 2014).

A number of models go some way to capturing the links between the transport sector and other sectoral impacts, for example the ITHIM model (Woodcock et al, 2013) to explore transport-health interrelations. These cross-sectoral models have, however, been largely developed based on traditional (and mainly aggregate) data sources such as travel diaries, fixed based traffic counts etc. This type of modelling is based on a 'snapshot' of the transport system at fixed-base locations. The models are prone to a number of weaknesses, such as over-reporting in the input data (PHE, 2014) and difficulties in identifying distributional impacts as a result of the granularity of the inputs. As a result, the outputs are necessarily at a high spatial level of aggregation. The Health Effects Institute recommends models that account for personal exposure or include time-activity data can produce the "best" estimates of human exposure (Health Effects Institute, 2010). The ability to understand the health equity consequences at a much more refined level than city-wide or at a fixed base location in the city is an important next step in policy development. Analysis to reveal whether particular residential locations or corridors in the city have a greater adverse health outcome from a transport initiative than other locations is key information in order to direct mitigating resources in a more targeted way. In comparing the expected benefits between alternative transport schemes ex-ante, an equity indictor of the health consequences will give a further dimension to decision making according to policies concerning transport, health and an equitable society.

The emergence of a next generation of transport and other data with extended use of ICT not just in the transport sector, but across society as a whole, has given rise to 'unprecedented volumes of data across all modes and transport systems' (EC 2016). The volume is not the only key characteristic however. It has a granularity, immediacy and geo-location aspects that are fundamentally of a different order to traditional data streams. PHE (2014) identify technology as a key strategic area to improve the accuracy in reporting activity and travel effectively without huge investment costs or imposition. The research presented here harnesses the opportunity provided by mobile phone sensor-generated location data to collect the day-to-day movement patterns of individuals (with their informed consent), including the mode, route and other information. This is referred to as Track and Trace (T&T) data. The technology to capture and collect location data operates in different ways, some requiring the download of a software app offering specific location based services provided by stakeholders such as city authorities, transport suppliers or third parties. The characteristics of a dataset generated depends on a number of factors such as: the branding and key features of the app, the subset of the population who engage with the smartphone app, the subset that actively retain the app over time, the subset who take part in additional voluntary data collection, such as short period experience sampling, voluntarily offering personal data such as age and gender and voluntarily using features such as correction of mode where the app has mis-recorded. Correction factors can be applied in some cases to reflect known biases in the data. The quality and representativeness of this type of data is, however, a subject matter for another paper (forthcoming).

A mobility profile of individuals' choices before and after introduction of a transport related initiative can be constructed based on the location data. This typically shows the origin, destination, route choice and mode choice over time (Figure 1). From the data, the individuals home location is either implicit or is sometimes accurately volunteered.

Figure 1: Spatial patterns in T&T trip data

The geographic sub-area that an individuals' home location lies within has (within the UK) an assigned multiple indicator of social and economic welfare, as reflected by the Index of Multiple

Deprivation, or IMD score (Gov.UK, 2015). The IMD score is a relative measure of deprivation for small areas (Lower-layer Super Output Areas) across England, based on seven different domains of deprivation: Income Deprivation, Employment Deprivation, Education, Skills and Training Deprivation, Health Deprivation and Disability, Crime, Barriers to Housing and Services, and Living Environment Deprivation. Whilst the social welfare of any individual isn't necessarily related to or reflected by the IMD score of the home location, the IMD score is reflective of general social conditions in that area and will approximate the conditions for many.

One of the potentially useful features of T&T data is therefore the ability to detect journey patterns for individuals living in sub areas with particular IMD scores and how these are impacted (positively or negatively) by new transport initiatives. From a policy perspective, this allows the possibility of either 1) assessing the impacts of targeted transport initiatives intended to improve welfare in deprived areas or 2) detecting unintended consequences of generalised transport initiatives on low IMD areas.

Similarly, health statistics (mortality, disease adjusted life years) is available at national and regional level in many countries. When T&T data is generated and interfaced with health statistics at sub-region level (where available), this allows a more nuanced understanding of how travel choices (and changes in travel choices) can improve individuals health outcomes.

In summary, T&T data (when collected in critical mass) holds the potential for a cumulative micro-level analysis of how travel choices are influenced by schemes and what the distributional (equity) outcomes of particular schemes are in relation to health and social welfare. This can be a far more complete analysis of transport choice (for a particular individual) than is currently possible, without significant survey resource or burden of requirement on the public.

<u>Case study for Newcastle City Region, UK</u>

The case study presented here is based on T&T data generated from a large scale implementation involving members of the public in the Newcastle City region (UK). The dataset covers an implementation period from July 2017 onwards, relating to circa 2000 individuals, who downloaded a smartphone app (EMPOWER project, 2015). The nature of the app was a broad spectrum, information and incentivisation source relating to city activities, travel news, events etc. The broad app function was attractive to a wide range of the public, who were, in return, asked to voluntarily record their age and gender. Age and gender were particularly pertinent to the calculation of micro-changes in health burden outcomes from changes in travel choice to more active modes. Whilst a number of other apps are available, the majority of these have a specific focus (such as fitness apps, cycling apps etc) and as a result have a propensity to attract very specific subsets of the population. For the purpose of this type of analysis, individuals' data arising from use of those apps is much less useful and more highly biased.

An example of some of the variables collected by the app is shown in Figure 2 below. This is an extremely rich datasource with more than 40 variables generated for each trip.

Unique trip ID	Start hour	Start day	Start week	End hour	Trip quality, Missing indicates an inserted trip	Average speed	Total trip distance	Trip modality, possibly corrected by the traveller	Total travelling cost, estimate	Total calories spend during trip, estimate	Emitted particle	Emitted amount in grams
1706027132	16	27	35	17	Good	1.3	1600	Foot	0.02	102	CO2	39
1706027132	17	27	35	18	Unknown	0.4	430				CO2	0
1706027132	18	27	35	18	Good	1.2	1700	Foot	0.02	111	CO2	44
1706027132	18	27	35	9	Missing	0	0				CO2	0
1706027132	9	28	36	9	Bad	0.3	40				CO2	0

Unique trip ID	Weather temperature in degrees Celsius	Amount of rain in mm	Weather condition	Latitude in WGS84	Longitude in WGS84	Accuracy in meters	Place name of origin	Place type	Postal code	PC4	Total stay duration in origin	Total stay percentage in origin	Number of visits to this place
1706027132	19	0	clear	-21.537	15.09316	10	The grove	home	LS2	BD21	1228445	78.1	28
1706027132	0	0		-11.544	15.084416	10	City	place	LS2	BD21	13796	0.9	2
1706027132	17	0	clear	-11.544	15.084416	10	City	place	LS2	BD21	13796	0.9	2
1706027132	0	0		-21.537	15.09316	10	The grove	home	LS2	BD21	1228445	78.1	28
1706027132	0	0		-21.537	15.09316	10	The grove	home	LS2	BD21	1228445	78.1	28

Figure 2: Trip data generated by smartphone application (anonymised here)

Whilst not all individuals were willing to declare their age and gender, those that did revealed a broad age distribution (including more senior ages) and a slight bias towards females. The gender bias may have arisen due to a number of factors – interest in the app, willingness to declare their gender, commitment to retain the app, as examples. However it is noteworthy that similar findings emerged in the data of another large city in Italy (EMPOWER project, 2018) where all individuals declared their age and gender on registration ie compulsory declaration.

In order to calculate micro changes in health burden outcomes arising from mode switch to active travel, a new transport-health model was derived, based on the macro level IHITM model. The new HABITS model interfaces the T&T data at individual level with the underlying transport-health models, enabling more disaggregate level outputs. It outputs the percentage change in all-cause mortality and disability adjusted life years (DALYs) following transport scheme implementation and based on age, gender and MET-hours. These calculations are made on a weekly trip-basis, aggregated from individual trips in order to accommodate a threshold model for the health outcomes. The T&T data also provides full O-D data, a range of data on externalities, habitual patterns, dwell time, phone type, confidence/quality indicators and more.

Based on inferred home location, the average change in mortality and DALYs for different IMD areas can be calculated, by aggregating the set of individuals' outcomes. This allows an equity analysis and understanding of whether (and to what extent) particularly deprived or affluent areas have been impacted by behavioural change policies.

In order to generate an equity index for those transport-health outcomes, we have propose dthe following statistic, adapted from Thomopoulos and Grant-Muller (2013).

Each of the q health-indicator values (ie for each IMD category) are normalised using a re-scaling method (OECD-JRC, 2008). Each indicator x^{t}_{qc} for a given sub-area of the City c, and time t, is transformed:

$$I^{t}_{qc} = \frac{x^{t}_{qc} - \min_{c}(x^{t}_{q})}{\max_{c}(x^{t}_{q}) - \min_{c}(x^{t}_{q})} \quad \ldots\ldots\ldots.. \quad (1)$$

Where minc (x^{t}_{q}) = minimum value of x^{t}_{qc} across all sub-areas c at time t

$\max_c(x^t_q)$ = maximum value of x^t_{qc} across all sub-areas c at time t

In this way the normalised indicators I^t_{qc} have values between 0 and 1. Alongside the standardised sub-region values for each indicator, a composite transport-health impact equity indicator (THEQ indicator) is given by:

$$THEQ = \sum w_q I^t_{qc} \qquad (2)$$

where w_i is the weight for the ith health indicator and can be calculated according to a number of approaches. Whilst the calculation of an overall indicator loses some of the important granular information, this reporting both disaggregate and aggregate equity outcomes is analogous to the calculation of a single transport scheme BCR value, whilst also reporting granular cost/benefit category outcomes.

In the analysis presented, we describe empirical findings relating to transport-health outcomes (and the equity implications of those outcomes) for sub-regions of Newcastle, based on empirical T&T data collected from July 2017 onwards and involving circa 2000 individuals. Newcastle is a city region with wide variation in life expectancy and health outcomes between different super output areas. Life expectancy is 12.4 years lower (men) and 9.5 years lower (women) in the most deprived areas than in least deprived of Newcastle. There are additional inequalities related to incidence of alcohol, smoking, obesity and self-harm. During the period since July 2017, the city council introduced several behavioural measures (i.e. individual incentivisation schemes) to encourage mode switch from the use of conventionally fuelled vehicles to more active modes. An analysis based on a ‘before and after’ approach will be used to demonstrate the potential and challenges in using T&T data for this purpose and the equity related insights that arose.