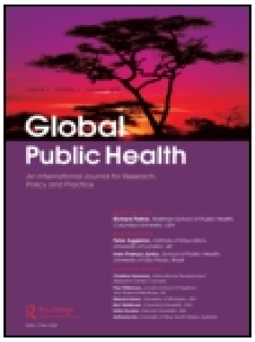

Global Public Health
An International Journal for Research, Policy and Practice

ISSN: 1744-1692 (Print) 1744-1706 (Online) Journal homepage: www.tandfonline.com/journals/rgph20

Large-scale road safety programmes in low- and middle-income countries: An opportunity to generate evidence

Adnan A. Hyder, Katharine A. Allen, David H. Peters, Aruna Chandran & David Bishai

Global Public Health, 2013
Vol. 8, No. 5, 504–518, http://dx.doi.org/10.1080/17441692.2013.769613

Large-scale road safety programmes in low- and middle-income countries: An opportunity to generate evidence

Adnan A. Hyder[a,b]*, Katharine A. Allen[a,b], David H. Peters[b], Aruna Chandran[a,b] and David Bishai[a,c]

[a]*Johns Hopkins International Injury Research Unit, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA;* [b]*Department of International Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA;* [c]*Department of Population Family and Reproductive Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA*

(*Received 29 July 2012; final version received 3 December 2012*)

The growing burden of road traffic injuries, which kill over 1.2 million people yearly, falls mostly on low- and middle-income countries (LMICs). Despite this, evidence generation on the effectiveness of road safety interventions in LMIC settings remains scarce. This paper explores a scientific approach for evaluating road safety programmes in LMICs and introduces such a road safety multi-country initiative, the Road Safety in 10 Countries Project (RS-10). By building on existing evaluation frameworks, we develop a scientific approach for evaluating large-scale road safety programmes in LMIC settings. This also draws on '13 lessons' of large-scale programme evaluation: defining the evaluation scope; selecting study sites; maintaining objectivity; developing an impact model; utilising multiple data sources; using multiple analytic techniques; maximising external validity; ensuring an appropriate time frame; the importance of flexibility and a stepwise approach; continuous monitoring; providing feedback to implementers, policy-makers; promoting the uptake of evaluation results; and understanding evaluation costs. The use of relatively new approaches for evaluation of real-world programmes allows for the production of relevant knowledge. The RS-10 project affords an important opportunity to scientifically test these approaches for a *real-world, large-scale road safety evaluation* and generate new knowledge for the field of road safety.

Keywords: road safety; injury prevention; road traffic injuries; developing countries; evaluation

Introduction

More than 1.2 million people die every year in road traffic crashes around the world and an additional 20–50 million are injured or disabled (Peden, World Health Organization, & World Bank, 2004). This burden falls mostly on low- and middle-income countries (LMICs) where the rates of road traffic deaths are twice as high as those in developed countries (World Health Organization Department of Violence and Injury Prevention and Disability, 2009). Ninety per cent of road traffic deaths occur in LMICs even though they account for $<50\%$ of the world's registered automobiles (World Health Organization Department of Violence and Injury

*Corresponding author. Email: ahyder@jhsph.edu

Prevention and Disability, 2009). Between 2004 and 2020, these rates are predicted to increase by 27% in LMICs while decreasing in high-income countries by 83%, augmenting the global gap (World Health Organization Department of Violence and Injury Prevention and Disability, 2009). This burden results in significant economic losses; road traffic injuries (RTIs) have been estimated to cost LMICs an estimated US$100 billion every year. These numbers are based on calculations that are already several years old and, therefore, are likely underestimates, but even this estimate is almost twice what these countries receive in total development assistance on an annual basis, or equivalent to 1–3% of their gross domestic product, which is also comparable to total health spending in many developing countries (World Bank, 2002; World Health Organization, Etienne, Asamoa-Baah, & Evans, 2010).

Despite the burden of RTIs on LMICs, the generation of evidence from these countries has not kept pace. A recent Cochrane Injury Group analysis of reviews on road safety interventions shows that of the 236 trials utilised in the group's 13 systematic reviews on road safety interventions, only 6 were conducted in LMICs (Perel, Ker, Ivers, & Blackhall, 2007). Moreover, while the 236 trials comprised a wide range of road safety interventions, all 6 LMIC-based studies focused on helmet-wearing among motorcyclists, illustrating the limited nature of research on RTIs in such countries (Perel et al., 2007). Another recent study highlighted the high-income country dominance of RTI research, whereby Australia, Europe and the USA produced, on average, over 82 RTI-related articles per 1000 road traffic-related deaths (Borse & Hyder, 2009). In comparison, rates for the Middle East, South America and Sub-Saharan Africa were between two and five, and in North Africa, India and China, there was less than one article on RTIs per 1000 road traffic-related deaths (Borse & Hyder, 2009).

The World Health Organization's (WHO) Global Status Reports on Road Safety highlights the potential of efficacious RTI interventions, but what remains missing is evidence on the effectiveness of programmes in LMIC settings (Peden et al., 2004; World Health Organization Department of Violence and Injury Prevention and Disability, 2009). The need to evaluate such programmes in LMICs is urgent (Perel et al., 2007) in order to inform injury prevention policies that are crafted to suit the sociocultural and economic climates unique to each country (Borse & Hyder, 2009; Evans & Ebrary Inc., 2001). Considering the diversity of LMICs, the need for local evidence and research is especially important. The effectiveness of programmes often depends on the context in which they are delivered, and in the case of behaviour-targeting interventions, is often tied to cultural norms and conduct (Chinnock, Siegfried, & Clarke, 2005; Perel et al., 2007). Finally, in settings where resources are scarce and health burdens high, utilising effective interventions is key to maximising resource use and improving health outcomes (Chinnock et al. 2005; Volmink, Swingler, & Siegfried, 2001).

The goal of this paper is to present a scientific approach for evaluating road safety programmes in LMIC settings. We review selected approaches for evaluating health programmes and discuss their applicability to road safety in developing countries. To illustrate a real world application, this paper describes the plan for a large-scale multi-country initiative, the Road Safety in 10 Countries Project (RS-10). This paper highlights the opportunity to generate much needed evidence for preventing road deaths in LMICs over the next five years (The Public Library of Science Medicine Editors, 2010).

RS-10 project

The RS-10 project is a multi-country programme aimed at reducing the global burden of RTIs funded by the Bloomberg Philanthropies with US$125 million over a five-year timeline (2010–2014). The RS-10 project involves a consortium of six partners: WHO, Johns Hopkins International Injury Research Unit (IIRU), the World Bank Global Road Safety Facility (WB), Global Road Safety Partnership (GRSP), Association for Safe International Road Travel (ASIRT) and EMBARQ-the World Resources Institute Center for Sustainable Transport. The project focuses on 10 countries that account for almost half (48%) of all traffic deaths globally: Brazil, Cambodia, China, Egypt, India, Kenya, Mexico, Russia, Turkey and Vietnam (World Health Organization Department of Violence and Injury Prevention and Disability, 2009). RS-10 project's overarching goal is to reduce death and serious injury caused by RTIs in LMICs by concentrating on evidence-based interventions. Therefore, an evaluation of the programme and assessment of its impact offers a unique opportunity for large-scale evaluation in the field of road safety.

RS-10 consists of a relatively standardised approach to designing interventions and evaluation in all participating countries while providing considerable flexibility and autonomy to each country (Hyder et al., 2012). At each national level, a working group of stakeholders has been created and a national work plan developed. Road safety legislation is being reviewed, police are being trained, civil society is being engaged and national awareness and safety campaigns are being launched. Two or more focused intervention sites (i.e. cities, districts or regions) have been identified in each country based on high RTI death rates, local political support, presence of appropriate collaborators, data availability and the potential to serve as a model for neighbouring regions. Each site is being encouraged to address two of four potential risk factors (for example, drink driving, excessive speed, lack of seat belt/child restraint use, lack of helmet use; Table 1). The selection of the intervention sites and targeted risk factors are undertaken in collaboration with national and international partners.

The project is attempting to ensure that at least some comparison area (or data) is defined in each country. In some countries (like Brazil and Vietnam), specific cities or districts have been identified as comparison sites; this selection was based on perceived similarity of the sites by national stakeholders and the willingness of local governments to participate in the project. In other RS-10 countries (like Egypt and Kenya), secondary sources are being used to collect historical data of several potential comparison areas, or national-level trends will be used as comparison for site-specific data. Matched selection of control sites or random allocation of interventions was not possible in this programme.

In each site, an evidence-based and nationally relevant set of interventions is being implemented to address the chosen risk factors. Generally they include: (1) legislative review and changes at the national level and at intervention sites; (2) police enforcement at intervention sites; (3) social marketing and public education campaigns nationally and at intervention sites; and (4) mobilisation of civil society to create demand for road safety at national and at intervention sites. For example, in the case of helmets, national legislation to ensure universal requirement of helmets, strict police enforcement of helmet wearing for all riders including random checks, promotion of helmet wearing by changing social norms and working through local

Table 1. RS-10 project sites and risk factors.

Country	Risk factors	Intervention sites
Brazil	Alcohol	Palmas
	Speed	Teresina
		Belo Horizonte
		Curitiba
		Campo Grande
Cambodia	Alcohol	Phnom Penh
	Helmets	Kandal
		Kampong Speu
China	Alcohol	Dalian
	Speed	Suzhou
Egypt	Speed	Cairo
	Seat belts	Alexandria
India	Alcohol	Jalandhar
	Helmets	Hyderabad
	Speed	
Kenya	Speed	Thika
	Helmets	Naivasha
Mexico	Alcohol	León
	Seat belts/child restraints	Guadalajara and Zapopan
Russia	Seat belts	Lipetsk Oblast
		Ivanovo Oblast
Turkey	Speed	Afyon
	Seat belts/child restraints	Ankara
	Helmets	
Vietnam	Alcohol	Ninh Bình
	Helmets	Ha Naam

NGOs may be done as a package of interventions. Local agencies participate in the planning and are responsible for guiding the project, coordinating the activities of national and international organisations and overseeing the implementation of each intervention, while national and state governments are partners. Actors from relevant sectors such as health, transport, police, law and education are being engaged as needed, and encouraged to participate and cooperate at country and site levels.

Given the complexity and enormity of the RS-10 project, the use of a suitable evaluation framework is required. In an effort to develop an evaluation framework for multi-country road safety programmes such as RS-10, we reviewed selected literature on conducting large-scale evaluations for health programmes.

Conceptual approach to evaluation: selected approaches

In this section, we provide a concise overview of selected large-scale public health evaluation approaches and their applicability to the field of road safety. We selected those approaches that have been successfully applied in a LMIC setting, that encompass a programme evaluation perspective as opposed to an intervention efficacy one, and have been used in public sector health programmes. Given the paucity of literature on evaluation frameworks and approaches that fit these inclusion criteria, we isolated four approaches to examine further: Habicht, Victora, and Vaughan (1999), Bryce & Victora (2005), Peters & World Bank (2009) and Victora, Black, Boerma, and Bryce (2011) (Table 2).

Key issues in evaluation design are the research question and the motivation behind an evaluation, which influence how it will be conducted (Habicht et al., 1999). Habicht et al. assert that the main objective behind conducting an evaluation is to influence decision-making. They proposed a framework that depends on a continuum defined by two axes: the first documents what is to be measured (i.e. impact or performance); the second axis refers to what sort of inference is needed (i.e. adequacy, plausibility or probability of an intervention's effectiveness). Hence, the nature of the evaluation hinges upon the effect to be measured and the desired precision of the findings (Habicht et al., 1999). If the question is whether behaviours or outcomes changed among the beneficiaries of a programme, then an *adequacy* approach could be used. Questions regarding whether the intervention or programme had effects beyond those expected by secular or external influences require at least a *plausibility* approach. These secular and external influences create a challenge to inference in a plausibility approach and it can be less persuasive to stakeholders who demand more rigorous evidence of impact. Finally, in this approach if one wants to quantify an intervention's effect with only a small known probability of change due to confounding (and other biases), then an evaluation design should seek a *probabilistic* inference. For example, if an evaluation were interested in assessing the plausibility of an intervention aimed at increasing health service coverage (i.e. a type of performance measurement), then a longitudinal-controlled (or with comparison groups) design might be appropriate. In contrast, if an evaluation wished to simply assess the adequacy of the same programme, then a before–after or longitudinal study without a comparison group might be suitable.

Peters & World Bank build on Habicht's framework by introducing a fourth level of inference that an evaluation design may address (Table 2). In this case, the type of inference desired is *explanatory* in nature. The focus is on how or why a particular intervention worked (or failed). Here, the question as to *how* the implementation of an intervention leads to measured effects may be answered using a mixed-methods approach. Quantitative designs such as cross-sectional studies, longitudinal studies and randomised-controlled trials may be supplemented with qualitative methods such as in-depth interviews, triangulation and focus groups (Peters & World Bank, 2009). As Habicht et al. (1999) illustrate, often evaluation projects call for more than one type of study and most robust evaluations will utilise more than one method.

Real-world application of these conceptual models has provided further lessons for the large-scale evaluation of health programmes. The evaluation of the *Integrated Management of Childhood Illness* (IMCI) programme, which was aimed at improving child health in LMICs over the past decade, is an important example yielding several methodological lessons and approaches for further consideration when conducting large-scale real-world evaluations (Bryce & Victora 2005; Bryce, Victora, Habicht, Vaughan, & Black, 2004). These include the importance of defining the breadth of the evaluation; selection of the intervention sites using well-defined criteria; creation of 'rules of engagement' to guide implementers; anticipation of validity problems; developing contingency plans; allowing appropriate time; crafting mechanisms for feedback; and budgeting adequately for evaluation (Bryce & Victora, 2005).

Table 2. Overview of selected evaluation approaches for large-scale health programmes and its application to the RS-10 project.

	Habicht et al. (1999)	Bryce & Victora (2005)	Peters & World Bank (2009)	Victora et al. (2011)
Description	A conceptual framework to address *how* an evaluation should be designed	Methodological considerations and lessons for large-scale public health interventions	Focuses on a using a multi-method evaluation to understand why or how an intervention succeeds or fails	An approach for the evaluation of scale-up of national programmes for maternal and child survival
Focus of evaluation	Performance and impact	Impact	Process	Impact
Example of areas of utilisation	Diarrhoea management	Child health and development	Health service delivery strategies	Child health and development Maternal health
Methodological considerations	Varied-depends on the research question	Varied-dependent on country under study	Mixed methods	Varied
Strengths	Provides a systematic and concise way of thinking about evaluation design	Based on real-world application of a large-scale evaluation	Builds on existing frameworks by focusing on process A multi-method approach	A thorough and comprehensive approach to evaluation of large-scale projects
Limitations	Focuses solely on outcomes not processes Assumes the main reason for evaluation is to influence decisions	Not a defined method but a guideline	Does not directly address effectiveness or efficacy questions Time and resource consuming approach	Relies heavily on cooperation and transparency between all partners A potentially costly and time-consuming approach
Application for RS-10	Use of a plausibility approach with a pre-/post-intervention design using controls (i.e. comparison sites)	Use of external evaluators while developing internal capacity	Use of a mixed methods approach with a qualitative arm	Data collection via multiple methods and use of interim and summative evaluation

Note: Please see references for full citation.

Victora et al. build on this approach and addressed the challenge of comparison groups and the difficulty of locating a 'virgin' locale with no existing interventions (Table 2). They advocate the use of baseline-level measurement and the continuation of these measurements as the comparison; a national platform approach where districts are used as the units of analysis; continuous monitoring throughout the intervention timeline; and increasing the range of analytical techniques employed to deal with gaps in data (Victora et al., 2011). We believe the mix of solid conceptual thinking, real-world application and length of time (over 10 years) of experience makes these lessons from IMCI unique in public health, and suitable for application to other areas such as road safety.

Taken together these various approaches provide a valuable base from which to create a suitable evaluation framework for a large-scale multi-country initiative to reduce RTIs and deaths (Table 2). The following section uses these approaches to develop such a framework for the RS-10 project.

An evaluation approach for RS-10

Our evaluation approach for RS-10 predominately follows the Habicht et al.'s (1999) framework since our focus is the impact on health outcomes (i.e. mortality and injuries) with the realisation that the programme operates within a context of other ongoing interventions and secular trends. Thus, we have selected a 'plausibility' approach. Governments and other non-governmental actors in RS-10 countries are likely doing something for road safety, though often little for the specific risk factors considered under RS-10. As a result, the plausibility approach will allow us to attribute some part of any detected change to the RS-10 programme. Therefore, a pre-/post-intervention design with controls (i.e. comparison sites) will be used with multiple methods for collecting data. However, it is important to note that we are also drawing on considerations developed by Peters & World Bank (2009), Bryce et al. (2004) and Victora et al. (2011); as a result our overarching evaluation framework employs a mixed methods approach and serves to build on previous approaches (Table 2).

The overall evaluation approach for RS-10 was designed to be pragmatic and to follow lessons learned from heath programme evaluation. Based on the work of Bryce et al. (2004) and Victora et al. (2011), we have summarised 13 main lessons from key large-scale child health programme evaluations and have applied them to our evaluation approach for RS-10. Table 3 shows the lessons and how they are being applied in RS-10.

The selection of study sites was not done randomly, but through specific criteria that were applied by each team and the set of stakeholders in their populations. For evaluation purposes, the non-random selection of sites introduces a bias because these factors are expected to be associated with both traffic fatality and programme participation (Angeles, Gullkey, & Mroz, 1998; Wooldridge, 2002). A *difference in difference* design reduces the bias due to time invariant factors, and the potential selection bias is further reduced by the incorporation of baseline data on site selection criteria from intervention and comparison sites into the analysis.

The use of *external evaluators* that are independent of the implementing agencies is generally necessary to enhance credibility and maintain objectivity (Bryce & Victora, 2005). RS-10 incorporates this approach by using a consortium, where one dedicated partner (Johns Hopkins-IIRU, authors) serves as the evaluating agent, whereas others are responsible for actually executing the interventions (Table 3). In addition, we have an independent contract with the donor and funds dedicated for each country,

Table 3. Thirteen lessons[a] of large-scale programme evaluation and their application to the RS-10 programme.

Lesson	Operational definition	Application to RS-10
Defining the scope	Define the depth and breadth of the evaluation at the outset	Decision at start to focus on plausibility and explanatory inferences since to assess the programme's effectiveness
Selecting study sites and developing research plans	Develop criteria for selecting study sites	Focused intervention sites identified in each country according to standardised criteria: high road traffic death rates, political support, readiness to implement interventions, presence of appropriate collaborators, data availability, potential to serve as model for neighbouring regions.
Maintaining objectivity	Conduct an independent and objective evaluation (use of external evaluators, separate funding stream)	Use of one partner (Johns Hopkins-International Injury Research Unit) as an independent evaluator who is funded directly, with no role in implementation
Developing an impact model	Collaborate with programme implementers to develop an impact model (pathways through which interventions are expected to lead to outcomes)	RS-10 consortium developing projections for likely impact of proposed interventions
Multiple data sources	Use of various methods for data (facility assessments, household surveys, longitudinal, qualitative research)	Primary and secondary data being collected including: observational studies, roadside interviews, focus group discussions, in-depth interviews, household surveys, in-country statistical data, surveillance data
Multiple analytic techniques	Variety of data analysis techniques are required to effectively deal with data gaps and biases	Use of multiple analytical techniques and approaches under way
Maximising external validity	Anticipate problems with external validity and develop mechanisms to address them	Detailed study of interventions sites and populations in reference to national data
Ensuring an appropriate time frame	Allowing sufficient time for the evaluation	Five-year project timeline allows for adequate time to implement interventions and assess their effectiveness
Importance of flexibility and a stepwise approach	Built in flexibility and a stepwise approach allow periodic assessment and mid-course corrections	Monitoring of selected indicators and risks with feedback to consortium to allow for mid-course adjustments.

Table 3 (*Continued*)

Lesson	Operational definition	Application to RS-10
Continuous monitoring	Evaluation as a continuous process aimed at improvement of implementation	Continuous monitoring of RS-10 progress through regular data collection on risk factors and analysis of secondary data
Providing feedback to implementers, policymakers	Include plans for feedback and dissemination to implementers and policy-makers	National working groups and committees briefed regularly by the consortium
Promoting the uptake of evaluation results	Develop opportunities for direct contact among researchers and policy-makers	Regular and direct contact between researchers and implementers at the international and national level through project working groups in each country
Evaluation costs	Budget for adequate support to analyse, interpret, write and disseminate results	Separate funds earmarked for monitoring and evaluation including analysis, interpretation, workshops and publications

[a]Based on the studies of Habicht et al. (1999), Victora et al. (2011) and Bryce & Victoria (2005).

ensuring that all costs will be adequately covered. Furthermore, funds have been designated to aid in the dissemination and uptake of the evaluation's findings.

Bryce and Victora have indicated that the sole use of external evaluators fails to develop *in-country capacity* (2005). Hence, the use of local collaborators who will work in tandem with us is a crucial element of RS-10 evaluation. In each participating country, we have identified a strong collaborating partner, usually a research or academic institution, to help conduct the evaluation. A programme for capacity building based on experiential learning, targeted training and focused mentoring has also been implemented (Potter & Brough, 2004). This will help ensure that technical capacity development goes hand in hand with the evaluation work using both face-to-face and electronic modalities (Table 3). The associated issue of sustainability of interventions is an important one for RS-10, and consortium partners will be addressing this element (Hyder et al., 2012).

The use of *standardised methods* for data collection is a key in large-scale evaluations that span multiple countries (Table 3). We have adopted this lesson in order to guarantee comparability of data across study sites. This is also important to avoid dependency on routinely collected data. Routine data, while important for the country, are often neither specific, nor timely enough for rigorous programme evaluation purposes, especially in settings where information systems are nascent. Throughout the project's five-year timeline, we will be monitoring selected indicators using these multiple methods. This will also permit continuous *feedback* and alterations during the course of the project (Table 3).

RS-10 will also use a variety of *analytic methods* to allow for both plausibility and explanatory inferences. This is central to identifying if any part of the change in road deaths, injuries or crashes can be attributed to the RS-10 project despite the presence of both secular trends and other related interventions. The use of comparison sites and data will play an important role in this analysis. Together we feel that RS-10 attempts to build on the lessons learnt from other large-scale public health evaluations and apply them (for the first time to our knowledge) in the field of road safety.

RS-10 data sources and methods

The inherent challenge of drawing inferences in the presence of secular and external influences creates the need for multiple, overlapping studies in the evaluation of RS-10 (Table 4). In this section, we will discuss these data collection methods along with their strengths and weaknesses. A combination of some or all of these data collection methods are being utilised in each RS-10 country, depending on the chosen interventions and risk factors, availability of existing data, and logistical feasibility and acceptability.

The use of *observational studies* remains a key component of the RS-10 evaluation. Such studies 'observe' or record the behaviours or risks exhibited by a random sample of those actually on the road – thus exposed to the risk of RTIs (Lange & Voas, 1998; Segui-Gomez et al., 2001). This will involve, for example, observing motorcycle riders to count if they are wearing helmets on different types of roads; or observing drivers and front-seat passengers in vehicles to count if they are wearing seat belts. These studies are relatively straightforward to execute and do not necessitate any direct human subject interaction. Furthermore, they permit the execution of multiple rounds of data collection over relatively short periods of time, allowing for rapid assessment of change.

We will also be utilising *roadside interviews* of a random sample of drivers on selected roadways (Table 4). A short set of questions on the behaviour and practices of road users is administered on a variety of risks that range from drink driving, motorcycle helmet use, seat belt use and speeding, in addition to a few demographic indicators (for example, age range, gender) (Huang, Zhang, Murphy, Shi, & Lin, 2011; Kulanthayan et al., 2010). These necessitate human subject interaction and can be more challenging since travelling motorists must be safely stopped, and they have to be willing to participate. However, their advantage is that they allow for inquiry to occur as close as possible to the exposure of interest (i.e. the road); and they can be done quickly, so multiple rounds of data collection are possible over relatively short periods of time.

In addition, the evaluation aims to uncover opinions and attitudes behind specific road behaviours; therefore, it includes a *qualitative arm* with focus group discussions and in-depth interviews. These methods can help uncover why certain behaviours and risky practices occur, how they may be connected to social and cultural beliefs, and how might they change over time (Flick, Kardorff, & Steinke, 2004). More importantly, they are being used to understand the perspectives of various stakeholders (such as motorcycle riders and passengers) and their interests in road safety at the national and local levels (Khorasani-Zavareh et al., 2009). The inclusion of this qualitative arm follows evaluation approaches recommended by Peters & World Bank, which promote the collection of more in-depth information. These methods require intensive human subject interactions, but focus on gathering a representation of ideas and opinions from those at risk of RTIs or involved in the amelioration of risks for RTIs.

Household surveys collect population-based data in a well-defined community and are powerful tools for generating in-depth data on recent injury, risk behaviours

Table 4. Data collection methods for RS-10 evaluation.

Method	Focus	Description	Proposed frequency
Observational studies	Risk factors	Records the risks or behaviours on a random sample of those exposed to road injuries (on a road)	3–6 per year
Roadside surveys	Knowledge and practices	Surveys of a random sample of drivers on selected roadways	3–5 per year
In-depth interviews/ focus group discussions	Attitudes and perceptions	Semi-structured interviews or discussion groups with key informants on national/local road safety	2–3 times per year
Injury surveillance systems	Hospital burden of injury and treatment details	System to collect standardised information for each injured patient arriving at an emergency department	Continuous
Household surveys	Community burden, perceptions, behaviours, socio-economic profile	Surveys that collect population-based data in a sample of households in each intervention site	2 per 5 years

and norms, and denominator data on road use and RTI risk exposure (Babbie, 1973; Fowler, 2009). We will use a random sample of households in each site to ascertain detailed demographic, social, behavioural, risk, economic and injury data (Table 4). Such an approach allows us to define the community experiences of RTIs, document the health consequences and capture the economic burden of RTIs at the household level (Dandona, Kumar, Raj, & Dandona, 2006). In the field of RTI research, large sample sizes are required to capture adequate data, making them expensive and time-consuming; as a result, one round of surveys will be done in select countries in the five years of the RS-10 project.

Surveillance data collected at health facilities provides important RTI data on those who access hospitals, and helps strengthen hospital-based health information systems (Table 4). Such an injury, surveillance system is a continuous effort to collect standardised information for each injured patient arriving at an emergency department and in-depth information on pre-hospital events and injury characteristics (Mutto, Lawoko, Nansamba, Ovuga, & Svanstrom, 2011). In addition, it includes data on elements of the care provided and outcomes in the emergency departments (Murray, Lopez, Barofsky, Bryson-Cahn, & Lozano, 2007). However, the development of a functioning surveillance system requires extensive coordination at the hospital-level and dedicated staff time. We are in the process of establishing such a surveillance system in the largest public sector hospital in several sites and will use the data as part of our RS-10 evaluation.

All of the methods discussed above highlight the importance of collecting primary data and its use in evaluation designs. However, secondary data obtained through routine information systems have also been identified and its strength and validity assessed (Murray & Lopez, 2010). This data come from the police, hospital medical records, forensic offices, ministries of health, national statistical offices, ambulance systems, ministries of transport and local governmental agencies. Each of these secondary data sources provides information on specific aspects of the RTI burden in each site and also offers comparison data from other parts of each RS-10 country. Hence, in line with the principles espoused by Habicht et al. (1999) we will draw on a variety of data sources and utilise both primary and secondary data to inform our evaluation.

Analytic approaches for RS-10 evaluation

The effectiveness of the RS-10 programme will be primarily measured in terms of the reduction in road traffic fatalities, injuries and crashes in the intervention sites over the five-year observation period and compared to comparison data or comparison sites. In order to enhance rigour, three analytic approaches will be used. First, regular review of the overall road fatality rate per 100,000 population from collected data will allow us to test for differences in the average road fatality rate between the RS-10 sites and appropriately chosen comparison sites. We will be using a *difference in difference* approach to test whether the change in road fatality rates in the intervention sites is significantly different from that in comparison sites; this approach helps reduce the biases that result from non-randomised selection of sites and from changes in repeat observations that occur naturally over time (Wooldridge, 2002). We will also estimate the *adjusted relative risk* for road fatality in each RS-10 country's programme site vs. comparison site or data. Means and proportion of each background characteristic (i.e. numbers of cars, population, urbanity and vehicle density) will be estimated for each site, and compared across intervention and comparison sites by appropriate statistical tests to assess baseline differences. Standard denominators of 100,000 population and where possible, one million travelled will also be applied.

In a second analytic approach, we will also base the difference in differences estimator utilising a *regression* model to quantify the differences between pre- and post-measurements in the intervention vs. control settings. Utilisation of a regression model enables controlling for regional and other factors. Since the road fatality rate is a statistically uncommon event, we will first model our multivariate analysis using a *Poisson* distribution, as this distribution is most commonly assumed for events as statistically rare as traffic fatalities. Tests of over-dispersion will be run and, if needed, we will use a mixture model such as a *negative binomial model* to allow for adjustment of variance independent of the mean, and retest for over-dispersion. We will report results based on the preferred model from the over-dispersion test.

Finally, we will conduct an interrupted *Time Series Analysis* (TSA) using the monthly number of deaths from traffic crashes. TSA will map the count of deaths over time in the intervention area and test for differences before and after the intervention. TSA is complementary to the difference in differences methods above because it takes into account that progressive observations are serially correlated. More importantly, TSA 'avoids the loss of information about variability in incidence over time that occurs when rates are aggregated into one before and one after rate', and TSA is a more efficient method for obtaining estimates, due to lower standard errors. TSA has been used in similar studies including assessment of community-based interventions on child injuries, alcohol-related injuries and occupational injuries (Monforton & Windsor 2010; Novoa, Perez, & Borrell, 2009; Razvodovsky, 2012). By utilising all three of these analytic approaches, we hope to provide a rigorous assessment of the RS-10 programme's effect.

While the approach described above provides an impact evaluation, we realise the importance of also providing process and formative evaluations. As a result, other consortium partners will be addressing process and formative evaluations for the work each one is doing (Table 2).

Conclusion

The dearth of road safety research in LMICs needs to be addressed while evidence-based programmes are implemented. The use of relatively new approaches for evaluation of real-world programmes allows for the production of relevant knowledge, which can be useful locally and also contribute to the global pool of knowledge. The approaches used in child health and health systems research provide much needed frameworks for road safety programme evaluation. The RS-10 project affords an important opportunity to scientifically test these approaches for a *real-world, large-scale road safety evaluation* and generate new knowledge for the field of road safety. Such evidence is more likely to influence local and national policy matters than externally transported knowledge. This is an important outcome, especially for the scale-up of proven interventions in the developing world (Mangham & Hanson, 2010).

Acknowledgements

The Road Safety in Ten Countries Project is supported by Bloomberg Philanthropies. This paper was developed as part of the evaluation plans for the project. We would also like to thank our colleagues at Johns Hopkins Bloomberg School of Public Health and selected members of consortium partners, and anonymous reviewers for their helpful comments.

References

Angeles, G., Gullkey, D. K., & Mroz, T. A. (1998). Purposive program placement and the estimation of family planning program effects in Tanzania. *Journal of the American Statistical Association*, *93*(443), 884–899. doi:10.1080/01621459.1998.10473745

Babbie, E. R. (1973). *Survey research methods*. Belmont, CA: Wadsworth.

Borse, N. N., & Hyder, A. A. (2009). Call for more research on injury from the developing world: Results of a bibliometric analysis. *Indian Journal of Medical Research*, *129*(3), 321–326.

Bryce, J., & Victora, C. G. (2005). Ten methodological lessons from the multi-country evaluation of integrated management of childhood illness. *Health Policy & Planning*, *20*(Suppl. 1), i94–i105. doi:10.1093/heapol/czi056

Bryce, J., Victora, C. G., Habicht, J. P., Vaughan, J. P., & Black, R. E. (2004). The multi-country evaluation of the integrated management of childhood illness strategy: Lessons for the evaluation of public health interventions. *American Journal of Public Health*, *94*(3), 406–415. doi:10.2105/AJPH.94.3.406

Chinnock, P., Siegfried, N., & Clarke, M. (2005). Is evidence-based medicine relevant to the developing world? Systematic reviews have yet to achieve their potential as a resource for practitioners in developing countries. *Evidence-Based Complementary and Alternative Medicine*, *2*(3), 321–324. doi:10.1093/ecam/neh114

Dandona, R., Kumar, G. A., Raj, T. S., & Dandona, L. (2006). Patterns of road traffic injuries in a vulnerable population in Hyderabad, India. *Injury Prevention*, *12*(3), 183–188. doi:10.1136/ip.2005.010728

Evans, T., & Ebrary Inc. (2001). *Challenging inequities in health from ethics to action*. Oxford, UK; New York, NY: Oxford University Press.

Flick, U., Kardorff, E. V., & Steinke, I. (2004). *A companion to qualitative research*. London; Thousand Oaks, CA: Sage.

Fowler, F. J. (2009). *Survey research methods*. Thousand Oaks, CA: Sage.

Habicht, J. P., Victora, C. G., & Vaughan, J. P. (1999). Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. *International Journal of Epidemiology*, *28*(1), 10–18. doi:10.1093/ije/28.1.10

Huang, Y. H., Zhang, W., Murphy, L., Shi, G., & Lin, Y. (2011). Attitudes and behavior of Chinese drivers regarding seatbelt use. *Accident Analysis & Prevention*, *43*(3), 889–897. doi:10.1016/j.aap.2010.11.009

Hyder, A. A., Allen, K. A., Di Pietro, G., Adriazola, C. A., Sobel, R., Larson, K., & Peden, M. (2012). Addressing the implementation gap in global road safety: Exploring features of an effective response and introducing a 10-country program. *American Journal of Public Health*, *102*(6), 1061–1067. doi:10.2105/AJPH.2011.300563

Khorasani-Zavareh, D., Mohammadi, R., Khankeh, H. R., LaFlamme, L., Bikmoradi, A., & Haglund, B. J. (2009). The requirements and challenges in preventing of road traffic injury in Iran. A qualitative study. *Bio Medical Central Public Health*, 9(486), 1–9. doi:10.1186/1471-2458-9-486

Kulanthayan, S., Razak, A., & Schenk, E. (2010). Driver characteristics associated with child safety seat usage in Malaysia: A cross-sectional study. *Accident Analysis & Prevention*, *42*(2), 509–514. doi:10.1016/j.aap.2009.09.015

Lange, J. E., & Voas, R. B. (1998). Nighttime observations of safety belt use: An evaluation of California's primary law. *American Journal of Public Health*, *88*(11), 1718–1720. doi:10.2105/AJPH.88.11.1718

Mangham, L. J., & Hanson, K. (2010). Scaling up in international health: What are the key issues? *Health Policy and Planning*, *25*(2), 85–96. doi:10.1093/heapol/czp066

Monforton, C. & Windsor, R. (2010). An impact evaluation of a federal mine safety training regulation on injury rates among US stone, sand, and gravel mine workers: An interrupted time-series analysis. *American Journal of Public Health*, *100*(7), 1334–1340. doi:10.2105/AJPH.2009.178301

Murray, C. J., & Lopez, A. D. (2010). Production and analysis of health indicators: The role of academia. *Public Library of Science Medicine*, *7*(11), e1001004. doi:10.1371/journal.pmed.1001004

Murray, C. J., Lopez, A. D., Barofsky, J. T., Bryson-Cahn, C., & Lozano, R. (2007). Estimating population cause-specific mortality fractions from in-hospital mortality: Validation of a new method. *Public Library of Science Medicine*, *4*(11), e326. doi:10.1371/journal.pmed.0040326

Mutto, M., Lawoko, S., Nansamba, C., Ovuga, E., & Svanstrom, L. (2011). Unintentional childhood injury patterns, odds, and outcomes in Kampala City: An analysis of surveillance data from the National Pediatric Emergency Unit. *Journal of Injury & Violence Research*, *3*(1), 13–18. doi:10.5249/jivr.v3i1.56

Novoa, A. M., Perez, K., & Borrell, C. (2009). Evidence-based effectiveness of road safety interventions: A literature review. *Gaceta Sanitaria*, *23*(6), 553.e1–553.e14. doi:10.1016/j.gaceta.2009.04.006

Peden, M. M., World Health Organization, & World Bank. (2004). *World report on road traffic injury prevention*. Geneva: WHO.

Perel, P., Ker, K., Ivers, R., & Blackhall, K. (2007). Road safety in low- and middle-income countries: A neglected research area. *Injury Prevention*, *13*(4), 227. doi:10.1136/ip.2007.016527

Peters, D. H., & World Bank. (2009). *Improving health service delivery in developing countries from evidence to action. Directions in development*. Washington, DC: World Bank.

Potter, C. & Brough, R. (2004). Systemic capacity building: A hierarchy of needs. *Health Policy and Planning*, *19*(5), 336–345. doi:10.1093/heapol/czh038

Razvodovsky, Y. E. (2012). Contribution of alcohol in accident related mortality in Belarus: A time series approach. *Journal of Injury and Violence Research*, *4*(2), 58–64. doi:10.5249/jivr.v4i2.100

Segui-Gomez, M., Wittenberg, E., Glass, R., Levenson, S., Hingson, R., & Graham, J. D. (2001). Where children sit in cars: The impact of Rhode Island's new legislation. *American Journal of Public Health*, *91*(2), 311–313. doi:10.2105/AJPH.91.2.311

The Public Library of Science Medicine Editors. (2010). Preventing road deaths – time for data. *Public Library of Science Medicine*, 7(3), e1000257. doi:10.1371/journal.pmed.1000257

Victora, C. G., Black, R. E., Boerma, J. T., & Bryce, J. (2011). Measuring impact in the millennium development goal era and beyond: A new approach to large-scale effectiveness evaluations. *The Lancet*, *377*(9759), 85–95. doi:10.1016/S0140-6736(10)60810-0

Volmink, J., Swingler, G., & Siegfried, N. (2001). Where to practise evidence-based medicine? *The Lancet*, *357*(9257), 724. doi:10.1016/S0140-6736(05)71488-4

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.

World Bank. 2002. *World Bank* [online]. Retrieved from http://www.worldbank.org/transport/roads/safety.htm

World Health Organization, Etienne, C., Asamoa-Baah, A., & Evans, D. B. (2010). *The World health report: Health systems financing: The path to universal coverage*. Geneva: World Health Organization.

World Health Organization Department of Violence and Injury Prevention and Disability. (2009). *Global status report on road safety: Time for action*. Geneva: Author.