ORIGINAL ARTICLE

Robust cross-country comparison of children meeting 24-HR movement guidelines: an odds solution for binary effect efficiency measures

Simon Eckermann[1] · Andrew R. Willan[2] · Anthony Okely[1] · Mark S. Tremblay[3] · Tim Coelli[4]

Abstract

Aim International comparisons of public health measures relative to observed best practice are increasingly important for evaluating community health promotion strategies and policies such as meeting or not meeting public health guidelines. This study aimed to identify methods enabling robust evaluation with such binary effect measures at a population level.

Subject and methods Conventional efficiency comparisons of binary effect proportions are problematic due to a lack of consistency with alternate framing of resulting relative risks. In this paper, we illustrate such inconsistent efficiency measures comparing the proportion of school age children (9–11 years) meeting or not meeting integrated movement guidelines (IMGs) across the 12 countries from the International Study of Childhood Obesity, Lifestyle and the Environment (ISCOLE) study. IMGs jointly consider physical activity, sleep and sedentary behaviours. An odds method is developed to enable consistent efficiency comparison with alternative framing of binary effects.

Results A novel odds solution to relative risk problems arising with conventional efficiency comparison of binary effects with alternative framing is shown to provide consistent efficiency measures relative to best practice. Furthermore, this technical advancement is shown to extend to consistent indirect comparison and evidence translation.

Conclusion Robust methods for international cross-country comparison of binary effect measures such as meeting or not meeting guidelines are identified with a novel odds ratio method. This novel solution is particularly important for health promotion evaluation of IMGs given the need for consistent comparison in evaluating practice evidence of what works now and consistent evidence translation of treatment effects as and when they emerge.

Keywords Health promotion evaluation · Binary effect efficiency measures · Relative risk fallacy · Odds ratio solution

Background

Robust processes of health economic evaluation generally need to jointly consider principles of coverage and comparability given an underlying objective of optimising budget constrained community net benefit that underlies societal decision-making across related research, reimbursement and regulatory decisions (Eckermann 2017; Eckermann and Pekarsky 2014; Eckermann and Coelli 2013; Eckermann et al. 2010; O'Brien 1996). For robust evaluation of health promotion strategies in community settings, satisfying adequate coverage in practice needs to allow for population level impacts in communities and in particular the potential of community network and multiplier effects across community networks (Shiell and Hawe 1996; Shiell et al. 2008; Hawe and Ghali 2008; Hawe et al. 2009; Eckermann et al. 2014; Eckermann and McCaffrey 2017). The public health and health economic potential of health promotion strategies, programs and policies with community ownership of effective strategies are substantial at a population level given community network effects can naturally broaden both the population exposed and their duration of effects without increasing intervention costs. The potential of network effects is particularly

✉ Simon Eckermann
seckerma@uow.edu.au; simoneck4@hotmail.com

[1] School of Health and Society, Faculty of Arts Social Sciences and Humanities, University of Wollongong, Wollongong, Australia

[2] University of Toronto, Toronto, Canada

[3] Children's Hospital of Eastern Ontario Research Institute, Ottawa, Canada

[4] School of Economics, University of Queensland, St. Lucia, Australia
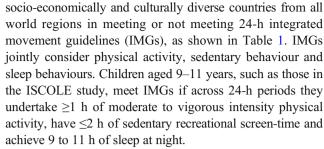
large where community ownership of health promotion extends to integrating strategies into practice and policies of community institutions. For example, integration of the Stephanie Alexander Kitchen Garden Program kitchen and garden lessons into school curricula (Eckermann et al. 2014; Yeatman et al. 2014) or the potential for policy integration of early childhood movement guidelines into early childhood care settings (Okely et al. 2017).

This potential for community network impacts and broadening of effects of health promotion strategies and policies with community ownership and integration do not arise with individual focussed strategies (Eckermann and McCaffrey 2017) and point to potential for significant underfunding of health promotion research and strategies in optimising budget constrained health system outcomes (Shiell and Jackson 2018). However, bottom-up methods to map the network and multiplier impacts over time of health promotion strategies with community ownership across populations are not yet well developed. This has frustrated calls for adequate coverage of impacts (effects and downstream cost savings) in evaluating health promotion policies and strategies in practice and associated research pathways and adequate investment in identifying and unlocking the health system potential of successful health promotion policies and strategies (Shiell and Jackson 2018). While health promotion evaluation guidelines generally recognise the key need to triangulate across evidence and levels of analysis in community settings (Zaza et al. 2005), adequate coverage of population level effects allowing for network multiplier impacts of alternative health promotion policies and strategies is required. In practice this would be best achieved if robust methods were available for top-down population level monitoring across countries of the effects of alternative strategies and policies over time. This is particularly the case given inadequacy of current bottom-up methods to capture community network effects (as considered at length in discussion).

Nevertheless, comparability as well as coverage principles need to be satisfied for robust evaluation of health promotion as in any public health evaluation (Shiell and Jackson 2018; Eckermann 2017; Eckermann and McCaffrey 2017). While efficiency analysis with binary data such as meeting or not meeting guidelines are feasible in adequately allowing for population coverage of effects, such analysis faces distinct methods challenges to satisfy comparability and in particular to enable robust consistent analysis with alternative framing of binary effects. This paper's aim is to address these challenges.

To illustrate these challenges in relative performance efficiency measures and their solution, we consider a public health international comparison in meeting integrated movement guidelines, the International Study of Childhood Obesity, Lifestyle and the Environment (ISCOLE) (Roman-Viñas et al. 2016; Katzmarzyk et al. 2013). The ISCOLE compared school-aged children (age 9–11 years) across 12 socio-economically and culturally diverse countries from all world regions in meeting or not meeting 24-h integrated movement guidelines (IMGs), as shown in Table 1. IMGs jointly consider physical activity, sedentary behaviour and sleep behaviours. Children aged 9–11 years, such as those in the ISCOLE study, meet IMGs if across 24-h periods they undertake ≥1 h of moderate to vigorous intensity physical activity, have ≤2 h of sedentary recreational screen-time and achieve 9 to 11 h of sleep at night.

The extent to which populations of children jointly meet or do not meet IMGs has been shown to be an important population level public health issue given these joint behaviours largely determine long-term habits, lifestyle and risk of non-communicable diseases. International evidence syntheses have shown child populations meeting IMGs have better body composition, cardiorespiratory and musculoskeletal fitness, cardiovascular and metabolic health, academic achievement and cognition, mental health and quality of life, emotional regulation, and pro-social behaviours (Dumuid et al. 2018a, b; Katzmarzyk et al. 2013; Okely et al. 2017; Roman-Viñas et al. 2016; Saunders et al. 2016; Tremblay et al. 2016; Walsh et al. 2018; WHO 2017, 2019). Hence, whether children meet or do not meet IMGs and their joint consideration of sleep, physical activity and sedentary behaviour are key to child health and development and trajectory of their lifestyle and health behaviours more generally.

Nevertheless, when undertaking efficiency comparisons relative to best practice based on binary effects data such as meeting or not meeting IMGs in Table 1, a distinct challenge for robust consistent analysis arises. While binary data such as that from ISCOLE are readily available and make efficiency comparisons relative to observed best practice feasible, they can always be framed as in Table 1 in two alternate ways; meeting or not meeting a guideline, as with any such binary data (success or failure; survival or mortality etc.).

Table 1 ISCOLE international comparison of 9–11 year-old children meeting or not meeting integrated movement guidelines

Country	Proportion meet IMGs	Proportion do not meet IMGs
Australia	0.149	0.851
Brazil	0.034	0.966
Canada	0.140	0.860
China	0.015	0.985
Columbia	0.078	0.922
Finland	0.067	0.933
India	0.060	0.940
Kenya	0.065	0.935
Portugal	0.020	0.980
South Africa	0.119	0.881
UK	0.114	0.886
USA	0.021	0.979

For robust efficiency measures relative to best practice, such binary effects efficiency measures need to be consistent with alternate framing of the binary effect to avoid selection bias or paradoxical results. In other words, to satisfy comparability efficiency measures relative to best practice analysed with the same binary data framed as meeting a guideline or not meeting a guideline, the results should be the same. In this paper, we illustrate the distinct lack of consistency with alternate conventional presentation of data as the probability of meeting or not meeting a guideline. More importantly, we provide methods for an odds solution in methods that is shown in results to: (i) consistently estimates efficiency scores relative to best practice with alternate framing of a binary effect; (ii) provide an intuitive interpretation of relative scores as odds ratios, and; (iii) extend to consistent indirect comparisons, evidence translation and standardisation. Finally, we return to discuss methods required to enable robust bottom-up evaluation of IMGs to aid triangulation, while robust top-down methods for public health efficiency measures are identified, before concluding.

The problem – Inconsistency of relative risk efficiency measures with binary data

For a binary effect such as meeting or not meeting IMGs in Table 1, efficiency measures relative to best practice across countries can be analysed either in terms of trying to maximise population proportion meeting guidelines or minimise the population proportion not meeting guidelines. However, output orientated efficiency analysis comparing the proportion meeting guidelines and input orientated efficiency analysis comparing the proportion not meeting guidelines lead to very different results, as shown in Table 2.

Differences between relative performance in efficiency analysis with alternate framing of the same binary data such as meeting (output orientation) or not meeting guidelines (input orientation) arise because:

(i) Relative performance measures represent relative risk as the population probability of meeting guidelines (P_i for any country i) compared to the best practice peer (P^*) from an output perspective or as the population probability of not meeting guidelines ($1-P_i$) relative to the best practice peer ($1-P^*$) from an input perspective and;

(ii) relative risk is not a symmetric metric; that is, relative risk of meeting guidelines (P_i/P^*) is not the reciprocal of not meeting guidelines ($P_i/P^* \neq (1-P^*)/(1-P_i)$).

Hence, inconsistency of efficiency measures with alternative framing of the same binary effect with relative risk arises due to inherent lack of symmetry of relative risk compared to the best practice peer that relative efficiency measures represent. This inconsistency makes interpretation of relative risk efficiency measures for any binary effect highly problematic given that resulting differences arising with alternate framing of the same binary effect either:

(i) introduces selection bias where only one framing is considered (for example, meeting or not meeting guidelines alone); or

(ii) creates a paradox where both framings are presented making efficiency measures uninterpretable.

To satisfy comparability and avoid either a selection bias or paradox that arises with relative risk efficiency measures requires a method enabling consistent estimation of alternative

Table 2 Relative risk efficiency measures for countries meeting or not meeting 24-h integrated movement behaviour guidelines (IMGs)

Country	Proportion meet IMGs	Efficiency relative to best practice in meeting IMGs	Proportion do not meet IMGs	Efficiency relative to best practice in not meeting IMGs
Australia	0.149	1.000=0.149/0.149	0.851	1.000=0.851/0.851
Brazil	0.034	0.228=0.034/0.149	0.966	0.881=0.851/0.966
Canada	0.140	0.940=0.140/0.149	0.860	0.990=0.851/0.860
China	0.015	0.101=0.015/0.149	0.985	0.864=0.851/0.985
Columbia	0.078	0.523=0.078/0.149	0.922	0.923=0.851/0.922
Finland	0.067	0.450=0.067/0.149	0.933	0.913=0.851/0.933
India	0.060	0.403=0.060/0.149	0.940	0.905=0.851/0.940
Kenya	0.065	0.436=0.065/0.149	0.935	0.910=0.851/0.935
Portugal	0.020	0.134=0.020/0.149	0.980	0.868=0.851/0.980
South Africa	0.119	0.799=0.119/0.149	0.881	0.966=0.851/0.881
UK	0.114	0.765=0.114/0.149	0.886	0.960=0.851/0.886
USA	0.021	0.141=0.021/0.149	0.979	0.869=0.851/0.979

output (meeting guideline) and input (not meeting guideline) presentation.

Methods

Odds and odds ratio solution

To overcome inconsistency of efficiency measures with binary effects arising with the non-symmetric relative risk measures, use of the symmetric odds and odds ratio metrics are suggested, as with consistent evidence synthesis and evidence translation methods (Eckermann et al. 2009, 2011). The key advantage of consistency arises with odds because odds represent probability divided by 1-probability and hence are symmetric for binary measures such as meeting or not meeting a guideline.

That is, undertaking efficiency analysis for meeting or not meeting the IMGs with odds comparison, performance relative to best practice for each country represents:

(i)   from an output perspective, odds of meeting guidelines (Pi/(1-Pi) for any country i relative to the best practice peer with highest odds of meeting guideline (P*/(1-P*)) and hence relative efficiency score Pi/(1-Pi)/P*/(1-P*); and

(ii) from an input perspective best practice peer with lowest odds of not meeting guideline ((1-P*)/P*) relative to odds of not meeting guidelines ((1-Pi)/Pi for any country i) and hence relative efficiency score ((1-P*)/P*)/((1-Pi)/Pi) = Pi/(1-Pi)/P*/(1-P*).

This novel odds ratio method is now shown in results to overcome the inconsistency of relative risk based efficiency relative to best practice highlighted in Table 1, and can be extended to solutions in any indirect comparison and translation of evidence of relative treatment effects to any jurisdiction of interest, employing results from previously developed odds ratio methods (Eckermann et al. 2009, 2011). These methods are illustrated in empirical analysis comparing school-aged children (age 9–11 years) in meeting or not meeting 24-h IMGs across all countries from the ISCOLE (Roman-Viñas et al. 2016, Katzmarzyk et al. 2013). The 12 countries in the ISCOLE study were selected to be representative of socio-economically and culturally diverse countries in every continent and world region.

Results

Consistent efficiency measures

The odds ratio solution for consistent binary effect efficiency measures with comparison of their odds in methods is illustrated in Table 3 for the international comparison of ISCOLE country's odds for meeting or not meeting guidelines relative to observed best practice. That is, efficiency measures comparing 9–11 year old children from each of the ISCOLE study countries relative to Australia in maximising odds of meeting IMGs (Australia's odds 0.175) or minimising odds of not meeting IMGs (Australia's odds 5.711).

Comparing the population odds of meeting or not meeting guidelines results in the same efficiency measure; a consistent estimate of input and output orientated efficiency relative to best practice. The resulting consistent efficiency measures respectively represent the odds ratios of meeting the guidelines relative to best practice in maximising (output orientated) and the reciprocal of the odds of not meeting the guidelines relative to best practice in minimising (input orientated), as shown in Table 3 for the ISCOLE study. Importantly, this implies efficiency measures are consistent with alternate output and input framing in comparing odds of meeting and not meeting guidelines, overcoming the problematic inconsistency arising with relative risk.

The consistent efficiency result arises with use of odds as there are two inversions that cancel each other with alternate framing for odds of binary effects given:

(1)   Odds are symmetric, that is, odds of not meeting guidelines are the reciprocal of the odds meeting guidelines, hence:

odds of not meeting guidelines
= p(not meeting guideline) / (p(meeting guidelines)
= 1/(p(meeting guidelines)/p(not meeting guidelines)
= 1/odds meeting guidelines;

(2)   Relative efficiency measures are relative to the best practice country maximum in meeting guidelines and minimum in not meeting guidelines, the same best practice country with reciprocal odds from (1).

For example, consider efficiency for Brazil. From Table 3 Brazil has odds of meeting guidelines of 0.035 relative to Australian best practice odds of meeting guidelines of 0.175. Alternatively, Australia's odds of not meeting guidelines of 5.711 = 1/0.175 can be compared relative to Brazil's odds of not meeting guidelines of 28.412 = 1/0.035. Hence, while efficiency relative to best practice for Brazil with odds of meeting guidelines = 0.035/0.175 = 0.201, this is also the efficiency relative to best practice for Brazil with odds of not meeting guidelines (1/0.175)/ (1/0.035) = 5.711/28.412 = 0.201.

For each country, inversion arising from the symmetry of odds between meeting and not meeting guidelines (or any binary effect) cancel with the inversion of relative performance relative to best practice for the same comparator in maximizing versus minimizing in meeting vs not meeting

Table 3 Efficiency comparison of IMGs odds in school age children

Country	Odds meeting IMGs	Relative efficiency meeting IMGs (output framing)	Odds not meeting IMG	Relative efficiency not meeting IMGs (input framing)
Australia	0.17509 = 0.149/0.851	1.000 = 0.17509/0.17509	5.711 = 0.851/0.149	1.000 = 5.711/5.711
Brazil	0.03520 = 0.034/0.966	0.201 = 0.03520/0.17509	28.412 = 0.966/0.034	0.201 = 5.711/28.412
Canada	0.16279 = 0.140/0.860	0.930 = 0.16279/0.17509	6.143 = 0.860/0.140	0.930 = 5.711/6.143
China	0.01523 = 0.015/0.985	0.087 = 0.01523/0.17509	65.667 = 0.985/0.015	0.087 = 5.711/65.667
Columbia	0.08460 = 0.078/0.928	0.483 = 0.08460/0.17509	11.821 = 0.928/0.078	0.483 = 5.711/11.821
Finland	0.07181 = 0.067/0.933	0.410 = 0.07181/0.17509	13.925 = 0.933/0.067	0.410 = 5.711/13.925
India	0.06383 = 0.060/0.940	0.365 = 0.06383/0.17509	15.667 = 0.941/0.060	0.365 = 5.711/15.667
Kenya	0.06952 = 0.065/0.935	0.397 = 0.06952/0.17509	14.385 = 0.935/0.065	0.397 = 5.711/14.385
Portugal	0.02041 = 0.020/0.980	0.117 = 0.02041/0.17509	49.000 = 0.980/0.020	0.117 = 5.711/49.000
South Africa	0.13507 = 0.119/0.881	0.771 = 0.13507/0.17509	7.403 = 0.881/0.119	0.771 = 5.711/7.403
UK	0.12867 = 0.114/0.886	0.735 = 0.12867/0.17509	7.772 = 0.886/0.114	0.735 = 5.711/7.772
USA	0.02249 = 0.021/0.979	0.123 = 0.02249/0.17509	46.619 = 0.979/0.021	0.123 = 5.711/46.619

guidelines. In general, comparing odds of meeting or not meeting guidelines leads to the same consistent relative performance efficiency estimate, and in each case are meaningfully interpreted for each country as the odds ratio of meeting or not meeting guidelines relative to best practice.

We now show that the consistent comparison of relative performance with symmetric odds metrics also extend advantages over relative risk to consistent indirect comparison and consistent translation of evidence of treatment effect.

Extension of consistent efficiency odds ratio result to indirect comparison

Given efficiency measures such as those in Table 3 are each measured relative to the same best practice peer, and consistently represent the odds ratio for meeting guidelines with alternate framing, they also enable consistent indirect comparison of relative performance as odds ratios (Eckermann et al. 2009).

To illustrate, consider indirect comparison between the UK and Canada from the odds ratio relative efficiency measures in Table 3 relative to the best practice common comparator, Australia, in each case. From the results in Table 3, efficiency measures are consistently estimated with alternate framing of meeting or not meeting a guideline as 0.735 for the UK and 0.930 for Canada. In each case, this represents the odds ratio for meeting or not meeting guidelines relative to Australia as the common best practice comparator. Now, undertaking indirect comparison, the odds ratio for the UK relative to Canada in meeting or not meeting integrated movement guidelines is consistently calculated from the efficiency measures for the UK relative to Australia and Canada relative to Australia as 0.735/0.930 = 0.790. This is the same calculation as if we had calculated the odds ratio for meeting guidelines directly from odds between the UK and Canada (0.12867/0.16279 = 0.790) in Table 2. That is, 0.12867/0.17509)/(0.16279/0.17509) = 0.12867/0.16279 = 0.790, with the odds for common best practice comparator Australia effectively cancelling in indirect comparison of odds ratios for the UK and Canada in the same way regardless of the framing of effect as meeting or not meeting guidelines.

Hence, odds ratios for meeting guidelines and relative to the same best practice peer also enable consistent relative performance comparison between any countries in any indirect comparison. Furthermore, note that given the symmetry of odds and odds ratios with alternate framing, the odds ratio of not meeting guidelines for the UK relative to Canada is simply the inverse, 1/(odds ratio of meeting guidelines), and hence 1/0.790 = 1.265. In general, the odds ratio comparison in Table 2 enables consistent and interpretable comparison of relative performance in meeting or not guidelines with any direct or indirect comparison.

In contrast, employing relative risk as in Table 2, lack of symmetry and inconsistency between meeting or not meeting IMGs leads to very different results in indirect comparisons for relative performance such as the UK relative to Canada with alternate framing. Framed as meeting or not meeting guidelines relative performance for the UK relative to Canada with indirect comparison is estimated as 0.765/0.940 = 0.813 and 0.960/0.990 = 0.970, respectively. This inconsistency reflects the same lack of consistency for each of the UK and Canada relative to Australia as a best practice common comparator. Similarly, for any bilateral country comparison while odds ratios (Table 3) will always be consistent, relative risk comparison (Table 2) is expected to be inconsistent. This advantage is again enabled by the symmetry of odds and odds ratios relative to a common best practice comparator and extend previous indirect comparison advantages of odd ratios for treatment effects (Eckermann et al. 2009).

Extension of odds result to evidence translation

Eckermann et al. (2011) show that for any binary effect, translating evidence of a treatment effect to settings with different baseline risk with relative risk leads to inconsistent estimation of absolute effect differences with alternative framing of the same binary outcome, while use of odds ratios is consistent. We now extend this result to demonstrate that odds relative performance comparison also naturally extends to consistently translating relative treatment effect evidence.

To illustrate, consider for example a hypothetical treatment effect found in school-aged populations where integration of IMGs into early childhood centres had an odds ratio (OR) for meeting guidelines of say 1.25 (or 2), while equivalently the OR for not meeting guidelines would be 0.8 = 1/1.25 (or 0.5). These respective ORs can then be applied to prior odds of meeting or not meeting guidelines in any country where that policy had yet to be adopted, to estimate the posterior odds and absolute population risk impact expected with such policy adoption.

For example, Canada with a prior probability of meeting guidelines of 0.14 (Table 1) and associated odds of meeting guidelines of 0.14/0.86 = 0.1628 or not meeting guidelines of 0.86/0.14 = 6.143 (Table 3). Now, applying Bayes Theorem, the odds with changed practice (posterior odds) in Canada is estimated in translating this evidence by simply applying the treatment OR to the odds of current practice.

In terms of meeting guidelines, applying an OR of 1.25 (or 2) with integration of guidelines into early childhood centres to current (prior) odds of 0.1628 in Canada leads to an expected (posterior) odds of 0.2035 (or 0.3256). This posterior odds can then be simply converted back to probabilities to estimate the incremental effect as a difference in absolute probabilities or risk, noting that in converting the relationship between probability (P) and odds (O), P=O/(1 + O) is equivalent to O=P/(1-P).

Converting the expected odds back to a probability (P=O/(1 + O)) leads to the expected rate of meeting guidelines with this policy of 0.2035/1.2035 = 0.169 in Canada in translating evidence with an OR of 1.25 (or 0.3256/1.3256 = 0.246 if the OR were 2). Hence, integrating IMGs into early childhood centres with associated OR of meeting guidelines of 1.25 (or 2) applied to prior odds of 0.16279 in Canada leads to a 2.9% (or 10.6%) absolute increase in the proportion of the population meeting guidelines, from 0.140 to 0.169 (or 0.246).

Alternatively, undertaking evidence translation in terms of not meeting guidelines, applying an OR of 0.8 (or 0.5) with integration of guidelines into early childhood centres to current (prior) odds of not meeting guidelines of 6.143 leads to an expected (posterior) odds of 0.8 × 6.143 = 4.914 (or 0.5 × 6.143 = 3.0714). Converting this back to a probability (P=O/(1 + O)) leads to an expected rate of meeting guidelines with this policy of 4.914/5.914 = 0.831 (or 3.0714/4.0714 = 0.754) and hence a 2.9% (or 10.6%) absolute population decrease in not meeting guidelines relative to base rate of 0.860.

Importantly this demonstrates that translating evidence with treatment effect expressed as an OR applied to the odds of meeting or not meeting guidelines results in the same consistent estimation of absolute differences. More generally, Eckermann, Coory and Willan (2011) show that while relative risk does not allow consistent evidence translation with alternative framing of binary effects precisely for the conditions where evidence translation is required (there is a treatment effect and baseline risk differs from that in the trial setting), odds ratios always do.

Consistent absolute effect estimates from evidence translation naturally can also be used to consistently estimate incremental costs associated with incremental effects, noting that appropriate cost consideration allows for downstream cost implications of effects as well as direct costs. Hence, odds methods are required for consistent evidence translation of binary effects in estimating absolute incremental effects, cost or cost effectiveness analysis. This in turn is key to undertaking robust clinical or health economic analysis of the expected net clinical benefit and net benefit of a strategy with binary effects data in translating evidence of effects to any given jurisdiction of interest.

Consequently, for any binary effect data, use of odds and odds ratios has been shown to provide a robust method that solves the lack of consistency of alternative framing with risk or probability presentation and resulting relative risk efficiency measures relative to best practice, but also more generally extends interpretability of relative performance measures to indirect comparison or translating a treatment effect to any jurisdiction of interest. Establishing odds as the appropriate metric to enable consistent analysis in comparison of binary effect data such as meeting or not meeting guidelines is a primary consideration in satisfying comparability for robust decision making.

Discussion

Top-down and bottom-up methods for robust IMGs evaluation

A robust method has been identified in this paper for consistently comparing jurisdiction's top-down population level efficiency comparisons relative to best practice for any binary effect measure, such as populations meeting or not meeting public health guidelines.

Developing such robust top-down methods is particularly important to evaluating health promotion and disease prevention policies and strategies, such as those aimed at improving 24-h movement behaviours in community settings given current limitations of conventional bottom-up methods. Trial-based bottom-up evaluation methods for estimating effects and downstream costs currently do not allow for network and multiplier effects across populations over time and hence community population level potential of such strategies (Shiell and Hawe 1996; Shiell et al. 2008; Moore et al. 2006; Hawe and Ghali 2008; Hawe et al. 2009; Eckermann et al. 2014; Eckermann and McCaffrey 2017). Therefore, community level potential for expanded scope and duration of population effects with community ownership are not covered or captured by conventional bottom-up methods, but are naturally covered with monitoring of health promotion strategies with top-down population level analysis over time. Such top down analysis can be used to monitor population, community level, effects of alternative health promotion strategies over time as well as associated downstream costs savings in assessing their cost effectiveness.

Consequently, robust top down analysis is key to enabling adequate coverage in identifying effective and cost effective health promotion strategies and addressing chronic underfunding of health promotion strategies internationally (Shiell and Jackson 2018; Eckermann 2017). Such underfunding in turn reflects a lack of profit motivated vested interests in supporting research for health promotion strategies such as IMGs which are either not patentable or have fewer apparent commercialization opportunities (Pekarsky 2012, 2015; Eckermann and Pekarsky 2014).

In evaluation of health promotion strategies in community settings, the robust population-level efficiency comparisons relative to best practice can also consistently, as shown in evidence translation results, combine with bottom-up evaluation methods as they emerge (Zaza et al. 2005). Monitoring compliance with 24-h movement guidelines in early childhood using guidelines launched in Canada and Australia in 2017 (Okely et al. 2017; Tremblay et al. 2016) and by WHO in 2019 (WHO 2019) make these methods particularly timely. These measures are important to identifying and evaluating relative efficiency of health promotion policies and strategies across health systems internationally. They can be used to highlight best relative performance in practice, the policies and practices associated with best practice, and the potential for, and factors associated with, improving policy and practice across suboptimal jurisdictions.

Ideally this would coincide with developing robust methods for bottom-up evaluation of health promotion approaches, environmental and whole of government strategies, and policies and practices aimed at improving compliance with guidelines. These strategies include, but are not limited to: the integration of guidelines into educational settings; building networks of walking and bike paths; creating public gardens, green spaces and public places; whole of government policy support across education, transport, parks and recreation, sport, and urban planning; and the running of culturally appropriate activities and programs (WHO 2017; Okely et al. 2017; Tremblay et al. 2016).

Nevertheless, distinct challenges remain for robust bottom-up evaluations that have yet to be jointly addressed. Those challenges relate to the interrelated time constrained nature of 24-h movement behaviours across physical activity, sedentary behaviour and sleep domains but also multiple domains of health effects that integrated movement behaviours map to and across community populations with multiple alternative health promotion and environmental strategies, modalities and pathways. Robust evaluation methods need to jointly consider physical activity, sedentary and sleep movement behaviours constrained by the 24 h day in identifying optimal population behaviours with compositional data analysis methods (Dumuid et al. 2018a, b) in mapping to multiple health effects. Such mapping to multiple health effects then need to combine with best evaluation methods for multiple strategy and joint multiple domain of effect cost effectiveness and efficiency comparisons (Eckermann et al. 2008; Eckermann and Willan 2011; Eckermann and Coelli; McCaffrey et al. 2015).

Finally, in evaluating expected population level effects and cost implications across alternative health promotion and environmental strategies, methods should allow for community ownership, network and multiplier effects (Eckermann et al. 2014; Shiell et al. 2008; Hawe et al. 2009). In this respect, it should be noted that environmental and whole of government strategies (e.g. improving community walkability and access to public recreation spaces) are not restricted to single target populations in their scope of expected community-level effects. For example, the same types of environmental and community accessible policies aimed at supporting children meeting movement guidelines have been shown to have health and lifestyle benefits for aged populations in being physically, mentally and socially active (WHO 2002, 2007; Kalache 2013; Eckermann 2017 chapter 12; Dyer et al. 2018; Eckermann et al. 2019). Such analyses point to whole of government environmental strategies for active communities providing potential multiplier benefits in all community populations across the lifecycle. Importantly, this implies that environmental strategies to promote movement

behaviours for children must ensure community ownership to have high potential multiplier effects. This potential is both in terms of population scope with network effects across populations of all ages in the community as well as duration of health effects with improving trajectories of integrated movement behaviours, lifestyle and habits.

Developing and applying robust methods for bottom-up as well as the robust top-down evaluation of population level effects with consistent odds method for evaluating the movement guidelines identified in this paper are important to supporting appropriate research and investment in health promotion strategies such as those for IMGs. However, the consistent odds ratio method for monitoring populations IMG compliance and comparing countries strategies in practice identified in this paper is key to robustly enabling adequate coverage in assessing effectiveness or cost effectiveness of strategies and policies both currently and in combining with such robust bottom-up methods as they develop.

More generally, the finding of consistent odds ratio relative performance efficiency comparison for binary effects in this paper have also been illustrated to extend to avoiding inconsistency with evidence translation and indirect comparison that arise with relative risk comparison.

Conclusion

We have shown that the use of odds overcomes inconsistency between input and output population level efficiency measures relative to best practice that arises for binary effects such as meeting or not meeting guidelines with use of probability or proportions. This finding for any binary effect results from the symmetrical nature of odds and odds ratios and their inversion with relative efficiency measures in maximising output (e.g. meeting guidelines) or minimising input (e.g. not meeting guidelines) relative to the same best practice comparator. Noting that binary effects can always be expressed in one of two ways (e.g. meeting or not meeting guidelines; success or failure; survival or mortality) this finding is key to consistent comparisons of population evidence with any binary effect data.

Furthermore, by employing odds for comparison, efficiency measures have been shown to represent odds ratios for meeting guidelines relative to best practice and as such are also highly interpretable. These findings overcome relative risk inconsistencies and fallacies with alternative framing of binary effects and extend distinct advantages previously shown for odds ratios over relative risk in indirect comparison and evidence translation (Eckermann et al. 2009, Eckermann et al. 2011; Eckermann 2017:65-75).

The odds method identified for consistent efficiency measures with binary effects has been highlighted as particularly important to aid top-down population evaluation in practice of health promotion and disease prevention programs, policies and strategies such as those for meeting IMGs. This is particularly the case given inadequacy of funding for such non-patentable or not easily commercialisable public health strategies (Pekarsky 2012, 2015; Eckermann and Pekarsky 2014; Eckermann 2017; Shiell and Jackson 2018) and bottom-up methods which currently fail to cover the network and multiplier potential over time of such community-level strategies where they have community ownership (Shiell and Hawe 1996; Shiell et al. 2008; Hawe and Ghali 2008; Hawe et al. 2009; Eckermann et al. 2014; Eckermann and McCaffrey 2017).

Declarations

References

Dumuid D, Stanford TE, Martin-Fernandez JA, Pedišić Z, Maher C, Lewis LK, Hron K, Katzmarzyk PT, Chaput J-P, Fogelholm M, Hu G, Lambert EV, Maia J, Sarmiento OL, Standage M, Barreira TV, Broyles ST, Tudor-Locke C, Tremblay MS, Olds T (2018a) Compositional data analysis for physical activity, sedentary time and sleep research. Stat Methods Med Res 27(12):3726–3738. https://doi.org/10.1177/0962280217710835

Dumuid D, Stanford TE, Pedišić Z, Maher C, Lewis L, Martín-Fernández J-A, Katzmarzyk PT, Chaput J, Fogelholm M, Standage M, Tremblay MS, Olds T (2018b) Adiposity and the isotemporal substitution of physical activity, sedentary time and sleep amongst school age children: a compositional data analysis approach. BMC Public Health 18:311. https://doi.org/10.1186/s12889-018-5207-1

Dyer SM, Liu E, Gnanamanickam ES, Milte R, Easton T, Harrison SL, Bradley CE, Ratcliffe J, Crotty M (2018) Clustered domestic residential aged care in Australia: fewer hospitalisations and better quality of life. Med J Aust 208(10):433–438. https://doi.org/10.5694/mja17.00861

Eckermann S (2017) Health economics from theory to practice: optimally informing joint decisions of research, reimbursement and regulation with health system budget constraints and community objectives. Springer, Cham. https://doi.org/10.1007/978-3-319-50613-5

Eckermann S, Briggs A, Willan A (2008) Health technology assessment in the cost-disutility plane. Med Decis Mak 28:172–181. https://doi.org/10.1177/0272989X07312474

Eckermann S, Coelli T (2013) Including quality attributes in efficiency measures consistent with net benefit: creating incentives for evidence based medicine in practice. Soc Sci Med 76:159–168. https://doi.org/10.1016/j.socscimed.2012.10.020

Eckermann S, Coory M, Willan AR (2011) Consistently estimating absolute risk difference when translating evidence to jurisdictions of interest. PharmacoEconomics 29(2):87–96. https://doi.org/10.2165/11585910-000000000-00000

Eckermann S, Coory M, Willan AR (2009) indirect comparison: relative risk fallacies and odds solution. J Clin Epidemiol 62:1031–1036. https://doi.org/10.1016/j.jclinepi.2008.10.013

Eckermann S, Dawber J, Yeatman H, Quinsey K, Morris D (2014) Evaluating return on investment in a school based health promotion and prevention program: the investment multiplier for the Stephanie Alexander kitchen garden National Program. Social Sci Med 114:103–112. https://doi.org/10.1016/j.socscimed.2014.05.056

Eckermann S, Karnon J, Willan AR (2010) The value of value of information: best informing research design and prioritization using current methods. PharmacoEconomics 28(9):699–709. https://doi.org/10.2165/11537370-000000000-00000

Eckermann S, McCaffrey N (2017) Beyond the individual: evaluating community-based health promotion and prevention strategies and palliative care. In: Eckermann S (ed) Health economics from theory to practice. Springer, Cham, pp 91–106. https://doi.org/10.1007/978-3-319-50613-5_4

Eckermann S, Pekarsky B (2014) Can the real opportunity cost stand up: displaced services the straw man outside the room. PharmacoEconomics 32(4):319–325. https://doi.org/10.1007/s40273-014-0140-3

Eckermann S, Phillipson L, Fleming R (2019) Reform and re-design of aged care environments are key to improved care quality and cost effective aged and health system care. Appl Health Econ Health Policy 17:127–130. https://doi.org/10.1007/s40258-018-0435-1

Eckermann S, Willan AR (2011) Presenting evidence and summary measures to best inform societal decisions when comparing multiple strategies. PharmacoEconomics 29(7):563–577. https://doi.org/10.2165/11587100-000000000-00000

Hawe P, Ghali L (2008) Use of social network analysis to map the social relationships of staff and teachers at school. Health Educ. Res 23(1):62–69. https://doi.org/10.1093/her/cyl162

Hawe P, Shiell A, Riley T (2009) Theorising interventions as events in systems. Am. J. Community Psychol 43:267–276. https://doi.org/10.1007/s10464-009-9229-9

Katzmarzyk PT, Barreira TV, Broyles ST, Champagne CM, Chaput J, Fogelholm M, Hu G, Johnson WD, Kuriyan R, Kurpad A, Lambert EV, Maher C, Maia J, Matsudo V, Olds T, Onywera V, Sarmiento OL, Standage M, Tremblay MS, Tudor-Locke C, Zhao P, Church TS (2013) The international study of childhood obesity, lifestyle and the environment (ISCOLE): design and methods. BMC Public Health 13:900 https://bmcpublichealth.biomedcentral.com/articles/10.1186/1471-2458-13-900

Kalache A (2013) The longevity revolution: creating a society for all ages. Adelaide thinker in residence 201-2013. Adelaide: government of South Australia. https://www.dunstan.org.au/wp-content/uploads/2018/12/Kalache_Report_compressed.pdf

McCaffrey N, Agar M, Harlum J, Karnon J, Currow D, Eckermann S (2015) Better informing decision making with multiple outcomes cost-effectiveness analysis under uncertainty in cost-disutility space. PLoS ONE 10(3):e0115544. https://doi.org/10.1371/journal.pone.0115544

Moore S, Haines V, Hawe P, Shiell A (2006) Lost in translation: a genealogy of the "social capital" concept in public health. J Epidemiol Comm Health 60(8):729–734. https://doi.org/10.1136/jech.2005.041848

O'Brien B (1996) Economic evaluation of pharmaceuticals: Frankenstein's monster or vampire of trials? Med Care 34(12 Suppl):DS99–D108

Okely A, Ghersi D, Hesketh KD, Santos R, Loughran SP, Cliff DP, Shilton T, Grant D, Jones RA, Stanley RM, Sherring J, Hinkley T, Trost SG, McHugh C, Eckermann S, Thorpe K, Waters K, Olds TS, Mackey T, Livingstone R, Christian H, Carr H, Verrender A, Pereira JR, Zhang Z, Downing KL, Tremblay MS (2017) A collaborative approach to adopting/adapting guidelines. The Australian 24-hour movement guidelines for the early years (birth to 5 years): an integration of physical activity, sedentary behaviour, and sleep. BMC Public Health 17(Suppl 5):869. https://doi.org/10.1186/s12889-017-4867-6

Pekarsky BAK (2015) The new drug reimbursement game: a regulator's guide to playing and wining. Springer, Cham. https://www.springer.com/gp/book/9783319089027

Pekarsky BAK (2012) Trusts, constraints and the counterfactual: reframing the political economy of new drugs. Dissertation, University of Adelaide. https://digital.library.adelaide.edu.au/dspace/handle/2440/79171

Roman-Viñas B, Chaput J-P, Katzmarzyk PT, Fogelholm M, Lambert EV, Maher C, Maia J, Olds T, Onywera V, Sarmiento OL, Standage M, Tudor-Locke C, Tremblay MS for the ISCOLE Research Group (2016) Proportion of children meeting recommendations for 24-hour movement guidelines and associations with adiposity in a 12-country study. Int J Behav Nutr Phys Act 13:123. https://doi.org/10.1186/s12966-016-0449-8

Saunders TJ, Gray CE, Poitras VJ, Chaput J-P, Janssen I, Katzmarzyk PT, Olds T, Gorber SC, Kho ME, Sampson M, Tremblay MS, Carson V (2016) Combinations of physical activity, sedentary behaviour and sleep: relationships with health indicators in school-aged children and youth. Appl Physiol, Nutr, Metab 41(6 suppl.3):S283–S293. https://doi.org/10.1139/apnm-2015-0626

Shiell A, Hawe P (1996) Health promotion community development and the tyranny of individualism. Health Econ 5(3):241–247. https://doi.org/10.1002/(SICI)1099-1050(199605)5:3<241::AID-HEC197>3.0.CO;2-G

Shiell A, Hawe P, Gold L (2008) Complex interventions or complex systems? Implications for health economic evaluation. Br Med J

336(7656):1281–1283. https://doi.org/10.1136/bmj.39569.510521.AD

Shiell A, Jackson H (2018) How much does Australia spend on prevention and how would we know whether it is enough? Health Promot J Austr 29(Suppl 1):7–9. https://doi.org/10.1002/hpja.165

Tremblay MS, Carson V, Chaput JP, Gorber SC, Dinh T, Duggan M, Faulkner G, Gray CE, Gruber R, Janson K, Janssen I, Katzmarzyk PT, Kho ME et al (2016) Canadian 24-hour movement guidelines for children and youth: an integration of physical activity, sedentary behaviour, and sleep. Appl Physiol, Nutr, Metab 41(6 suppl.3): S311–S327. https://doi.org/10.1139/apnm-2016-0151

Walsh JJ, Barnes JD, Cameron JD, Goldfield GS, Chaput J-P, Gunnell KE, Ledoux A-A, Zemek RL, Tremblay MS (2018) Associations between 24-hour movement behaviours and global cognition in US children: a cross-sectional observational study. Lancet Child Adolesc Health 2(11):783–791. https://doi.org/10.1016/S2352-4642(18)30278-5

World Health Organisation (2019) Guidelines on physical activity, sedentary behaviour and sleep for children under 5 years of age. Geneva: World Health Organization. Licence: CC BY-NC-SA 3.0 IGO. https://apps.who.int/iris/handle/10665/311664

World Health Organisation (2017) Best buys and other recommended interventions for the prevention and control of noncommunicable diseases updated (2017), appendix 3 of the global action plan. http://www.who.int/ncds/management/WHO_Appendix_BestBuys.pdf

World Health Organization (2007) Global age-friendly cities, a guide. Geneva. https://www.who.int/ageing/publications/Global_age_friendly_cities_Guide_English.pdf

World Health Organization (2002) Active ageing: a policy framework. WHO/NMH/NPH/02.8 Geneva. https://www.who.int/ageing/publications/active_ageing/en/

Yeatman H, Quinsey K, Dawber J, Nielsen W, Condon-Paoloni D, Eckermann S, Morris D, Grootemaat P, Fildes D (2014) Combining realism with rigour: evaluation of a National Kitchen Garden program in Australia primary schools. Eval J Australasia 14(2):17–24. https://doi.org/10.1177/1035719X1401400203

Zaza S, Briss PA, Harris KW, Task Force on Community Preventive Services (2005) The guide to community preventive services: what works promote health? Oxford University Press, Oxford. https://doi.org/10.1093/acprof:oso/9780195151091.001.0001