Alignment of systematic reviews published in the Cochrane Database of Systematic Reviews and the Database of Abstracts and Reviews of Effectiveness with global burden-of-disease data: a bibliographic analysis

Sze Lin Yoong, Alix Hall, Christopher M Williams, Eliza Skelton, Christopher Oldmeadow, John Wiggers, Chante Karimkhani, Lindsay N Boyers, Robert P Dellavalle, John Hilton, Luke Wolfenden

ABSTRACT

Background Systematic reviews of high-quality evidence are used to inform policy and practice. To improve community health, the production of such reviews should align with burden of disease. This study aims to assess if the volume of research output from systematic reviews proportionally aligns with burden of disease assessed using percentages of mortality and disability-adjusted life years (DALYs).

Methods A cross-sectional audit of reviews published between January 2012 and August 2013 in the Cochrane Database of Systematic Reviews (CDSR) and Database of Abstracts of Reviews of Effects (DARE) was undertaken. Percentages of mortality and DALYs were obtained from the 2010 Global Burden of Disease study. Standardised residual differences (SRD) based on percentages of mortality and DALYs were calculated, where conditions with SRD of more than or less than three were considered overstudied or understudied, respectively.

Results 1029 reviews from CDSR and 1928 reviews from DARE were examined. There was a significant correlation between percentage DALYs and systematic reviews published in CDSR and DARE databases (CDSR: r=0.68, p=0.001; DARE: r=0.60, p<0.001). There was no significant correlation between percentage mortality and number of systematic reviews published in either database (CDSR: r=0.34, p=0.14; DARE: r=0.22, p=0.34). Relative to percentage of mortality, mental and behavioural disorders, musculoskeletal conditions and other non-communicable diseases were overstudied. Maternal disorders were overstudied relative to percentages of mortality and DALYs in CDSR.

Conclusions The focus of systematic reviews is moderately correlated with DALYs. A number of conditions may be overstudied relative to percentage of mortality particularly in the context of health and medical reviews.

INTRODUCTION

The production of intervention research is key to informing efforts at improving population health. Randomised controlled trials (RCTs) are considered the most internally valid study design and provide the strongest evidence of causal relationships between intervention (exposure) and outcome(s) examined. Observational study designs including cohort and case–control studies also provide useful information particularly in circumstances where it is not feasible to undertake RCTs, such as when the outcome is rare (eg, adverse events), or when long-term follow-up is needed. Systematic reviews of RCTs represent the highest quality study design in evidence hierarchies. By integrating and providing global estimates of the available evidence in context of the broader literature, systematic reviews overcome some of the shortcomings of individual studies including systematic bias and statistical imprecision. While there are some limitations with systematic reviews that can occur as a result of inadequate methodology (eg, poor search strategy) and potential challenges with generalisability and applicability of results, these reviews are fundamental tools for identifying gaps in the literature, informing evidence-based policy formulation and clinical and public health practice. Importantly, policymakers increasingly use systematic reviews as the preferred source of evidence to guide decision-making.

To maximise the benefits of investments in research activity to improve the health and social well-being of the community globally, international organisations including the Institute of Medicine have recommended that government-funded organisations prioritise research activities on the basis of disease burden. Burden of disease is a measure used to assess and compare the relative impact of different diseases and injuries on population health. Mortality and morbidity (assessed using disability-adjusted life years (DALYs)) are frequently used indices of burden of disease. Total mortality is defined as the number of deaths due to a specific cause (disease or injury) within the total population. DALYs take into account both years of life lost from premature death and years of life lived with a disability from a condition, providing a measurement of the discrepancy between existing and ideal health status. Given the role of systematic reviews in facilitating evidence-based health policy and practice, the production of systematic reviews that focus on conditions contributing to greatest health burden is critical for community health improvement. Key producers of systematic reviews should align with burden of disease.
reviews including Cochrane,12 Johanna Briggs Institute13 and the Campbell collaboration14 have similarly recognised this as an important consideration in generating evidence to best meet global health needs.

Previous studies examining the focus of systematic reviews have found only a modest correlation between the production of research and burden of disease.15 16 A recent study (2014) examining the representation of specific skin conditions within the Cochrane Database of Systematic reviews (CDSR) found that review output and DALYs were well matched for 5 of the 15 skin conditions examined.17 In 2003, an audit of the focus of systematic reviews more broadly found moderate correlation between the number of systematic reviews published in the Cochrane database and the 1990 estimates of DALYs from the Global Burden of Disease (GBD) study (r=0.54).15 This study only described the correlation between the number of reviews and DALYs data and did not attempt to identify specific conditions which have been understudied or overstudied relative to DALYs. Recent estimates from the 2010 GBD study have identified substantial shifts in the leading causes of mortality and DALYs, globally, since 1990. Non-communicable diseases account for 63.5% of mortality and 54% of DALYs in 2010 compared with 34% mortality and 44% DALYs in 1990.18 19 20

Given such changes, there is a lack of recent evidence assessing the extent to which systematic review output aligns with international estimates of burden of disease. Therefore, this study aimed to (1) examine the correlation between percentage of reviews published in two international databases: CDSR and Database of Abstracts of Review of Effects (DARE) with GBD estimates (mortality and DALYs) from the GBD 2010 study, (2) identify conditions that were overstudied and understudied relative to global mortality and DALYs and (3) assess discrepancies between the databases. CDSR and DARE were chosen because they employ stringent quality measures to ensure that included reviews are of high rigour and are thus recognised as the primary publishers of high-quality systematic reviews in health and medical research.

METHODS

Data source
All authors, titles and abstracts of full reviews published in CDSR and DARE between January 2012 and August 2013 were extracted. These dates were chosen to allow for potential diffusion of these new estimates for the GBD 2010 study. Both CDSR and DARE include systematic reviews that evaluate the effects of healthcare interventions and delivery of health services. The production of systematic reviews in CDSR is carried out by 53 Cochrane Review Groups, with over 26 000 contributors from over 120 countries.21 DARE includes systematic reviews not carried out by Cochrane. Reviews published in the broader literature rated as high quality and of sufficient methodological rigour are selected and assembled by the Centre for Reviews and Dissemination at the University of York, UK.

Data extraction

Conditions/diseases examined
The types of conditions and diseases examined were coded according to the classifications used in the GBD 2010 study. The study is a systematic and epidemiological dataset produced by the Institute of Health Metrics and Evaluation since 1990, and provides ongoing estimates of mortality and morbidity for 291 diseases from 187 developed and developing countries.22 Consistent with the GBD 2010 classifications,22 the focus of the reviews was grouped as communicable, non-communicable, injury-related and other (for conditions not included in the taxonomy or not specified). Using a data extraction form, one reviewer (ES) with a PhD in behavioural sciences and more than 5 years of research and/or health promotion experience, a second reviewer (SLY) checked coding of all diseases, and a consensus process was used to resolve any discrepancies.

<table>
<thead>
<tr>
<th>Disease</th>
<th>CDSR</th>
<th>DARE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Disease</td>
<td>Number of publications*</td>
<td>Per cent</td>
</tr>
<tr>
<td>Communicable</td>
<td>199</td>
<td>22</td>
</tr>
<tr>
<td>HIV, AIDS and tuberculosis</td>
<td>21</td>
<td>2.4</td>
</tr>
<tr>
<td>Diarrhoea, lower respiratory infections, meningitis and other common infectious diseases</td>
<td>64</td>
<td>7.2</td>
</tr>
<tr>
<td>Neglected tropical diseases and malaria</td>
<td>11</td>
<td>1.2</td>
</tr>
<tr>
<td>Maternal disorders</td>
<td>34</td>
<td>3.8</td>
</tr>
<tr>
<td>Neonatal disorders</td>
<td>52</td>
<td>5.8</td>
</tr>
<tr>
<td>Nutritional deficiencies</td>
<td>8</td>
<td>0.9</td>
</tr>
<tr>
<td>Other communicable</td>
<td>9</td>
<td>1.0</td>
</tr>
<tr>
<td>Non-communicable</td>
<td>638</td>
<td>71</td>
</tr>
<tr>
<td>Neoplasms</td>
<td>104</td>
<td>12</td>
</tr>
<tr>
<td>Cardiovascular and circulatory diseases</td>
<td>93</td>
<td>10</td>
</tr>
<tr>
<td>Chronic respiratory diseases</td>
<td>46</td>
<td>5.2</td>
</tr>
<tr>
<td>Cirrhosis of the liver</td>
<td>9</td>
<td>1.0</td>
</tr>
<tr>
<td>Digestive diseases</td>
<td>32</td>
<td>3.6</td>
</tr>
<tr>
<td>Neurological disorders</td>
<td>52</td>
<td>5.8</td>
</tr>
<tr>
<td>Mental and behavioural disorders</td>
<td>73</td>
<td>8.2</td>
</tr>
<tr>
<td>Diabetes, urogenital, blood and endocrine disorders</td>
<td>86</td>
<td>9.6</td>
</tr>
<tr>
<td>Musculoskeletal disorders</td>
<td>39</td>
<td>4.4</td>
</tr>
<tr>
<td>Other non-communicable</td>
<td>147</td>
<td>16</td>
</tr>
<tr>
<td>Injury</td>
<td>57</td>
<td>6.4</td>
</tr>
<tr>
<td>Transport injuries</td>
<td>4</td>
<td>0.5</td>
</tr>
<tr>
<td>Unintentional injuries</td>
<td>51</td>
<td>5.7</td>
</tr>
<tr>
<td>Self-harm and interpersonal violence</td>
<td>2</td>
<td>0.2</td>
</tr>
<tr>
<td>War and disaster</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

*Total number more than sum of reviews due to some reviews examining multiple conditions.

CDSR, Cochrane Database of Systematic Review; DARE, Database of Abstracts and Reviews of Effectiveness.
Burden of disease assessed using mortality and DALYs
Total mortality rates and DALYs from the GBD 2010 were used as a measure of burden of disease, as recommended by WHO.\textsuperscript{10} Mortality and DALY metrics for each of the 21 conditions examined in the GBD 2010 study were obtained using the GBD interactive time plot tool using the search parameters ‘time plot’, ‘DALYs’ metric, ‘global’ place, ‘all ages’, ‘both’ sexes, and ‘%’ units for each condition.\textsuperscript{20} The methods used to generate mortality and DALY statistics in the GBD 2010 estimates have been previously described.\textsuperscript{19}

Statistics
All data analyses were conducted using SAS (V9.4, SAS Institute, Cary, North Carolina, USA).\textsuperscript{24} Frequencies and percentages of studies were generated for each condition examined in CDSR and DARE. As the data were not normally distributed, the correlation between the ranking order of the number of reviews focusing on each disease condition published in CDSR and DARE, and percentages of mortality and DALYs, was estimated using Spearman’s rank correlation coefficient. To assess if the distribution of publications was consistent with the distribution of number of deaths and DALYs for each condition, the observed (O) number of publications for each condition was compared with the expected (E) number of publications based on the proportion of death and DALYs, and assessed for goodness of fit using the Pearson $\chi^2$ test. The E count for each condition was calculated by multiplying the total number of publications by the corresponding proportion of deaths or DALYs for that condition. Standardised residual differences (O−E)/E were displayed using a bar chart, where positive numbers on the bar chart indicate conditions where the number of observed studies was greater than the number of expected studies (overstudied), and negative numbers indicate conditions where the number of expected studies was smaller than the number of observed studies (understudied). To identify conditions that were disproportionally studied, conditions with standardised residual differences (SRD) exceeding 3 ($or$−3) were considered statistically significantly overstudied or understudied. This threshold was chosen as it is accounted for multiple testing using a conservative Bonferroni adjustment. To assess whether the discrepancies in observed and expected number of publications were different between the two databases, SRDs from each condition were compared using the paired $t$ test.

RESULTS
Overall, 1029 reviews from CDSR and 1928 reviews from DARE were published between January 2012 and August 2013. One hundred and forty-eight reviews (14%) from CDSR and 284 (14%) reviews from DARE did not examine a condition included in the taxonomy and were excluded from the subsequent analysis. Only 56 (5.4%) publications in CDSR and 4.5% of publications in DARE contributed to multiple categories.

Number and percentage of studies assessing each condition in CDSR and DARE
The majority of reviews published in CDSR (71.4%) and DARE (81.5%) examined non-communicable diseases (table 1). The most frequently examined conditions for both databases were neoplasms (11.6% CDSR; 21.0% DARE) and cardiovascular and circulatory diseases (10% CDSR; 12% DARE). There was a significant correlation between the number of publications in CDSR and DALYs ($r=0.68, p=0.001$) and DARE and DALYs ($r=0.60, p<0.001$). No significant correlation was observed between number of reviews and percentage of mortality for CDSR ($r=0.34, p=0.14$) or DARE ($r=0.22, p=0.34$).

Figures 1 and 2 show the SRDs of the reviews published in CDSR and DARE. For reviews published in the CDSR (figure 1), overstudied conditions (those with a SRD $>3$) relative to percentage of mortality, were mental and behavioural disorders (SRD=16.7), musculoskeletal conditions (SRD=13.4), other non-communicable disease (SRD=10.5) and maternal disorders (SRD=6.6). Maternal disorders were also overstudied relative to percentage of DALYs (SRD=4.58).

For DARE (figure 2), the conditions overstudied relative to percentage of mortality were musculoskeletal (SRD=36.9), mental and behavioural diseases (SRD=16.0) and other non-communicable disease (SRD=4.2). None of the conditions were understudied relative to percentage of DALYs.

When comparing SRDs between the number of expected minus observed studies between CDSR and DARE for each condition, no significant differences between the databases for DALYs ($p=0.5777$) or mortality ($p=0.6592$) were found.

DISCUSSION
Using burden-of-disease metrics from the GBD 2010 study, this study found a significant correlation between DALYs and systematic reviews published in CDSR and DARE in 2012–2013. There was no significant correlation between percentage of mortality and the numbers of systematic reviews published in either database. Mental and behavioural disorders, musculoskeletal conditions and other non-communicable diseases were overstudied relative to percentage of global mortality. Collectively, the study suggests that, for most conditions, production of systematic reviews is well aligned with burden-of-disease estimates of DALYs.

A previous study published by Swingler in 2003 found modest correlation ($r=0.54$) between the number of reviews published in CDSR and percentage of DALYs, using estimates from the GBD 1990. The primary causes of DALYs have changed considerably since publication of the 1990 estimates. In 2010, the leading causes of DALYs, globally, were ischaemic heart disease (fourth in 1990, an increase of 29%), followed by lower respiratory infections (first in 1990, a decrease of 44%), stroke (fifth in 1990; an increase of 19%), diarrhoeal diseases (second in 1990; a decrease of 51%), and HIV/AIDS (3.3rd in 1990; an increase of 351%).\textsuperscript{19} Despite such changes, a similar moderate association between reviews published in CDSR and percentage of DALYs was identified in the current study ($r=0.68$).\textsuperscript{15}

This study is the first to examine the systematic review output and percentage of mortality, and has found that the number of publications in CDSR and DARE did not correlate with the total percentage of mortality. Such findings are consistent with studies examining association between research funding allocation, where positive correlations with percentage of mortality have not been consistently reported.\textsuperscript{8 15–27}

The higher than expected number of published reviews focusing on mental and behavioural disorders, musculoskeletal disorders and other non-communicable disease relative to mortality may reflect the complexity of these conditions and interventions needed for management and prevention. Relative to other conditions which may be more suited to surgical or pharmacological intervention, a broader and more heterogeneous range of strategies may need to be tested to address behavioural and other non-communicable diseases given the complex aetiology of these conditions. Further, as the effectiveness of these
interventions is typically dependent on a range of contextual factors, the need to understand the specific circumstances in which an intervention impacts on health outcomes may have resulted in larger numbers of primary trials and reviews. The specific diseases examined within these conditions are also likely to be the focus of research investment for many higher income countries, and are priority research areas for government and philanthropic sources. As research funding is correlated with publication output, this is likely to result in an increased number of primary studies and reviews examining these conditions.

For CDSR, specifically, maternal disorders were categorised as overstudied relative to both mortality and DALYs. These findings are consistent with the study by Swingler, which reported that CDSR had a better coverage of maternal issues compared with DARE in 2000. This could reflect the presence of more established Cochrane Review Groups that support the production of systematic reviews in the area of maternal health. The higher than expected number of reviews in this area could also reflect targeted collaboration with international organisations to achieve specific outcomes as part of the prioritisation process. For example, the Cochrane Pregnancy and Childbirth Review Group worked in collaboration with WHO to update their recommendations for the prevention and treatment of postpartum haemorrhage guidelines in 2012/2013. Findings from our study suggests that a large number of reviews are produced in this area, and efforts to streamline production of reviews so that they address issues most relevant to areas of need in developed countries may be needed in order to achieve the millennium development goals by 2015.

Strengths and weaknesses of the study
There are a number of limitations that need to be considered in the interpretation of the study findings. First, although useful, the burden-of-disease estimates including percentage of mortality and DALYs represent just one source of data available to prioritise systematic review activity. Other factors, including the availability of primary studies for review, the number and type of treatment options for a given condition or population, disease trends, emerging evidence that may necessitate production of new reviews, public interest and the presence of disease types which have been established for a longer period of time are among other important considerations that may impact on systematic review output, and are not currently addressed in this manuscript. Second, owing to feasibility constraints, classification was performed using review titles and abstracts only. Examination of full texts may have improved the accuracy of study classification. Further, the inclusion of only reviews

Figure 1 The standardised residual difference (observed minus expected) for each cause of mortality and disability published in the Cochrane Database of Systematic Reviews in 2012/2013.
published between January 2012 and August 2013 may not reflect all reviews available in CDSR and DARE. Third, the data collection method of classifying reviews into a particular condition involved some degree of subjectivity. Fourth, this study was also not powered to specifically identify overstudied or understudied conditions. While we did not identify any understudied conditions, this could have been due to the insufficient number of reviews included in this study. Finally, this study does not take into account the scope of the included review. For example, some authors may include multiple intervention strategies for a condition within one review (lumping), whereas other authors may include only a single intervention in one review, often resulting in multiple reviews for the same condition (splitting).

Unanswered questions and future research

Our study highlighted that some disease categories may be overstudies in systematic reviews relative to percentages of mortality and DALYs. A closer examination of the specific conditions examined within each broad category of disease is warranted, particularly for conditions identified as overstudied, as this could highlight particular areas for research attention and provide an understanding of potential reasons for the disproportionately larger number of reviews for these conditions. For example, an examination of Cochrane and non-Cochrane reviews in the maternal health area found that while there were a substantial number of reviews overall, some gaps still existed, particularly for developing countries. Further, the production of systematic reviews represents only one proxy of research output and usage. The examination of other indicators of output and research usage (eg, citation rates) with burden-of-disease indicators may provide a more in-depth understanding of how well research production and use aligns more broadly with burden of disease. Despite its importance, little is known regarding the usefulness of strategies to prioritise production of systematic reviews to best inform practice and policy. Initiatives undertaken by the Cochrane collaboration to improve prioritisation of systematic reviews, including application of an equity lens to inform agenda setting and research decisions, engagement with relevant stakeholders and policymakers, establishing formal prioritisation partnerships with the James Lind Alliance, and development of priority review topics are commendable. Such efforts, however, are often varied and largely driven by individual review groups. While the current study did not identify any difference between CDSR and DARE in terms of number of reviews and alignment with mortality and DALYs, more formal evaluation of the impact of such initiatives is warranted and will inform our understanding on how to best prioritise research topics examined in systematic reviews.

CONCLUSIONS

The production of systematic reviews in CDSR and DARE moderately aligns with DALYs but not mortality. This study found that mental and behavioural disorders, musculoskeletal conditions and other non-communicable diseases were overstudied
Department of Veterans Affairs and grants from the Centers for Disease Control and Prevention (CDC 3U48DP001938-0451 (PI:Crane), and the National Institutes of Health (NCI R21 CA173654 (PI:Walkocz). JH is an employee of The Cochrane Collaboration which produces the Cochrane Database of Systematic Reviews.

Competing interests  None declared.

Provenance and peer review  Not commissioned; externally peer reviewed.

Open Access  This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

REFERENCES
1 Black N. Why we need observational studies to evaluate the effectiveness of health care. BMJ 1996;312:1215–18.


