
NIHR Health Services and Delivery Research Programme grant 11/2003/27

Bruce Guthrie,¹,∗ Alexander Thompson,² Siobhan Dumbreck,¹ Angela Flynn,¹ Phil Alderson,³ Moray Nairn,⁴ Shaun Treweek,⁵, Katherine Payne.²

* Corresponding author

1. Population Health Sciences Division, University of Dundee, Dundee UK
2. Manchester Centre for Health Economics, The University of Manchester, Manchester UK
3. Centre for Clinical Practice, National Institute for Health and Care Excellence, Manchester UK
4. Scottish Intercollegiate Guidelines Network, Edinburgh UK
5. Health Services Research Unit, University of Aberdeen, Aberdeen UK
Important

A ‘first look’ scientific summary is created from the original author-supplied summary once the normal NIHR Journals Library peer and editorial review processes are complete. The summary has undergone full peer and editorial review as documented at NIHR Journals Library website and may undergo rewrite during the publication process. The order of authors was correct at editorial sign-off stage.

A final version (which has undergone a rigorous copy-edit and proofreading) will publish as part of a fuller account of the research in a forthcoming issue of the Health Services and Delivery Research journal.

Any queries about this ‘first look’ version of the scientific summary should be addressed to the NIHR Journals Library Editorial Office – nihredit@soton.ac.uk

The research reported in this ‘first look’ scientific summary was funded by the HS&DR programme or one of its predecessor programmes (NIHR Service Delivery and Organisation programme, or Health Services Research programme) as project number 11/2003/27. For more information visit http://www.nets.nihr.ac.uk/projects/hsdr/11200327.

The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HS&DR editors have tried to ensure the accuracy of the authors’ work and would like to thank the reviewers for their constructive comments however; they do not accept liability for damages or losses arising from material published in this scientific summary.

This ‘first look’ scientific summary presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HS&DR programme or the Department of Health. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HS&DR programme or the Department of Health.
Scientific summary

Background: Multimorbidity is the presence of two or more long-term conditions in a single person. The majority of people with any long-term condition have multimorbidity, and the majority of people aged over 65 years are multimorbid. Multimorbidity matters because people with it are the most frequent users of healthcare, reflecting that multimorbidity is associated with higher mortality, lower quality of life, increased problems of care coordination, and increased treatment burden including polypharmacy. Clinical guidelines have significantly contributed to making healthcare more evidence-based and to reducing variation in treatment. However, guidelines are increasingly criticised for contributing to excessive treatment burden and sometimes frankly futile treatment because they do not properly account for multimorbidity. This reflects that almost all guidelines are single disease focused, at least partly because the evidence base on which guidelines draw is for single diseases.

Aim and objectives: The aim of this project was to test the methodological feasibility of new approaches to summarising and creating evidence for single-disease guidelines that better account for the management of people with multimorbidity. Specific objectives were:

1. To systematically collate and summarise the evidence of benefit, harm and cost effectiveness for guideline recommendations for three common conditions, including where recommendations are mutually reinforcing or contradictory, in order to examine the value and feasibility of making existing evidence and guideline recommendations more useful for people with multimorbidity.

2. To develop and evaluate exploratory modelling methods to estimate expected benefit, time to benefit, risk and healthcare costs for people with selected multiple conditions, in order to examine the value and feasibility of new approaches to evidence creation for guidelines for people with multimorbidity.

The overall design was a literature-based and an economic modelling project in collaboration with the National Institute for Health and Care Excellence (NICE) and the Scottish Intercollegiate Guidelines Network (SIGN). To ensure that the findings were as directly relevant as possible to guideline development, the specific focus of the project within the two broad objectives was determined by a multidisciplinary project reference group (PRG) whose
professional and public members all had experience of guideline development with NICE or SIGN.

**Methods:** Analysis focused where possible on guidelines for three exemplar conditions (type 2 diabetes, depression and heart failure), chosen because they are individually important, commonly comorbid, include both physical and mental health conditions, and have treatments whose benefits accrue over different periods of time. A series of linked studies were conducted:

(1) Examination of the extent to which comorbidity was accounted for in guidelines for the three exemplar conditions, in the clinical research questions (CRQs) underpinning guideline development, treatment recommendations, and research recommendations. The extent to which CRQs were underpinned by economic evidence was also examined. Potentially serious drug-disease and drug-drug interactions between drugs recommended for each of the three exemplar conditions and drugs recommended for eleven other conditions were systematically identified.

(2) Applicability: with the advice of the PRG, for each of the three exemplar conditions we examined applicability of evidence for selected first-line drug treatment recommendations, none of which were significantly qualified by age, comorbidity or other factors. Reports and protocols of key trials informing these recommendations were examined to define the trial population studied, and epidemiological data from two large studies of multimorbidity and polypharmacy were used to compare the trial population to the guideline population for whom recommendations were being made in terms of comorbidity and co-prescribing.

(3) Absolute benefit: after discussion with the PRG, we examined the absolute benefit of selected first-line treatments for heart failure and type 2 diabetes since for both the treatments involved were long-term and preventative. Pooled relative risk was extracted from guidelines or systematic reviews, and estimates of absolute risk reduction calculated using median trial baseline risk (as is already done in NICE guidelines), then using a range of baseline risk estimates from observational data and annualised to calculate absolute benefit per year of treatment. At each stage, we identified the assumptions required for the calculations to be valid. For the comparison of absolute QALY gain, we defined the key principles that have to be met to make such comparisons valid and illustrated this by
estimating absolute QALY gain for pharmacological treatment of hypertension and the use of statins for primary prevention of cardiovascular disease.

(4) Temporal dimension of benefit: time to benefit has been proposed as an important consideration in guideline development and clinical decision-making but in practice usually reduces to using trial median duration as an unreliable proxy for the time required for net benefit to accrue. The pay-off time approach provides an alternative, and is defined as the minimum time required for the expected cumulative net benefits of an intervention to exceed its expected cumulative harms. The pay-off time is straightforward to conceive of for surgical interventions where harm may be immediate but benefit deferred, but will also apply to long-term drug treatments if there is disutility from taking a tablet that is not due to specific adverse effects, for example from the inconvenience of taking regular medication, or ordering prescriptions, or collecting medication from pharmacists (direct treatment disutility – DTD). A second temporal dimension of benefit relates to accounting for competing risk of death due to comorbidity. As an exemplar of the kind of long-term preventive treatment with small individual absolute benefit that significantly drives treatment burden in multimorbidity, we used an existing model-based cost-effectiveness analysis (CEA) created to inform the NICE lipid-modification guideline to estimate change in QALYs over time for three vignettes defined to cover a range of baseline 10-year cardiovascular risk (10% (the new primary prevention treatment threshold), 15% and 20% (the old threshold)), and examined how sensitive the findings of the model was to DTD and to increased risk of death due to comorbidity.

(5) Development of a discrete event simulation model-based CEA for people with both depression and coronary heart disease. Model-based CEA that fail to account for the particular characteristics of a multimorbid population may lack validity in much the same way as clinical evidence from single disease populations. The inclusion of more than one condition of interest for the relevant patient population poses two key challenges in terms of how to: (1) identify the important conditions to model simultaneously and then; (2) capture the interactions between the various entities (for example, the progress of the diseases simultaneously) mathematically into a structured model. Guided by the PRG, the aim of this model-based CEA was defined to be how to estimate the relative cost-effectiveness of pharmacological treatments of major depressive disorder in primary care for patients who
are also likely to go on and receive treatment for coronary heart disease. An exploratory analysis was conducted to investigate disease interaction effects and the calculation of the absolute QALY. A discrete event simulation model with a life-time horizon was used to capture the costs and benefits of antidepressant treatment for patients with depression at risk of CHD and populated with existing evidence supplemented by expert opinion where necessary.

**Results:** (1) In the exemplar guidelines, comorbidity and older age were rarely accounted for in the clinical research questions (CRQ) which frame guideline development, with the exception of the depression in adults with a chronic physical health problem where all of them were. Only half of CRQs had any associated economic evidence, and only one in seven had an associated de novo model-based CEA. All the examined guidelines cross-referenced other NICE guidance, most commonly in relation to closely related conditions (with the exception that both physical health guidelines referenced depression guidance, and one of the depression guidance was specifically in people with chronic physical problems). Although treatment recommendations did sometimes address comorbidity and drug interactions, this was most often in terms of closely related conditions, and none of the treatment recommendations were qualified in terms of reduced life expectancy. There were no explicitly contradictory recommendations across the guidelines, and no research recommendation was related to comorbidity or age except for the depression with a chronic physical health problem guideline where all were. Examining drugs recommended for the three exemplar conditions in the context of eleven other conditions, we found that 27 of the 32 potentially serious drug-disease interactions were for comorbid chronic kidney disease (CKD). Potentially serious drug-drug interactions between exemplar condition recommended drugs and other condition recommended were common for all exemplar guidelines (133 for type 2 diabetes, 89 for depression, 111 for heart failure), although varied in their likely frequency.

(2) Applicability: For all three conditions, there were large differences between people with the condition who were eligible or not eligible for the trials informing treatment recommendations. The implications of these differences varied by condition. For example, approximately 40% of people newly diagnosed with type 2 diabetes in Scotland in 2008 would have been excluded based on age alone, and these excluded older people had much
higher levels of comorbidity (e.g., CKD 25.5% of the over-75s vs 2.8% of the under-65s eligible for the relevant trials of hypoglycaemic treatment, heart failure 10.7% vs 1.7%) with type 2 diabetes having very large differences in comorbidity whereas the implications for depression were more in terms of drug-drug interactions. However, there were no common and serious drug-drug interactions of great concern. Conversely for depression, most people treated with selected antidepressants would have been eligible for relevant trials based on age, but although older people were a small minority of the treated population, co-prescription of significantly interacting drugs was very common in those aged 65 years and over (e.g., SSRI antidepressants were co-prescribed with aspirin in more than one third of patients over 65).

(3) Absolute benefit: across plausible ranges of population baseline risk the annual absolute benefit of commonly used preventive treatments were shown to be very different. For example, in heart failure due to left ventricular systolic dysfunction, the number needed to treat (NNT) with beta-blockers compared to placebo for one year to prevent one death varied from 9 (95% CI 8-11) to 34 (95% CI 29-42) across the interquartile range of baseline risk. In comparison, in people with newly diagnosed type 2 diabetes, the NNT with metformin vs diet for one year to prevent one heart attack or stroke across a plausible range of baseline risk varied from 139 (95% CI 91-625) at a 10-year cardiovascular risk of 10%, to 278 (95% CI 182-1250) at 10-year cardiovascular risk of 20%. Of note is that such calculations require significant assumptions which are not valid in all circumstances. These include assumptions that relative risk of benefit and of harm are constant across populations, that competing risks of death are not significant, and that baseline risk has been accurately measured in the guideline population or its important subgroups. We additionally examined using model-based CEA to compare interventions in terms of absolute QALY gain (which addresses the problems of attempting to compare different clinical outcomes such as death and a cardiovascular event), demonstrating that this approach is feasible. However, the validity of such comparisons relies on the use of a set of broad principles which we have defined, namely that interventions have originally been evaluated against a do-nothing option, that the length of analysis is similar, that baseline risk is comparable or appropriate to the population being considered, that a standardised reference case has been used, and ideally that uncertainty has been quantified. In practice, comparing interventions using absolute QALY gain will be limited by the availability of appropriate model-based CEAs.

© Queen’s Printer and Controller of HMSO 2016. This work was produced by Guthrie et al. under the terms of a commissioning contract issued by the Secretary of State for Health. This ‘first look’ scientific summary may be freely reproduced for the purposes of private research and study and extracts may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.
(4) Temporal dimension of benefit: Using the model created for the NICE guideline on lipid modification, we showed that DTD was associated with pay-off times measured in years (range 2.2 to 4.4 years for even the lowest level of modelled DTD, with statins for primary prevention never paying off at the higher levels of DTD) and reduced lifetime absolute QALY gain. Graphical presentation of cumulative QALYs as a QALY-profile has the potential to improve the interpretation of such data by guideline development groups. Further model modification to account for the increased mortality risk associated with multimorbidity showed that lifetime absolute QALY gain was sensitive to varying competing risks of death.

(5) It was feasible to use a discrete event simulation-based model to represent the relevant care pathways to estimate the relative cost-effectiveness of pharmacological treatments of major depressive disorder in primary care for patients who are also likely to go on and receive treatment for coronary heart disease. The model-based CEA suggested that sertraline was likely to be the most cost-effective option for patients with the two selected conditions but there were extensive levels of uncertainty around the mean incremental costs and benefits. An exploratory analysis showed that the largest absolute QALY gained was generated from sertraline. The estimated absolute QALYs gained were in the range of 0.48 to 0.57 QALYs, which are larger than those estimated for treatment with statins (~0.2 QALYs) but less than those estimated for treating hypertension (~1.0 QALYs).

**Conclusions:** The project has shown that it is feasible to address several of the important problems faced by guideline developers when attempting to account for multimorbidity. We believe that single disease guideline developers could consider piloting or implementing within their existing processes (a) The use of epidemiological data characterising the guideline population to inform guideline development group consideration of applicability and extrapolation of evidence, and interactions; (b) Systematic comparison of the absolute benefit of long-term preventive treatments in order to inform decision-making in people with reduced life-expectancy and/or high treatment burden; (c) Modification of the output from economic models used in guideline development to identify the time to accrue a benefit from treatment in terms of the pay-off time and to consider benefit in people with competing risks of death from other conditions.

**Research recommendations:** Further research is needed to design and optimise ways of presenting comparative absolute benefit to clinicians and patients, to evaluate the use of
epidemiological data in the guideline development process, to generate robust empirical estimates of direct treatment disutility and define how best to incorporate them in economic models, and to improve the underlying evidence base for treatments in multimorbid and older populations, in terms of both trials of interventions and the creation of better data on baseline risk, competing risk of mortality and harm.

Scientific summary word count 2351