Better guidelines for better care: accounting for multimorbidity in clinical guidelines – structured examination of exemplar guidelines and health economic modelling

Bruce Guthrie,[1]* Alexander Thompson,[2] Siobhan Dumbreck,[1] Angela Flynn,[1] Phil Alderson,[3] Moray Nairn,[4] Shaun Treweek[5] and Katherine Payne[2]

[1]Population Health Sciences Division, University of Dundee, Dundee, UK
[2]Manchester Centre for Health Economics, 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

*Corresponding author b.guthrie@dundee.ac.uk

Scientific summary

Accounting for multimorbidity in clinical guidelines
Health Services and Delivery Research 2017; Vol. 5: No. 16
DOI: 10.3310/hsdr05160

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, as do the majority of people aged > 65 years. Multimorbidity matters because people with it are the most frequent users of health care, reflecting the fact that multimorbidity is associated with higher mortality, lower quality of life, increased problems of care co-ordination and increased treatment burden including polypharmacy. Clinical guidelines have significantly contributed to making health care 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 the fact that almost all guidelines are focused on single diseases, 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 (TTB), risk and health-care 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 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, are commonly comorbid, include both physical and mental health conditions, and have treatments of which benefits accrue over different periods. 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 interactions between drugs each of the three exemplar conditions and drugs recommended for it and 11 other conditions were systematically identified, as were drug–drug interactions between drugs recommended for the exemplar conditions and for the others.

2. Applicability: with the advice of the PRG, for each of the three exemplar conditions we examined the applicability of evidence for selected first-line drug treatment recommendations, none of which was 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 with the guideline population for which recommendations were being made in terms of comorbidity and coprescribing.

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 were calculated by using median trial baseline risk (as is already done in NICE guidelines) and 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 quality-adjusted life-year (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: TTB 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; pay-off time 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 taking a tablet causes disutility that is not due to specific adverse effects, for example from the inconvenience of taking regular medication, 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 preventative 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 were to DTD and to increased risk of death due to comorbidity.

5. Development of a discrete event simulation (DES) model-based CEA for people with both depression and coronary heart disease (CHD): model-based CEAs 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 (e.g. the simultaneous progress of the diseases) 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 CHD. An exploratory analysis was conducted to investigate disease interaction effects and the calculation of the absolute QALY gain. A DES model with a lifetime horizon was used to capture the costs and benefits of antidepressant treatment for patients with depression at risk of CHD, and was 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 CRQs that frame guideline development, except those for depression in adults with a chronic physical health problem, which did account for them. 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 (except that both physical health guidelines referenced depression guidance, and one of the pieces of depression guidance was specifically about 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 was 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 guideline on depression with a chronic physical health problem, which was related to both. Examining drugs recommended for the three exemplar conditions in the context of 11 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 drugs recommended for exemplar conditions and those recommended for other conditions were common for all exemplar guidelines (133 for type 2 diabetes, 89 for depression, 111 for heart failure), although they varied in their likely frequency.

2. Applicability: for all three conditions, there were large differences between people with the condition who were eligible and those who were 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 in 25.5% of those aged > 75 years vs. 2.8% of those aged < 65 years who were eligible for the relevant trials of hypoglycaemic treatment; heart failure in 10.7% vs. 1.7%). Type 2 diabetes showed 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, coprescription of significantly interacting drugs was very common in those aged ≥ 65 years (e.g. selective serotonin reuptake inhibitor antidepressants were coprescribed with aspirin in more than one-third of patients aged > 65 years).

3. Absolute benefit: across plausible ranges of population baseline risk, the annual absolute benefits of commonly used preventative treatments were shown to be very different. For example, in heart failure due to left ventricular systolic dysfunction, the number needed to treat with beta-blockers compared with placebo for 1 year to prevent one death varied from 9 [95% confidence interval (CI) 8 to 11] to 34 (95% CI 29 to 42) across the interquartile range of baseline risk. In comparison, in people with newly diagnosed type 2 diabetes, the number needed to treat with metformin versus diet for 1 year to prevent one heart attack or stroke across a plausible range of baseline risk varied from 139 (95% CI 91 to 625), at a 10-year cardiovascular risk of 10%, to 278 (95% CI 182 to 1250), at a 10-year cardiovascular risk of 20%. Of note is that such calculations require significant assumptions that are not valid in all circumstances. These include assumptions that relative risks 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 also used 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.

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 2.2–4.4 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 DES-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 CHD. 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 0.48–0.57 QALYs, which is more than those estimated for treatment with statins (≈ 0.2 QALYs) but fewer 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 (1) the use of epidemiological data characterising the guideline population to inform guideline development group consideration of both likely interactions and the wider applicability and extrapolation of evidence; (2) systematic comparison of the absolute benefit of long-term preventative treatments in order to inform decision-making in people with reduced life expectancy and/or high treatment burden; and (3) 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 DTD 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.

Funding

Health Services and Delivery Research

Criteria for inclusion in the *Health Services and Delivery Research* journal
Reports are published in *Health Services and Delivery Research* (HS&DR) if (1) they have resulted from work for the HS&DR programme or programmes which preceded the HS&DR programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

HS&DR programme

The Health Services and Delivery Research (HS&DR) programme, part of the National Institute for Health Research (NIHR), was established to fund a broad range of research. It combines the strengths and contributions of two previous NIHR research programmes: the Health Services Research (HSR) programme and the Service Delivery and Organisation (SDO) programme, which were merged in January 2012.

The HS&DR programme aims to produce rigorous and relevant evidence on the quality, access and organisation of health services including costs and outcomes, as well as research on implementation. The programme will enhance the strategic focus on research that matters to the NHS and is keen to support ambitious evaluative research to improve health services.

For more information about the HS&DR programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hsdr

This report