Do Economic Evaluations in Primary Care Prevention and the Management of Hypertension Conform to Good Practice Guidelines? A Systematic Review

Maria Cristina Peñaloza Ramos, MA, Pelham Barton, PhD, Sue Jowett, PhD, Andrew John Sutton, PhD

Background: *Results of previous research have identified the need for further investigation into the compliance with good practice guidelines for current decision-analytic modeling (DAM).* ***Objective:*** *To identify the extent to which recent model-based economic evaluations of interventions focused on lowering the blood pressure (BP) of patients with hypertension conform to published guidelines for DAM in health care using a five-dimension framework developed to assess compliance to DAM guidelines.* ***Methods:*** *A systematic review of English language articles was undertaken to identify published model-based economic evaluations that examined interventions aimed at lowering BP. The review covered the period January 2000 to March 2015 and included the following electronic bibliographic databases: EMBASE and Medline via Ovid interface and the Centre for Reviews and Dissemination's (CRD) NHS-EED. Data were extracted based on different components of good practice across five dimensions utilizing a framework to assess compliance to DAM guidelines.* ***Results:*** *Thirteen articles were included in this review. The review found limited compliance to good practice DAM guidelines, which was most frequently justified by the lack of data.* ***Conclusions:*** *The assessment of structural uncertainty cannot yet be considered common practice in primary prevention and management of hypertension, and researchers seem to face difficulties with identifying sources of structural uncertainty and then handling them correctly. Additional guidelines are needed to aid researchers in identifying and managing sources of potential structural uncertainty. Adherence to guidelines is not always possible and it does pose challenges, in particular when there are limitations due to data availability that restrict, for example, a validation process.*
Key words: *decision-analytic modeling; modeling; guidelines; good practice; cardiovascular disease.* ***(MDM Policy & Practice 2016;1:1–15)***

Cardiovascular disease (CVD), which incorporates coronary heart disease (CHD) and stroke, is the main cause of death worldwide[1] and in England and Wales.[2] Hypertension, defined as a persistent raised blood pressure (BP) of 140/90 mmHg,[3] has been recognized as the most important modifiable risk factor for CVD.[2,3] Poorly controlled high BP can damage artery walls and increase the risk of developing a blood clot. Moreover, if it is not treated it can also damage organs such as the kidneys, heart, and brain. Decision-analytic modeling (DAM) guidelines have recognized that randomized controlled clinical trials (RCTs) are good sources of evidence to judge the effectiveness of treatments; however, because the time horizon for trials often does not reflect the full duration of the impact of interventions, DAM is used to extend the results of a short-term trial over a longer time horizon.[4,5] A primary outcome used in RCTs that are focused on hypertension is often change in BP. However, this is only an intermediate outcome, and DAM can be used to examine the impact of change in BP on the risk of CVD events in the longer term.

Previous research has identified the need for further investigation into the compliance of DAM to good practice and its impact on the conclusions drawn from economic evaluations.[6] Our aim is to critically evaluate how DAM in primary prevention of CVD conforms to guidelines and, in doing so,

DOI: 10.1177/2381468316671724

validate a framework previously developed to assess compliance to guidelines. The focus here is on one particular clinical area since this makes it possible to remove some of the variation between models that is not relevant for the purpose of assessing compliance (e.g., different outcomes, treatment options, or sources of uncertainty). CVD prevention has been selected due to the wide number of recent and available model-based cost-effectiveness studies conducted in this topic area. We focused on interventions aimed at lowering BP, as a modifiable risk factor for CVD, and sought to answer the following research question: "To what extent do model-based economic evaluations of primary prevention interventions aimed at lowering BP in patients with hypertension or at risk of developing hypertension conform to the published guidelines for DAM?"

METHODS

Studies of interventions aimed at lowering BP were reviewed and the challenges faced when applying DAM methods were identified and discussed. A systematic review was conducted, meeting the UK Centre for Review and Dissemination guidance and the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines for reporting.[7]

The review followed a structured approach for framing research questions: patient population (P), intervention (I), the comparator group (C), outcome (O), and the study design (S), or PICOS.[7] Articles published from January 2000 to March 2015 and written in English were included in this review if they met all of the following conditions:

- The target population was individuals presenting with high BP or at risk of developing hypertension
- The intervention(s) aimed at lowering BP
- Management of hypertension, as a modifiable risk factor for CVD, was part of a primary prevention strategy (when studies also included secondary prevention, we have concentrated on the results for primary prevention)
- The study was a model-based economic evaluation

This review excluded systematic reviews, guidelines, trials, protocols, and conference abstracts. In addition, we also excluded studies where the interventions

- Were aimed at screening BP
- Were part of a polypill strategy
- Measured nonadherence to treatment
- Were part of a secondary prevention and treatment strategy

Searches were undertaken using terms identified by expert clinical opinion and a list of synonyms identified for each term that helped inform the final search terms used in this review ("cost effectiveness," "mathematical model," "decision analysis," "Markov model," "decision tree," "economic evaluation," "hypertension," and "lowering blood pressure"). The search was undertaken using truncations and wildcards, and all synonyms were subsequently combined with appropriate medical subject heading terms (MeSH) or subject terms using Boolean operators (see Online Appendices 1 and 2).

The following databases were searched: EMBASE and Medline via the Ovid interface, and the Centre for Reviews and Dissemination's (CRD) NHS Economic Evaluation Database (NHS-EED). In addition, we manually examined the reference lists of the studies included in this review. All articles identified by database searching were exported into ENDNOTE-X7 and duplicate references were removed.

Titles identified by the searches were screened by reading the abstract; this activity was completed by two reviewers (SJ and CP). Articles that appeared to be relevant at this point were obtained and screened against the inclusion and exclusion criteria (CP); several articles appeared relevant on reading the abstract but were subsequently excluded after reading the full article.

Received 13 June 2016 from the Health Economics Unit, University of Birmingham, UK (MCPR, PB, SJ); and the Health Economics Unit, Leeds Institute of Health Sciences, University of Leeds, UK (AJS). MCPR undertook the review and analyses and wrote the first draft. All authors contributed to refined manuscript and approved the final version. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Health Economics Unit or the University of Birmingham. No funding was received for this study. Revision accepted for publication 29 July 2016.

The online appendixes for this article are available on the *Medical Decision Making Policy & Practice* Web site at http://mpp.sagepub.com/supplemental.

Address correspondence to Maria Cristina Peñaloza Ramos, Health Economics Unit, Public Health Building, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK; telephone: +44 (0)121 414 7061; e-mail: m.c.penaloza@bham.ac.uk.

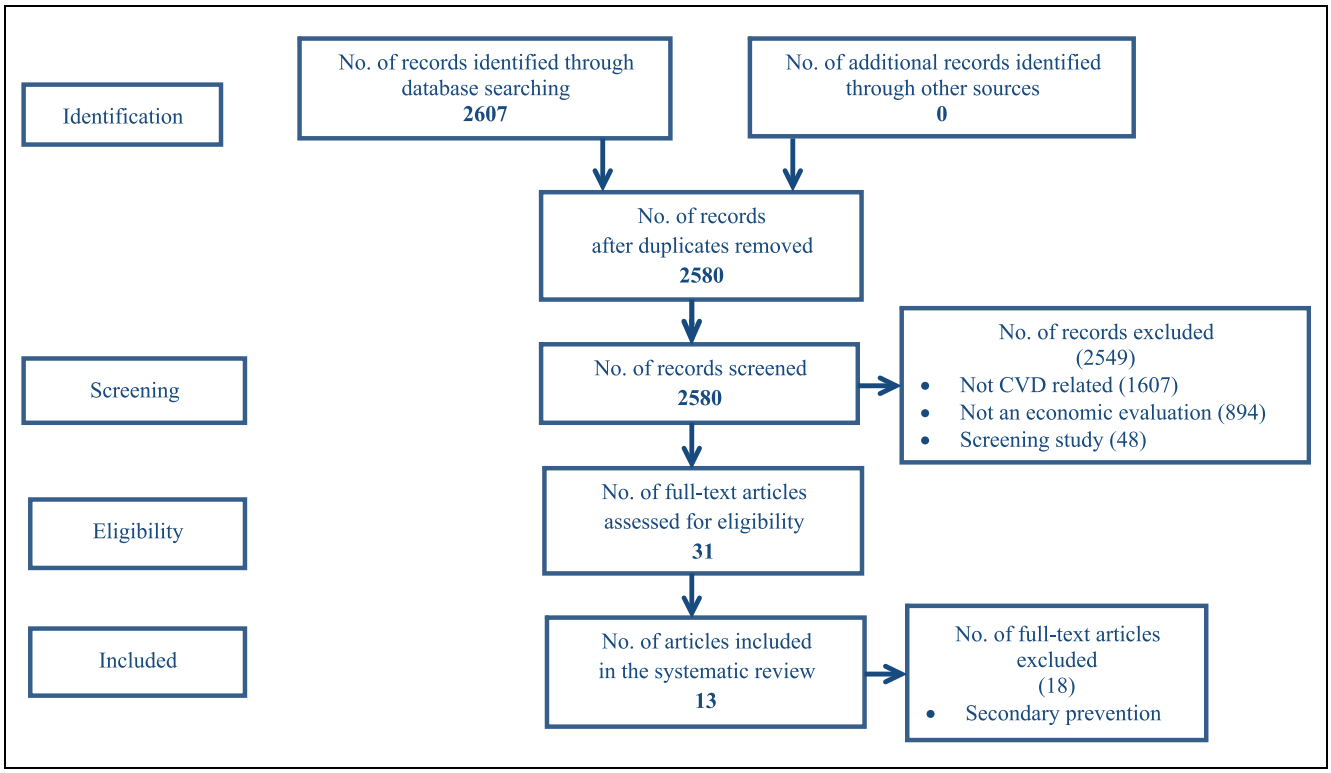

Figure 1 Flow chart using the PRISMA statement for the systematic review.

All studies were manually searched and data were extracted by a single reviewer (CP); any doubtful point(s) were checked with at least one another reviewer.[6] The extraction tool consisted of a framework[6] that synthesizes contemporary DAM guidelines in a single checklist instrument; this framework was developed to aid researchers assessing adherence to guidelines. The tool aided the retrieval and organization of information from each study across five dimensions (see Online Appendix 3):

1. Problem concept
2. Model concept
3. Synthesis of evidence
4. Analysis of uncertainty
5. Model transparency and validation

This approach ensured that the review did not miss any information related to the model building process. Data were extracted as free text and in the form of a ''yes/no'' response.

RESULTS

The database search yielded 2607 studies; after removing 27 duplicates, 2580 studies were left for screening. A total of 2549 studies were excluded because they did not consider a CVD-related intervention, were not a model-based economic evaluation, or were focused on screening (Figure 1). Thirty-one full-text articles were assessed for eligibility, of which 18 were rejected as a secondary prevention strategy. Thirteen studies were included in this review, none of which were identified through other sources (Figure 1).

Only two of the studies included were published prior to 2004. Thus, it can be seen that the majority of studies (11/13) would have had access to DAM guidelines at the time of their publication (e.g., Weinstein and others[8] or Philips and others[9]).

Four studies evaluated programs for the clinical prevention and treatment of hypertension[10–13] and nine evaluated antihypertensive drug treatments to lower BP (Table 1).[14–22] Ten studies were cost-utility analyses (CUA) or combined both CUA and a

Table 1 Summary of Analytic Framework, Methods, and Model Features of Studies Included.

Study (Year)	Research Question	Perspective/Country	Comparators	Target Population/ Intervention	Time Horizon/ Discount Rate	Reported Baseline Results	Health States	Study Design and Outcomes
Kaambwa and others (2014)[11]	LT CE of self-management of HPN	UK NHS	Self-management of HPN versus usual care	66-year-old with HPN	Lifetime; 3.5% for both	Self-management of HPN was CE	Well	CUA
						ICER £1624/QALY	Stroke MI Angina HF Death	QALYs
Stevanovic and others (2014)[14]	CE of lowering BP in patients with HPN and low CVD risk	Dutch HIS	Anti-HPN with HCTZ versus various combinations of HCTZ/Losartan (ACEIs) or HCTZ/ ARBs versus no-treatment	Various age groups: 40, 50, 60, and 65; gender and various HPN groups	10 year and lifetime; 4% for costs and 1.5% for health	Systolic BP reduction was found CE	Disease free-HPN	CEA
						A 65-year-old: • 10 year lifetime: HCT €6032/LYG man or €12,345/LYG woman • Lifetime: HCT €3076/LYG man or €3074/LYG woman	Acute CVD (nonfatal) Stable CVD (nonfatal) Fatal CVD Non-CVD death	LYG
Wu and others (2013)[15]	CE of amlodipine (CCB) versus ARB in the prevention of stroke and MI	Chinese third party payer	Amlodipine (CCB) versus ARB	Average 65-year-old cohort presenting HPN	5 years, 3% for both	Amlodipine was the dominant strategy	Disease free-HPN	CUA
							Stroke Poststroke MI Post-MI Dead	QALYs
Kourlaba and others (2013)[16]	CE of a BP lowering drug therapy in patients with mild-to-moderate HPN	Greek third party payer	Telmisartan/HCTZ compared to losartan/HCTZ and valsartan/HCTZ	Average 57-year-old cohort presenting HPN; analyses undertaken by gender	Lifetime, 3.5% for both	Telmisartan found to be CE	Disease free-HPN	CEA
						Males: €3002/QALY or €1765/LYG Females: €10,856/QALY or €7076/LYG	Nonfatal MI Post-nonfatal MI	CUA LYG
							Stroke Poststroke Death	QALYs

(continued)

Table 1 (continued)

Study (Year)	Research Question	Perspective/Country	Comparators	Target Population/ Intervention	Time Horizon/ Discount Rate	Reported Baseline Results	Health States	Study Design and Outcomes
Ekwunife and others (2013)[17]	CE of drugs in the management of HPN	Nigerian third party payer perspective for costs	Four classes of antihypertensive medications: HCTZ, pranolol (beta-blocker), lisinopril (ACE), and nifedipine (CCB)	Average 40-year-olds with HPN	30 years; 3% for both	In the low CVRS ACEI had highest (15,000 $/QALY) NMB; however, in the medium and high risk CVRSs, CCB had highest WTP (15,000 and 12,500 $/QALY, respectively)	Nonasymptomatic (disease free) Stroke Nonfatal stroke CHD Nonfatal CHD	CUA NMB EVPI US$/QALYs
Wisloff and others (2012)[12]	CE of various generic anti-HPN in the prevention of CVD	Norwegian HIS	CCB compared to no-treatment in various age groups and gender	HPN patients at different age groups (40, 50, 60, and 70)	Lifetime, 4% for both	CCB/male was CE: 40: −€456,838/LYG 50: −€445,018/LYG 60: −€410,510/LYG 70: −€352,875/LYG CCB/female was CE: 40: −€621,537/LYG 50: −€630,144/LYG 60: −€588,999/LYG 70: −€465,906/LYG	Disease free-HPN Stroke Stroke-sequelae AMI Angina HF Post-CVD Dead	CEA LYG NHB
Baker and others (2012)[22]	CE of initiating hypertension treatment with valsartan and then switching to generic losartan in the prevention of CVD	US third party payer perspective	Two comparative analyses: 1. Continual valsartan versus continual losartan 2. Continual valsartan versus valsartan switch to generic losartan	Moderate HPN patients; SBP 160–179; aged 18 and older	20-year time horizon and 3% discount for both	Treatment of moderate hypertension was considered CE with an ICER of $32,313/QALY or $27,268/LYG Switching treatment resulted in an ICER of £30,170/QALY and $25,460/LYG	CVD event free with treated HPN Post-CVD with treated HPN Death	CVD event rates per arm CEA CUA LYG QALYs

(continued)

Table 1 (continued)

Study (Year)	Research Question	Perspective/Country	Comparators	Target Population/ Intervention	Time Horizon/ Discount Rate	Reported Baseline Results	Health States	Study Design and Outcomes
Granstrom and others (2012)[21]	Long-term CE of candesartan versus losartan in the primary prevention of HPN	Swedish HIS	Candesartan versus losartan	Average 62-year-old cohort presenting with HPN	Lifetime, 3% for both	Candesartan was the dominant strategy	Disease free-HPN	CUA
							HF Chronic IHD Post-MI PAD Poststroke Arrhythmia Dead	QALYs
Perman and others (2011)[13]	CE multi-intervention program versus pharmacological strategy	Argentinian third party payer	HPN program compared to usual care	Two target groups: 65-year-old plus HPN; 65-year-old HPN and previous CVD	Lifetime, 5% for both	The HPN program was cost-effective: US$1124/LYG	Acute myocardial event	CEA
							No event Death	LYG
Ekman and others (2008)[19]	CE of irbesartan in combination with HCTZ in BP reduction	Swedish third party payer	Four strategies in male and female population: irbesartan; placebo; losartan; valsartan	55-year-old male cohort presenting with HPN	Lifetime, 3% for both	Irbesartan was CE when compared to placebo in males and females; ICERs of €3451/QALY and €7704/QALY respectively	Disease free-HPN	CUA
						Losartan and valsartan were dominated by irbesartan in males and females	Angina	QALYs
							MI Post-MI CHF Stroke Poststroke Dead	
Gandjour and Stock (2007)[10]	CE of a national HPN program for patients with essential HPN and without CVD	German HIS	National program versus no program (for low- and high-risk population)	Various age groups (40–49, 50–59, and 60–69); patients with essential HPN and without CVD	Lifetime, 3% for both	National program is CE	Disease free-HPN	CUA
						High-risk male: 40: €800/QALY 50: €880/QALY 60: €757/QALY High-risk female: 40: −€17,347/QALY 50: −€26,987/QALY 60: −€1,263/QALY	MI Stroke Renal failure Death	QALYs

(continued)

Table 1 (continued)

Study (Year)	Research Question	Perspective/Country	Comparators	Target Population/ Intervention	Time Horizon/ Discount Rate	Reported Baseline Results	Health States	Study Design and Outcomes
Montgomery and others (2003)[12]	Effectiveness and CE of LT BP lowering	UK Health Service perspective	Anti-HPN treatment versus nontreatment	Various population cohorts: 30–39; 40–49; 50–59; 60–69; 70–79. Hypertensive population	Lifetime, 6% for costs and 1.5% for effects	Treatment found more CE than nontreatment	Untreated	CUA
						ICER was higher for low-risk women compared to low-risk men	Treated T_se	QALYs
							U_cve_ua T_cve_ua T_se_cve_ua U_cve_af T_cve_af T_se_cve_af Death	
Nordmann and others (2003)[20]	CE of ACE as HPN first-line therapy versus conventional therapy	Canadian third party payer	4 strategies:	40-year old male cohort presenting with HPN but without CVD	Lifetime, 5% for both	Unfavorable results of CE:	Disease free-HPN (with or without LVH)	CEA
			• Control or conventional therapy			ECG versus control: US$0/QALY/LYG	CAD	CUA
			• ECG			EchoCar versus control: US$200,000/QALY/LYG	CVD	LYG
			• EchoCar			ACE versus control = US$700,000/QALY or US$525,000/LYG	CHF	QALYs
			• ACE				Dead	

Note: ACE = angiotensin-converting-enzyme inhibitor; AMI = acute myocardial infarction; ARB = angiotensin-II-receptor blocker; BP = blood pressure; CCB = calcium-channel blocker; CE = cost-effectiveness or cost-effective; CEA = cost-effectiveness analysis; CHF = congestive heart failure; CUA = cost-utility analysis; CVD = cardiovascular disease; EchoCar = echocardiography; EVPI = expected value of perfect information; HCTZ = hydrochlorothiazide; HF = heart failure; HIS = health insurance system; HPN = hypertension; IHD = ischemic heart disease; LYG = life year gained; MI = myocardial infarction; NHB = net health benefit; NMB = net monetary benefit; LT = long term; PAD = peripheral artery disease; T_se = treated, side-effects (health state); U_cve_ua = untreated, cardiovascular event, unaffected (health state); T_cve_ua = treated, cardiovascular event, unaffected (health state); T_se_cve_ua = treated, side-effects, cardiovascular event, unaffected (health state); U_cve_af = untreated, cardiovascular event, affected (health state); T_cve_af = treated, cardiovascular event, affected (health state); T_se_cve_af = treated, side-effect, cardiovascular event, affected (health state).

cost-effectiveness analysis (CEA)[10–12,15–17,19–22] while three studies were CEA[13,14,18] (Table 1). The intervention target (risk factor) examined was high BP. The remainder of this section describes the main findings.

Problem Concept and Model Concept

The decision problem and study objective(s) were stated in all the studies (Table 2), and all evaluated CE from a health care payer perspective. The target decision-maker audience was made explicit in 10/13 studies as that of the health care payer, that is, including only the health effects experienced by patients receiving the intervention and costs for the medical services required to provide the intervention.[23] For the remaining studies,[10,14,19] the perspective was left implicit. Ekman and others[19] commented that the analysis was ''in a Swedish health care setting'', while Stevanovic and others[14] were interested ''in the Dutch setting'' and Gandjour and Stock[10] focused on those ''insured by the German SHI,'' where SHI refers to the German Statutory Health Insurance.

For all studies, the target population was individuals with hypertension or at risk of developing hypertension (Tables 1 and 2), frequently stratified by gender, presence of hypertension, age groups, and mean age. The target population was always modeled as closed (reflecting members entering only at the start of the analysis).

Despite all the studies sharing a common aim, namely, primary prevention of CVD via lowering BP, these economic models compared a wide range of interventions and presented their results using outcome measures such as QALYs,[10–12,15–17,19–22] life years gained (LYG),[13,14,16,18,20,22] net health benefits (NHB),[18] net monetary benefits (NMB),[17] and expected value of perfect information (EVPI)[17] (Table 1).

Side effects were modeled in only one study.[15] Four studies[10,11,14,22] acknowledged the lack of adverse events as a limitation of their results due to lack of data. Two studies argued that since ''previous clinical trials found that first-line hypertensive drugs do not have more side effects than placebo''[13] or they have ''mild side effects,''[19] there was no need to model adverse effects. Similarly another study argued that fatal side effects would have been already captured in the clinical trials via the measure of effectiveness.[18]

All the studies commented on the reasons for the selection of their comparators, where their choice of comparators seems to have been governed by the scope of the study. Two studies acknowledged as a limitation the exclusion of relevant comparator(s) arguing that there may be more relevant comparators not included.[19,21] Furthermore, the ''do nothing'' option was considered in four of the studies.[10,14,18,19]

All the studies used Markov models and included a figure showing the model structure; in one study,[13] the structure of the Markov model shown in the figure did not seem to reflect the structure of the model described in the text. The model structures accounted for both acute and chronic health states. Five studies made explicit reference to how the structure of their models was defined either by using an existing generic model,[18] being based on disease progression,[10,11] or consisting of health states designed to reflect the course and history of CVD events.[22] One study reported that ''health states in the Markov model are based on cardiovascular events measured in the previously reported registry study.''[21] For the remaining studies it was inferred that the model structure was based on disease progression.

A lifetime time horizon was adopted in all but two studies: of these, one used a 5-year[15] time horizon for a population aged 65 years while the second used 20 years for a population aged 18 and older.[22] The 5-year time horizon was justified as matching the 5-year time span given to social security authorities in China for budget planning,[15] while the 20-year time horizon was not discussed.[22] Cycle length, though rarely justified in the studies, was always 1 year. Only one study[10] justified their choice as most of the data used in their model referred to a 1-year period.

Synthesis of Evidence

Patient heterogeneity was considered in most of the studies; results were presented by age cohorts[10,12,14,15,18] and gender.[10–12,14–16,18,19,21] Some studies added further analyses based on the risk of CVD,[10,12,17] scenarios of SBP reduction,[14,19] smoking,[14] and patient adherence.[14,22] The risks of secondary events were modeled in seven of the studies, for example, the risk of a further stroke after a first stroke.[12–14,18–21] In some instances, assumptions were acknowledged; for example, the study by Stevanovic and others[14] assumed the risk of secondary events to be equal to the risk of a first nonfatal CVD event. The authors acknowledged that this would lead to an underestimation of the

Table 2 Adherence to Good Practice Guidelines: Summary Results of Assessment.

Information	Review Question	Kaambwa and others (2014)[11]	Stevanovic and others (2014)[14]	Wu and others (2013)[15]	Kourlaba and others (2013)[16]	Ekwunife and others (2013)[17]	Wisloff and others (2012)[12]	Baker and others (2012)[22]	Granstrom and others (2012)[21]	Perman and others (2011)[13]	Ekman and others (2008)[19]	Gandjour and Stock (2007)[10]	Montgomery and others (2003)[12]	Nordmann and others (2003)[20]
Dimension 1: Problem concept														
Decision problem	Is there a written decision problem?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Are the study's objective(s) consistent with the decision problem and the study's scope?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Analytical perspective	Has the perspective being stated?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Target population	Has the target population being identified?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Health outcomes	Are model's outcome(s) consistent with the perspective, scope and model's objective(s)?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Have any adverse effect(s) be captured?	No	No	Yes	No	No	No	No	No	No	No	No	No	No
Interventions modeled	Are the options under evaluation clear?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Were the inclusion/exclusion of feasible options justified?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time horizon	Is it sufficient to reflect all important differences between options?	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Have time horizon, duration of the treatment and the treatment effect(s) described and justified?	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dimension 2: Model concept														
Choice of model type	Was the unit of representation given?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Does interaction(s) among individuals need to be model? If yes, was this described?	No	No	No	No	No	No	No	No	No	No	No	No	No
	Does the decision problem require a short time horizon?	No	No	No	No	No	No	No	No	No	No	No	No	No
	Is it necessary to model time in discrete cycles?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Was a type of model discussed and chosen?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Model structure	Was the starting cohort defined by demographic and clinical characteristics affecting transition probabilities or state values?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

(continued)

Table 2 (continued)

Information	Review Question	Kaambwa and others (2014)[11]	Stevanovic and others (2014)[14]	Wu and others (2013)[15]	Kourlaba and others (2013)[16]	Ekwunife and others (2013)[17]	Wisloff and others (2012)[12]	Baker and others (2012)[22]	Granstrom and others (2012)[21]	Perman and others (2011)[13]	Ekman and others (2008)[19]	Gandjour and Stock (2007)[10]	Montgomery and others (2003)[12]	Nordmann and others (2003)[20]
	Were health states and transitions reflecting the biological or theoretical understanding of the disease modeled?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Dimension 3: Synthesis of evidence														
Patient heterogeneity	Was patient heterogeneity required/considered?	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	Yes	Yes	Yes	Yes
Data sources	Were transition probabilities and intervention effects derived from representative data sources?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Were (all) methods and assumptions used to derive the model's inputs described?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Were parameters derived from observational studies controlled for confounding?	NA	NA	NA	NA	NA	NA	NA	Yes	NA	NA	NA	Yes	NA
	Was data's quality discussed?	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	If expert opinion was used, were its methods described and justified?	No	NA	NA	NA	NA	Yes	NA	NA	NA	NA	No	NA	Yes
Utilities (HSUV-weights and benefits)	Are the utilities incorporated into the model appropriate?	Yes	NA	Yes	Yes	Yes	NA	Yes	Yes	NA	Yes	Yes	Yes	Yes
	Is the source for the utility weights referenced?	Yes	NA	Yes	Yes	Yes	NA	Yes	Yes	NA	Yes	Yes	Yes	Yes
Half cycle correction	Was the use of a half cycle correction stated?	Yes	No	No	No	No	No	No	Yes	No	No	Yes	No	Yes
Resources including costs	Were the costs used in the model justified and its sources described?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Were discount rates reported and justified?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Communicating results	Did the report presented results using non-technical language aided by figures or tables?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parameter precision	Were mean value(s), distribution(s), source(s) of data and rationale for the supporting evidence described?	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Dimension 4: Analysis of uncertainty														
Analysis of uncertainty	Was analysis of uncertainty pertaining to the decision	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

(continued)

Table 2 (continued)

Information	Review Question	Kaambwa and others (2014)[11]	Stevanovic and others (2014)[14]	Wu and others (2013)[15]	Kourlaba and others (2013)[16]	Ekwunife and others (2013)[17]	Wisloff and others (2012)[12]	Baker and others (2012)[22]	Granstrom and others (2012)[21]	Perman and others (2011)[13]	Ekman and others (2008)[19]	Gandjour and Stock (2007)[10]	Montgomery and others (2003)[12]	Nordmann and others (2003)[20]
	problem included and reported?													
Parameter estimation and uncertainty	Were one-way or two-way DSA sensitivity analysis performed?	No	Yes	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Was a probabilistic sensitivity analysis (PSA) included?	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	No	Yes	No	No
Multivariate estimation and correlation	Was correlation among parameters considered?	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Structural uncertainty	Were there any discussion/ evidence of uncertainty in structural assumptions?	Yes	Yes	No	No	No	Yes	No	No	Yes	Yes	Yes	No	Yes
Other reporting of uncertainty analyses	Was EVPI measured/ discussed?	No	No	No	No	Yes	No	No	No	No	No	No	No	No
	If model calibration was used to estimate parameters, was uncertainty tested?	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Dimension 5: Model transparency and validation														
Transparency	Were the purpose, type and graphical description of the model provided?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Were the source(s) of funding and their role identified?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	No	Yes	Yes
	Were data sources identified/ described?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Were methods customized to specific application(s) and settings?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Were the effects of uncertainty measured?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Were limitations acknowledged/discussed?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Was any reference to the availability of model's documentation at request or terms and conditions to access it?	No	No	No	No	No	No	No	No	No	No	No	No	No
Validation	Was there any evidence of model's face validity?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	Was internal validity (verification or technical validity) assessed?	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No
	Was cross-validation (external consistency) assessed?	Yes	Yes	No	Yes	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes
	Was external validity assessed	No	Yes	No	No	No	No	No	No	No	No	No	No	No
	Was the model's predictive validity assessed?	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

CVD risk, and so an increased risk of death in patients experiencing nonfatal CVD events was adopted.[14] In Wisloff and others,[18] secondary nonfatal events were allowed, and a patient experiencing a secondary event was assumed to be in a health state that was worse than the state they were already in. For example, a patient with stroke sequelae that experiences a myocardial infarction (MI) will have the risk and costs associated with the stroke sequelae and not those related to MI. Perman and others[13] utilized expert opinion in the assessment of the risk of secondary events. Montgomery and others,[12] due to a lack of data, assumed that any second cardiovascular event was fatal and acknowledged this as a limitation. Some studies that did not use separate states to model secondary events[10,11,22] captured the increased mortality from secondary events through the mortality rate of patients surviving CVD events. Few of the studies acknowledged the lack of epidemiological data to model secondary events as a limitation.[11,22]

All studies applied discounting to their results: a discount rate of 3% was most common for costs and benefits;[10,15,17,19,21,22] two studies used a different discount rate for costs and benefits (Stevanovic and others[14] used 4% and 1.5%, while Montgomery and others[12] used 6% and 1.5%, respectively; Table 1). Information on the parameters used as inputs were most frequently presented in tables showing mean values and the type of distribution(s) while some studies also included 95% confidence intervals or range intervals.[10,11,20] The methods used to report the sources of information varied from reporting a detailed list of sources per parameter in a table to mentioning the sources of data in the main text.

Analysis of Uncertainty

The studies examined and reported uncertainty surrounding their identified outcomes through sensitivity analysis (SA). Uncertainty in parameter estimates was most commonly handled through deterministic sensitivity analysis (DSA) and probabilistic sensitivity analysis (PSA). Five studies used only one-way DSA,[12,15,19,20,22] while another four[11,16–18] only used PSA. Only one study measured EVPI[17] (Table 2).

Elements pertaining to structural uncertainty (SU) were acknowledged as such in six studies.[10,11,13,14,19,20] Most commonly SU was assessed through SA by varying the time horizon,[11] the duration of the effectiveness of the treatment,[11,14] the discount rate,[13,19] or by using alternative measures of outcomes.[19] One study examined the impact of assumptions related to secondary events.[18] Lack of clinical evidence for key parameters such as the treatment effect of drugs[10,11,14] was identified as a source of SU. Two studies acknowledged that they could have included more relevant comparators had they had more information,[19,21] and another two acknowledged that they had excluded a potentially relevant state due to lack of epidemiological data[10] or insufficient evidence on its relevance.[20]

The decision about which events and health states were included was partially discussed. Some studies acknowledged that they subdivided a health state[11] (CHD into MI, HF, and angina), or excluded a potentially relevant health state[15] (combined stroke and MI event). All studies included chronic health states (post events); however, few discussed having modeled the progression of disease.[10,11,22] Most frequently, the studies acknowledged the adoption of assumptions, that is, assuming the duration of treatment effects to be lifetime or as long as the time horizon in the model,[10,11,14] or 5 years[19] or varied.[21]

Model Transparency and Validation

All the studies included a graphical description of the Markov model they used (Table 2). Sources of funding were identified in 11 studies: five were funded by the pharmaceutical industry,[14–16,19,21,22] one benefited from joint funds from government and pharmaceutical sources,[18] three were exclusively government-funded,[11,12,20] and one was privately funded.[17] None of the studies stated any means for accessing more detailed information about the model. All the studies had a clear policy context with an explicit statement of funder and developer.

Validation, according to guidelines,[24] is a set of methods for judging the accuracy of a model in making relevant predictions; in other words, validation helps readers understand what a model does and how it does it. In this review, we checked for five main types of validation. All the studies were subjected to face validity checks (having been peer reviewed and published in a journal) and they were subjected to verification (internal validity checking). The methods used were justified to a greater or lesser extent in each study. All studies undertook SA of parameters as a way to double check that the direction and magnitude of their outputs were as expected.

In terms of cross-validation, results were mixed. Eight studies[10–14,17,19,20] examined different models that addressed the same problem and compared their results; however, the level of detail provided varied. Five studies presented limited or no evidence of cross-validation;[15,16,18,21,22] only Wisloff and others[18] undertook an exercise of external validation by comparing their estimated lifetimes to those reported by Statistics Norway and in doing so they found that the input into their model needed to be adjusted to fit Norwegian mortality data. An assessment of predictive validity was not included in any of the studies considered.

DISCUSSION

Using a previously developed practical framework,[6] we have critically evaluated how 13 published economic evaluations conformed to contemporaneous good practice guidelines. We found that published economic evaluations of interventions aimed at lowering BP in patients with hypertension, as part of a primary prevention strategy of CVD, demonstrated limited compliance to DAM guidelines, which has usually been explained by lack of data or imperfect data. This was particularly apparent in the assessment of SU (or lack of) and model external validation.

This review identified common grounds in terms of the adherence to, and use of, guidelines. The conceptual modeling in all the studies included in this review was based on a disease process where the focus was on the definition of the health states (conditions) as opposed to treatment (pathways) received and where the decision problem posed required the evaluation of the reduction in the risk of developing hypertension, thus explaining the use of Markov models.

It has been argued that alternative model structures can lead to variations in model predictions,[25] most importantly, in the context of a primary prevention strategy, an inappropriate model structure may lead to poorly informed policy decisions, resulting in inefficient allocation of scarce resources.[26] Models are by nature sensitive to choices made at every single stage during the model development process (i.e., model concept, model structure). There will almost always be more than one set of choices, and for this reason, guidelines have suggested assessing the extent to which model predictions are influenced by the choices made during the model development process and have suggested methods to do so, such as scenario analyses.[27,28]

Lifetime time horizons should be adopted (or be justified when constrained by the cohort's lifetime), or at the very least, time horizons should be ''long enough'' to capture relevant differences in outcomes across strategies.[23] Lack of data or imperfect data still poses important challenges for researchers, for example, when modeling the risk of secondary events and disease progression or to attempt the assessment of model validity. Even though elements pertaining to SU were identified by various authors, the assessment of SU cannot be considered common practice in this particular clinical area and additional guidelines are still needed to aid researchers identifying and quantifying SU.

External validity still poses a challenge to researchers and, more importantly, to future guidelines due to the apparent unavailability of actual extra data (from RCT or patient-level data) to undertake the exercise. It has been suggested that instead of using all the data available to create a model, some data be set aside to use during the validation process (e.g., one third of the data).[29] This may or may not always be possible, and will depend on how much data a researcher has to build a model.

Studies included in this review shared similar research questions and yet there was a great diversity in the structures of the Markov models used. Some of these were simple and some more complex, and they were generally developed with limited justification.[26] These indicate, as suggested by Squires and other,[30] that the methods for the development of the model structure are still underdeveloped. This can lead to errors including poor validity, credibility, and no basis for model verification and the analysis of structural uncertainty.

Caro and Möller[29] described the above as the disposable approach to modeling: models are built for a single use, focused on a particular product for a relatively short time. This explains—to some extent—the reduced motivation for undertaking model validation.[29] Future research should examine whether the development of ''generic models,'' or as proposed by Caro and Möller, the development of multi-use models over time, can capture sufficient detail to be realistic and avoid particulars for which there are no data, and thereby allow the economic evaluation of interventions targeting CVD in any setting, and whether this will bridge the knowledge gap and, most importantly, allow ease of comparison between the results obtained from different studies.

This is the first study that has critically reviewed compliance to DAM guidelines using a previously developed practical framework. It has covered more than a decade of published DAM studies of interventions aimed at lowering BP in patients with hypertension. We believe the inclusion of recent studies from European, American, and Asian countries has helped reflect current practice worldwide.

The exclusion criteria adopted may be considered as limitation; however, these were required to guarantee consistency in the analysis. Furthermore, a negligible number of non-English-language studies were identified pertaining to applied studies. The fact that none of the studies included was published after the release of the ''five-dimension framework'' and the selection of one particular clinical area (and any impact on generalizability this may have) may also be considered a limitation.

Our findings seem in line with recent debate around the methodological challenges being faced by DAM where model validation and SU have been identified as fundamental problems due to the lack of motivation, time, and data to validate models and, in the case of SU, a lack of methods.[29]

ACKNOWLEDGMENTS

We want to thank Professor Richard McManus and Professor Jonathan Mant for their insightful comments during the design of this study.

REFERENCES

1. World Health Organization. Global Status Report on Noncommunicable Diseases. Geneva, Switzerland: World Health Organization; 2014.

2. National Institute for Health and Clinical Excellence. Lipid Modification: Cardiovascular Risk Assessment and the Primary and Secondary Prevention of Cardiovascular Disease (NICE Guidelines CG67). London: National Institute for Health and Clinical Excellence; 2008.

3. National Institute for Health and Clinical Excellence. Hypertension: Clinical Management of Primary Hypertension in Adults (NICE Clinical Guideline 127). London: National Institute for Health and Clinical Excellence; 2011.

4. Gray AM. Cost-effectiveness analyses alongside randomised clinical trials. Clin Trials. 2006;3:538–42.

5. Ramsey S, Willke R, Briggs A, et al. Good research practices for cost-effectiveness analysis alongside clinical trials: the ISPOR RCT-CEA Task Force report. Value Health. 2005;8:521–33.

6. Peñaloza Ramos MC, Barton P, Jowett S, Sutton AJ. A systematic review of research guidelines in decision-analytic modeling. Value Health. 2015;18:512–29.

7. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med. 2009;151:264–9.

8. Weinstein MC, O'Brien B, Hornberger J, et al. Principles of good practice for decision analytic modeling in health-care evaluation: report of the ISPOR Task Force on Good Research Practices—Modeling Studies. Value Health. 2003;6:9–17.

9. Philips Z, Ginnelly L, Sculpher M, et al. Review of guidelines for good practice in decision-analytic modelling in health technology assessment. Health Technol Assess. 2004;8:1–158.

10. Gandjour A, Stock S. A national hypertension treatment program in Germany and its estimated impact on costs, life expectancy, and cost-effectiveness. Health Policy. 2007;83:257–67.

11. Kaambwa B, Bryan S, Jowett S, et al. Telemonitoring and self-management in the control of hypertension (TASMINH2): a cost-effectiveness analysis. Eur J Prev Cardiol. 2014;21:1517–30.

12. Montgomery AA, Fahey T, Ben-Shlomo Y, Harding J. The influence of absolute cardiovascular risk, patient utilities, and costs on the decision to treat hypertension: a Markov decision analysis. J Hypertens. 2003;21:1753–9.

13. Perman G, Rossi E, Waisman GD, et al. Cost-effectiveness of a hypertension management programme in an elderly population: a Markov model. Cost Eff Resour Alloc. 2011;9:4.

14. Stevanovic J, O'Prinsen AC, Verheggen BG, Schuiling-Veninga N, Postma MJ, Pechlivanoglou P. Economic evaluation of primary prevention of cardiovascular diseases in mild hypertension: a scenario analysis for the Netherlands. Clin Ther. 2014; 36:368–84.e5.

15. Wu Y, Zhou Q, Xuan J, et al. A cost-effectiveness analysis between amlodipine and angiotensin ii receptor blockers in stroke and myocardial infarction prevention among hypertension patients in China. Value Health Regional Issues. 2013;2:75–80.

16. Kourlaba G, Fragoulakis V, Theodoratou D, Maniadakis N. Economic evaluation of telmisartan, valsartan and losartan in combination with hydrochlorothiazide for treatment of mild-to-moderate hypertension in Greece: a cost-utility analysis. J Pharm Health Serv Res. 2013;4:81–8.

17. Ekwunife OI, Okafor CE, Ezenduka CC, Udeogaranya PO. Cost-utility analysis of antihypertensive medications in Nigeria: a decision analysis. Cost Eff Resour Alloc. 2013;11:2.

18. Wisloff T, Selmer RM, Halvorsen S, Fretheim A, Norheim OF, Kristiansen IS. Choice of generic antihypertensive drugs for the primary prevention of cardiovascular disease—a cost-effectiveness analysis. BMC Cardiovasc Disord. 2012;12:26.

19. Ekman M, Bienfait-Beuzon C, Jackson J. Cost-effectiveness of irbesartan/hydrochlorothiazide in patients with hypertension: an economic evaluation for Sweden. J Hum Hypertens. 2008;22: 845–55.

20. Nordmann AJ, Krahn M, Logan AG, Naglie G, Detsky AS. The cost effectiveness of ACE inhibitors as first-line antihypertensive therapy. Pharmacoeconomics. 2003;21:573–85.

21. Granstrom O, Levin L, Henriksson M. Cost-effectiveness of candesartan versus losartan in the primary preventive treatment of hypertension. Clinicoeco Outcomes Res. 2012;4:313–22.

22. Baker TM, Goh J, Johnston A, Falvey H, Brede Y, Brown RE. Cost-effectiveness analysis of valsartan versus losartan and the effect of switching. J Med Econ. 2012;15:253–60.

23. Roberts M, Russell LB, Paltiel AD, Chambers M, McEwan P, Krahn M. Conceptualizing a model: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force–2. Value Health. 2012;15:804–11.

24. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force–7. Value Health. 2012;15:843–50.

25. Jackson CH, Bojke L, Thompson SG, Claxton K, Sharples LD. A framework for addressing structural uncertainty in decision models. Med Decis Making. 2011;31:662–74.

26. Squires H, Chilcott J, Akehurst R, Burr J, Kelly MP. A systematic literature review of the key challenges for developing the structure of public health economic models. Int J Public Health. 2016;61:289–98.

27. Briggs AH. Handling uncertainty in cost-effectiveness models. Pharmacoeconomics. 2000;17:479–500.

28. Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ, Paltiel AD. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force–6. Value Health. 2012;15:835–42.

29. Caro JJ, Moller J. Decision-analytic models: current methodological challenges. Pharmacoeconomics. 2014;32:943–50.

30. Squires H, Chilcott J, Akehurst R, Burr J, Kelly MP. A framework for developing the structure of public health economic models. Value Health. 2016;19:588–601.