Simple but not simpler: a systematic review of Markov models for economic evaluation of cervical cancer screening

Juliana Yukari Kodaira Viscondi,[I] Christine Grutzmann Faustino,[I] Alessandro Gonçalves Campolina,[II] Alexander Itria,[III] Patricia Coelho de Soárez[I,*]

[I] Departamento de Medicina Preventiva, Faculdade de Medicina FMUSP, Universidade de Sao Paulo, Sao Paulo, SP, BR. [II] Instituto do Cancer do Estado de Sao Paulo (ICESP), Hospital das Clinicas HCFMUSP, Faculdade de Medicina, Universidade de Sao Paulo, Sao Paulo, SP, BR. [III] Instituto de Patologia Tropical e Saude Publica, Departamento de Saude Coletiva, Nucleo de Economia e Avaliacoes da Saude, Instituto de Avaliacao de Tecnologia em Saude, Universidade Federal de Goias, Goias, GO, BR.

*Corresponding author. E-mail: patricia.soarez@usp.br

The aim of this study was to critically evaluate the quality of the models used in economic evaluations of screening strategies for cervical cancer prevention. We systematically searched multiple databases, selecting model-based full economic evaluations (cost-effectiveness analyses, cost-utility analyses, and cost-benefit analyses) of cervical cancer screening strategies. Two independent reviewers screened articles for relevance and performed data extraction. Methodological assessment of the quality of the models utilized formal checklists, and a qualitative narrative synthesis was performed. Thirty-eight articles were reviewed. The majority of the studies were conducted in high-income countries (82%, n=31). The Pap test was the most used screening strategy investigated, which was present in 86% (n=33) of the studies. Half of the studies (n=19) used a previously published Markov model. The deterministic sensitivity analysis was performed in 92% (n=35) of the studies. The mean number of properly reported checklist items was 9 out of the maximum possible 18. Items that were better reported included the statement of decision problem, the description of the strategies/comparators, the statement of time horizon, and information regarding the disease states. Compliance with some items of the checklist was poor. The Markov models for economic evaluation of screening strategies for cervical cancer varied in quality. The following points require improvement: 1) assessment of methodological, structural, heterogeneity, and parameter uncertainties; 2) model type and cycle length justification; 3) methods to account for heterogeneity; and 4) report of consistency evaluation (through calibration and validation methods).

KEYWORDS: Uterine Cervical Neoplasms; Mass Screening; Decision Modeling; Markov Chains; Cost-benefit Analysis.

◼ INTRODUCTION

Cervical cancer continues to be an important public health problem, with an estimated 266,000 deaths from cervical cancer worldwide in 2012 (approximately 87% of cervical cancer deaths occur in less developed regions) (1). Screening programs have reduced the incidence and mortality of cervical cancer. However, substantial costs are involved in providing the infrastructure, training the manpower, buying consumables, elaborating surveillance mechanisms, and

treating and following up with patients (2). Therefore, successful programs will require using evidence-based, cost-effective approaches and strengthening national health systems (3).

Decision-analytic modeling (DAM) has increasingly been used to assess cancer prevention and control strategies in terms of their cost-effectiveness and to inform public policies. DAM supports decision makers in making choices related to the evaluated screening strategies for cervical cancer options.

Cervical screening models vary considerably in their degree of complexity. The Markov model is the most common model used to simulate the natural history of progression to cervical pre-neoplastic and neoplastic disease. This popularity is likely due to the apparent simplicity of its implementation and use.

Previous reviews (4-8) have specifically discussed the use of DAM to evaluate the cost effectiveness of cervical cancer screening, and others have discussed models that also evaluate the impact of human papillomavirus (HPV) vaccination on screening programs. However, none of these reviews critically evaluated the quality of the Markov models used in

Markov models for cervical cancer screening
Viscondi JY et al.

CLINICS 2018;73:e385

economic evaluations of screening strategies for cervical cancer using formal checklists. These instruments may identify flaws that influence the cost-effectiveness results (9). Thus, critical evaluation can confirm the credibility and reliability of the results being used by decision makers (10).

The aim of this review, which was performed as part of a health technology assessment project funded by the Brazilian Public Health System, was to provide an overview of the quality of Markov models for economic evaluation of screening strategies for cervical cancer prevention. We identify some of the most important methodological issues, reflect on the reasons for the poor report and discuss implications for research standards.

■ MATERIALS AND METHODS

Protocol and registration

This methodological systematic review was conducted based on the Centre for Reviews and Dissemination (CRD) guidelines (9) and reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist (11). A protocol was developed prior to the initiation of this review but was not registered with International Prospective Register of Systematic Reviews (PROSPERO) because this review does not contain direct patient or clinically relevant outcomes.

Eligibility criteria

Studies were included if they reported on the use of a Markov model to evaluate the costs and health outcomes of cervical cancer screening. Eligibility criteria were defined based on the components of the PICOS approach:

Participants: Markov model for economic evaluation of cervical cancer screening.

Intervention: Cervical cancer screening in settings with or without an HPV immunization program.

Comparators: Screening tests: Papanicolaou smear (Pap test), liquid-based cytology (LBC), hybrid capture (HC2), HPV-DNA, visual inspection with acetic acid (VIA), visual inspection with Lugol's iodine (VILI), and speculoscopy.

Outcome: Incremental cost-effectiveness ratio (ICER).

Study design: Model-based full economic evaluations (cost-effectiveness analyses, cost-utility analyses, and cost-benefit analyses).

This review included only English, Spanish, and German language publications. Editorials, abstracts of congress, review studies, studies that did not compare screening strategies in terms of costs and health consequences, and studies that exclusively analyzed vaccination strategies were excluded.

Electronic search

An electronic search was performed in the following databases: MEDLINE via PubMed (1946 to August (week 2) 2016), the NHS EED National Health Service Economic Evaluation Database (NHS EED) of the Centre for Reviews and Dissemination (CRD) (1994 to August week 2, 2016), EMBASE (1974 to August week 2, 2016) and Web of Science (1900 to August (week 2) 2016). The search included terms used in previous reviews (4,7,12,13) and relevant studies (14-23): ''Markov model'' AND (''Uterine Cervical Neoplasms'' OR ''Cervical Intraepithelial Neoplasia'' OR ''Squamous Intraepithelial Lesions of the Cervix'') AND ''Human Papilloma Virus'' AND ''Screening'' AND ''Costs and Cost

Analysis''. The electronic search strategies created specifically for each database are provided in Appendix 1.

Searching other sources

Additional relevant studies were identified by assessing the reference lists of major publications on the subject and the references of studies identified by electronic databases.

Study selection

This review included only Markov model-based full economic evaluations of cervical cancer screening in settings with or without an HPV immunization program. Two independent reviewers (JYKV and CGF) screened the titles and abstracts of the identified studies and selected them using specific inclusion and exclusion criteria. Any disagreements during this process were resolved by discussion or by a third reviewer (PCS).

Data collection process

Two reviewers (JYKV and CGF) independently extracted the data into a Microsoft Excel 2016 spreadsheet form tailored to this project. The data collection form was based on a prior publication (4) and was piloted in five studies.

The following data were extracted from all studies:

1. General study characteristics: authors, year of publication, country where the analysis was performed, screening tests for cervical cancer, target population, study type (cost-effectiveness analyses, cost-utility analyses), currency, year of reported costs, ICER, funding sources, conflicts of interest, health outcomes perspective and time horizon of analysis, cost-effectiveness thresholds, and HPV immunization program in place.
2. Model characteristics: use of own model, graphical representation, number of health states, cycle length, software used, calibration of parameters, model validation and types of sensitivity analysis.

Summary measures conversions

To enable comparisons across studies conducted in different countries and years and account for the effects of inflation over the designated period, the summary measures (ICERs) were updated to the year 2015. When the year of reported costs was not specified, the article publication year was used. Local currencies were initially inflated to 2015 values using specific consumer price indexes (24,25) and then converted into 2015 international dollars (I$) using purchasing power parity conversions provided by the World Bank (http://data.worldbank.org/indicator/PA.NUS.PPP) (26).

Quality assessment

We evaluated the reporting quality of the structuring and development of Markov models using items of the framework for quality assessment of DAM (27) and a previously described instrument (28). The adapted checklist is an 18-item measure of the overall quality assessment of a DAM and contains three dimensions: 1) structure, 2) data, and 3) consistency (see Appendix 2). We chose these instruments as they are widely accepted as a scientific standard for the reporting of DAM studies, and they can be applied to quality assessment of DAMs for health technology assessment (HTA). The response options for each item include 'yes', 'no' and 'not applicable'. Each reviewed study was evaluated

CLINICS 2018;73:e385

Markov models for cervical cancer screening
Viscondi JY et al.

individually, and we counted each properly reported item (answer = 'yes') and summed responses based on a maximum possible count of 18.

Synthesis of results

The more relevant results were summarized as a narrative synthesis. The study characteristics are presented in tables and figures. Due to study heterogeneity, meta-analysis or statistical pooling of the extracted summary measure (ICER) was not performed, given that this was neither feasible nor meaningful (29).

■ RESULTS

Search results

After the removal of duplicates, a total of 201 potentially relevant articles were identified. After assessment of the eligibility criteria, 38 studies (30-66) met these review inclusion criteria. Figure 1 presents the flowchart of the selection process.

Study characteristics

Table 1 presents the economic evaluation of included Markov model-based studies. The majority of the studies were conducted in high-income countries (82%, n=31). Greater than half of the studies were set in three high-income countries (USA=13, GBR=5, and CAN=3). Sixteen percent (n=6) of these studies were conducted in upper-middle income countries, and only one study included lower-middle and low-income countries (48).

The Pap test was the most commonly used screening strategy investigated and was employed in 86% (n=33) of the studies. The LBC, HC2 and HPV-DNA test were employed in 34% (n=13), 29% (n=11) and 24% (n=9) of the studies, respectively. Combined tests, such as Pap + HC2, Pap + HPV-DNA, Pap + speculum and HC2 + cytology, were employed in 26% (n=10) of the studies. Other technologies, such as VIA, VILI and self-collection, were also investigated (16%, n=6). Thirteen studies (34%) considered the effect of an HPV immunization program on the analysis.

The majority of the studies (53%, n=20) were cost-effectiveness analyses, followed by cost-utility analyses (34%, n=13) and a combination of both (13%, n=5). The lowest ICER estimate (I$156.91) was obtained in the African region (48), and the highest (I$1,173,080.66) was noted in Taiwan (46). Most of the calculated ICERs (67%, n=24) could be considered cost-effective strategies.

Half of the studies (n=19) used a previously published Markov model. In particular, five studies (36,43,49,51,66) used the model developed by Myers et al. (67). A graphical representation was presented in 68% (n=26) of the studies. The number of health states considered when stated (n=31, 82%) ranged from 4 to 23 states (mean of 12). Among the studies that reported the duration of the Markovian cycle used (n=31, 82%), the majority (n=20, 65%) considered annual cycles. Among the studies that reported (n=19, 50%) the use of some software, most studies (n=11, 58%) used TreeAge (TreeAge Software Inc., Williamstown, MA), whereas Excel (Microsoft Corp., Redmond, WA) was used by 47% (n=9) of the studies. One study used software developed by the WHO, PopMod (68), and one study implemented the model using the C ++ programming language (35) (Figure 2).
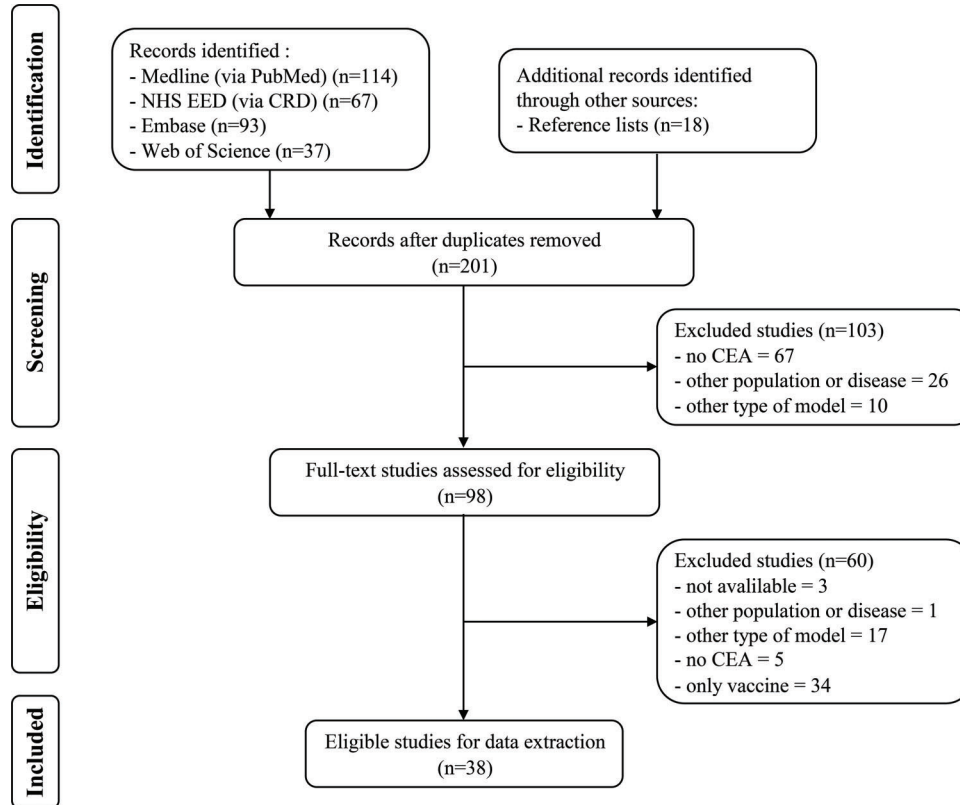

Figure 1 - Flowchart of systematic review selection process.
HEE: health economic evaluation

Markov models for cervical cancer screening
Viscondi JY et al.

CLINICS 2018;73:e385

Table 1 - Characteristics of the economic evaluation of the included Markov model-based studies.

Study	Country [1]	Screening tests [2]	Target pop [3]	Study type [4]	Currency (year) [5]	ICER (I$) [6]
(McCrory et al. 1999)	USA	PAP, NewCyto	15-85 years	CEA	USD (1997)	32,503.08/LY
(Hutchinson et al. 2000)	USA	PAP, AutoPap, LBC	20-65 years	CEA	USD (1997)	73,837.07/LY
(Myers et al. 2000)	USA	PAP, NewCyto	15-85 years	CEA	USD (1997)	4,310.60/LY
(Taylor et al. 2000)	USA	PAP+speculoscopy	18-65 years	CEA	USD (2001)	not calculated
(Montz et al. 2001)	USA	PAP, LBC	20-80 years	CEA	USD (1997)	26,532.61/LY
(Mandelblatt et al. 2002)	USA	PAP, LBC, HC2	≥20 years	CUA	USD (2000)	96,826.00/QALY
(Kulasingam and Myers 2003)	USA	PAP	12-85 years	CEA	USD (2001)	60,076.00/LY
(Goldie et al. 2004a)	USA	PAP, LBC, HC2, HC2+Cyto	≥30 years	CEA	USD (2001)	41,488.03/LY
(Goldie et al. 2004b)	USA	PAP, LBC, HC2	≥12 years	CUA	USD (2002)	77,073.34/QALY
(Karnon et al. 2004)	GBR	PAP, LBC	15-95 years	CEA	GBP (NI)	16,782.16/LY
(Kim et al. 2004)	HKG	PAP, LBC	≥15 years	CEA	USD (2000)	12,821.96/LY
(Sherlaw-Johnson and Philips 2004)	GBR	PAP, LBC, HPV test	≥15 years	CEA	GBP (2001)	6,010.34/LY
(Kim et al. 2005)	GBR, NLD, FRA, ITA	PAP, HC2, PAP+HC2	NI	CEA	USD (2004)	23,590.95 - 106,985.77/LY
(Anderson et al. 2008)	AUS	PAP	15-85 years	CEA	USD (NI)	252,485.84/LY
(Andres-Gamboa et al. 2008)	COL	PAP, HPV test	15-76 years	CEA	USD (2006)	262,302.32/LY
(Bistoletti et al. 2008)	SWE	PAP+HPV test	≥32 years	CEA	USD (2005)	7,132.05/LY
(Gutierrez-Delgado et al. 2008)	MEX	PAP, HC2, PAP+HC2	12-64 years	CUA	MXN (2006)	209,410.00/DALY
(Rogoza et al. 2008)	CAN, NLD, TWN, GBR, USA	PAP	≥12 years	CEA and CUA	CAD,EUR,NTD,GBP,USD (2006)	9,203.22 - 1,173,080.66/QALY
(Coupe et al. 2009)	NLD	PAP, HPV test	12-100 years	CUA	EUR (2006)	11,810.39/QALY
(Ginsberg et al. 2009)	WHO 14 regions	PAP, HPV test, PAP+HPV test, VIA	≥15 years	CUA	I (2000)	156.91 - 63,616.09/DALY
(Kulasingam et al. 2009)	CAN	PAP, HC2, PAP+HC2	NI	CEA and CUA	CAD (2006)	7,060.14/LY
(Reynales-Shigematsu et al. 2009)	MEX	PAP	12-85 years	CEA	USD (2004)	2,075.00/LY
(Balasubramanian et al. 2010)	USA	LBC, HPV test, Self-collection	12-85 years	CUA	USD (2007)	15,142.00/QALY
(Chuck 2010)	CAN	PAP, LBC, HC2	12-80 years	CUA	CAD (2007)	16,528.54/QALY
(Creighton et al. 2010)	AUS	PAP	10-84 years	CEA	AUD (NI)	30,583.80/LY
(Sroczynski et al. 2010)	DEU	PAP, PAP+HC2	≥15 years	CEA	EUR (2007)	107,285.43/LY
(Chen et al. 2011)	TWN	PAP, HPV test	NI	CEA	USD (NI)	64,249.86/LY
(Kitchener et al. 2011)	GBR	LBC, HC2, Self-colletion	10-84 years	CEA and CUA	GBP (2007)	49,193.51/LY
(Praditsitthikorn et al. 2011)	THA	PAP, VIA	≥15 years	CUA	THB (2007)	543,574.08/QALY
(Shi et al. 2011)	CHN Rural	VIA, VIA/VILI, HPV test	NI	CEA and CUA	USD (2009)	30,206.17/LY
(Sopina and Ashton 2011)	NZL	LBC, HPV test	12-85 years	CEA	NZD (2009)	9,164.45/QALY
(Sroczynski et al. 2011)	DEU	PAP, HC2, PAP+HC2	≥15 years	CEA	EUR (2007)	177,473.23/LY
(Vokó et al. 2012)	HUN	PAP	≥25 years	CUA	HUF (NI)	33,339.15/QALY
(Yamamoto et al. 2012)	JPN	PAP	≥11 years	CUA	JPY (2010)	28,865.17/QALY
(Fonseca et al. 2013)	BRA Amazon	PAP	≥12 years	CUA	USD (NI)	1,341.06/QALY
(Ostensson et al. 2013)	SWE	PAP, HPV test	≥15 years	CEA	EUR (2011)	8,795.19/LY
(Huh et al. 2015)	USA	PAP, HPV test, PAP+HPV test	30-70 years	CUA	USD (2013)	7,800.62/QALY
(Nghiem et al. 2015)	USA	LBC, HPV test	≥12 years	CUA	USD (2012)	18,854.49/QALY

[1] Countries classified according to the list of country names and 3-letter codes abbreviations by the United Nations (http://unstats.un.org/unsd/methods/m49/m49alpha.htm).
[2] Screening tests: PAP = conventional cytological test also known as Papanicolaou (pap) smear test; New Cyto = hypothetical new cytological test; AutoPap = automated read cytological test; LBC = liquid-based cytology test; HC2 = Digene high-risk HPV Hybrid Capture© 2 test (Qiagen); HPVtest = Human Papillomavirus (HPV) DNA detection with genotyping high-risk types by polymerase chain reaction (PCR); Self-collection = high-risk HPV DNA testing of self-collected vaginal samples; VIA = visual inspection with acetic acid; VIA/VILI = VIA in combination with Lugol's iodine.
[3] Target population: women within the age range indicated.
[4] Economic study type: CEA = cost-effectiveness analysis; CUA = cost-utility analysis.
[5] Currencies classified according to the International Organization for Standardization, ISO 4217:2015 (http://www.iso.org/iso/home/standards/currency_codes.htm).
[6] ICER = incremental cost-effectiveness ratio; I$: Geary-Khamis dollar, more commonly known as the international dollar; LY = life years; QALY = quality-adjusted life years; DALY = disability-adjusted life years; NI = Not informed.
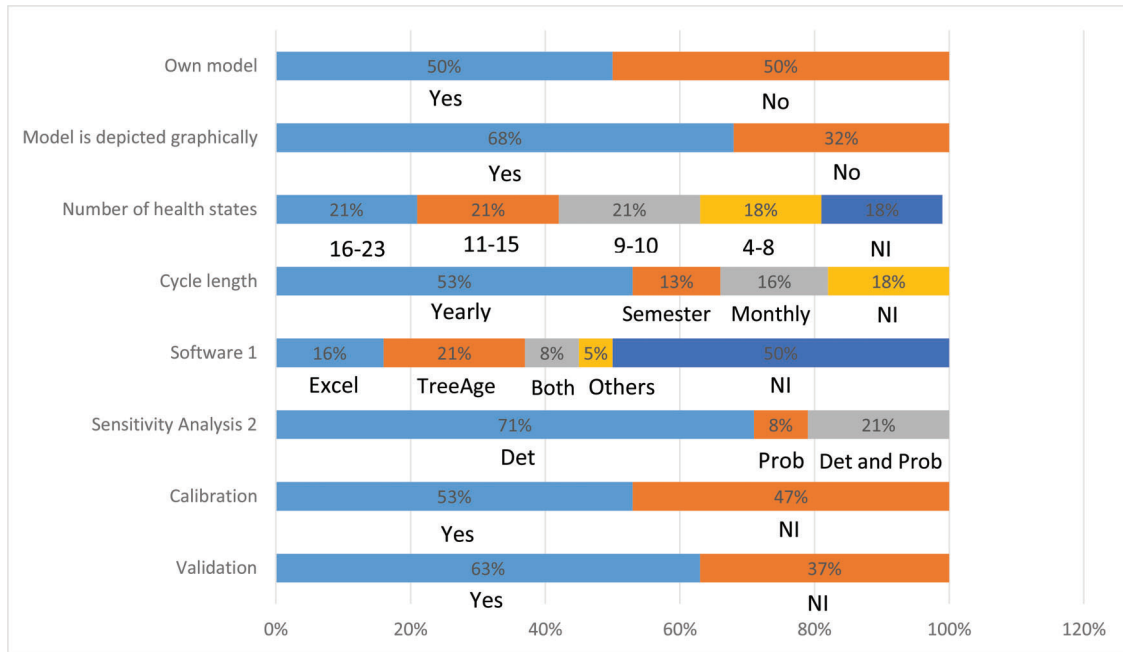
CLINICS 2018;73:e385

Markov models for cervical cancer screening
Viscondi JY et al.

Figure 2 - Decision-model characteristics of included studies.
[1] Software: Others = WHO PopMod or C++ Program.
[2] Sensitivity analysis: Det = deterministic, Prob = probabilistic.
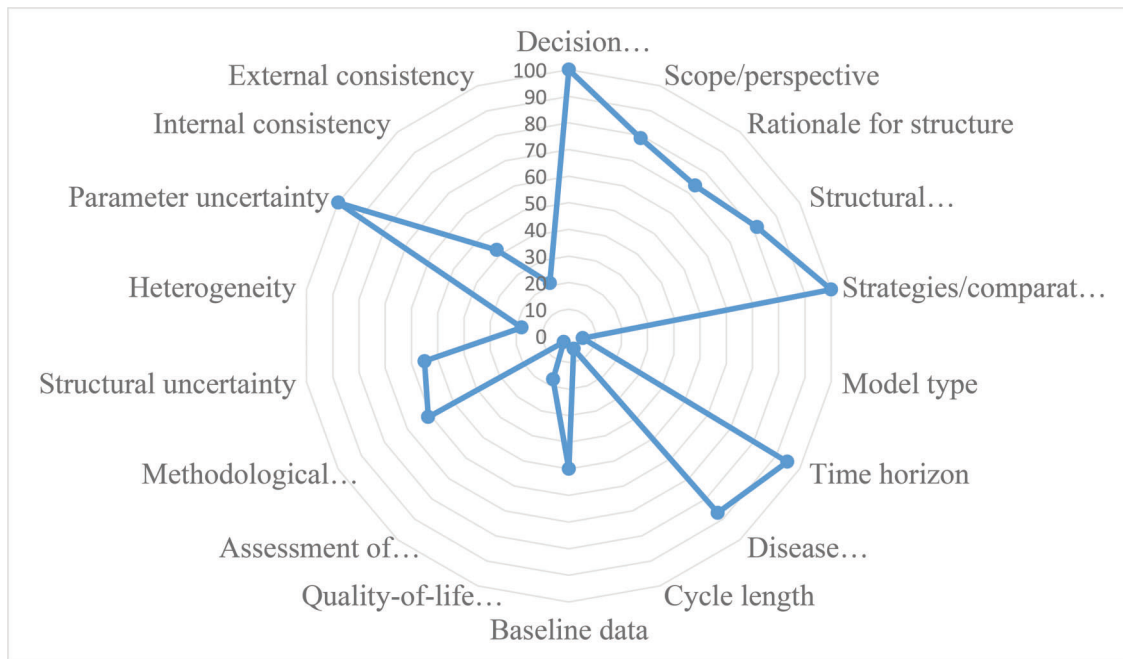NI = not informed.

Figure 3 - Proportion of economic evaluation studies adequately reporting checklist items (n=38).

Deterministic sensitivity analysis was performed in 92% (n=35) of the studies, of which 23% (n=8) also performed probabilistic analysis. The validation of the model was informed by 24 (63%) studies, whereas 53% (n=20) of the studies mentioned that model parameters were calibrated (Figure 2).

Figure 3 presents the proportion of economic evaluation studies (n=38) that properly complied with the 18 items of the checklist domains. The detailed assessment is reported in Appendix 3. The mean number of properly reported checklist items was 9 (SD 2.0) out of the maximum possible 18. Items that were better reported than others were the statement of decision problem (item 1, 100%), the description of the strategies/comparators (item 5, 100%), the statement of the time horizon (item 7, 95%) and informing the disease states (item 8, 87%). Only one study simultaneously assessed

Markov models for cervical cancer screening
Viscondi JY et al.

CLINICS 2018;73:e385

the methodological, structural, heterogeneity, and parameter uncertainties (item 12) (61). Compliance was poor for the assessment of structural uncertainty (55%, n=21) and extremely poor for the justification of model type (5%, n=2), cycle length (5%, n=2), assessment of heterogeneity (18%, n=4), the appropriateness of utilities (17%, n=4), and assessment of external consistency (21%, n=8).

■ DISCUSSION

This systematic review was the first study to comprehensively assess the methodological quality of the models of previously published studies using items of formal checklists. We evaluated 38 decision-analytic cost-effectiveness models, and the results demonstrated poor compliance with these checklists.

As noted in a previous review (12), only one study has been conducted in lower-middle and low-income countries (48), which exhibit the greatest cervical cancer burden. Approximately 84% of cervical cancer cases occur in less developed countries, with the highest incidences of cervical cancer noted in Africa, Latin America and the Caribbean. This finding reflects a lack of technical expertise and shortage of trained health economists in these regions. This finding also highlights the importance of local studies and enforces the need for strengthening the local modeling capacity.

Model structure

Half of the included studies (n=19) used a previously published Markov model. Only two studies justified the choice of model type (43,54), and the overwhelming majority did not provide reasons or explain why the use of a Markov model was appropriate. The choice of model type should be appropriate for the problem. In the case of cervical cancer, a Markov model may be suitable if the objective of the study is to assess alternative screening strategies in a setting in which disease prevalence is constant. The Markov model will simulate disease progression for a particular cohort of patients, assigning a probability of progression and regression between each of the classifications of dysplasia and invasive cancer (69). One limitation of the closed population model (such as a Markov cohort model) is that it may predict an increased cancer incidence compared with an open model. If the analysis incorporates the effect of an HPV immunization program, the ideal model would be a dynamic model that follows an entire population, allowing for evaluation of the impact of herd immunity (i.e., indirect protection of susceptible individuals by a significant proportion of immune individuals in the population) (69). Thirteen studies (36,38,45-48,50,55,57,59,62,63,66) reported that the effect of an HPV immunization program was considered in the analysis but did not explain how herd immunity was incorporated using a static cohort model.

The Markov model can be more transparent and easy to understand and provides more conservative estimates than dynamic models. In contrast, because the latter model type allows for the inclusion of more detail, it can generate several uncertainties in the evaluation process in addition to requiring more input and computational resources that may not be available in all settings. The direct and indirect effects of vaccination may not be observed in surveillance data for many years. Thus, although dynamic models are still developed by a small group of modelers (70), the development of these models will become increasingly important

to explore the impact on screening as the first vaccinated cohorts approach the age of cervical cancer screening (12). Previous studies have reported an increased screening rate among vaccinated women and the lowest proportion of cervical abnormalities compared with those not vaccinated (71,72). Future model-based economic evaluations will need to take into account the continuum interaction between screening and vaccination to predict the effects of vaccination on screening programs (6).

Only two studies justified the choice of cycle length (56,64). The cycle length should reflect the clinical problem and be the shortest interval at which the pathologies and/ or diagnosis typically occurs (73), and its justification should be based on the natural history of the disease (74). In the case of cervical cancer, often the only source of information regarding cases is the clinical examination results. However, this information may be under-reported given that HPV infections and precursor lesions may regress in less than a year (75) and screening is typically performed annually. Therefore, ideally, the definition of the cycle should not be based on the intervals between exams (74). However, occasionally, these are the only available data. The other option would be to use data from another setting, and both approaches would impact the analysis results.

Model data

Although half of the included studies presented the transition probabilities, none of them explained how the probabilities were calculated or whether the cycle correction was used. Concern has been raised in the DAM literature regarding confusion about the appropriate use of rates and probabilities. Depending on the model, this misconception may introduce important errors, impacting the validity of the model results (76,77). Various approaches can be used to estimate transition probabilities for the natural history of cervical cancer in Markov models, including a literature review of HPV and cervical intraepithelial neoplasia (CIN) progression and regression rates, data from observational studies, and fitting approaches (78). Although some relevant publications exist, no formal guidelines are available for the estimation of transition probabilities for use in Markov models (79). The understanding of the difference between rates and probabilities and how to transform them correctly is essential for those developing Markov models.

According to international guidelines, if health benefits are measured through utility measures, the methods used (e.g., time trade-off, standard gamble, specific questionnaires) and the subjects in whom the assessments were performed (e.g., patients, members of the general public, health professionals) need to be reported (80,81). Only 17% of the reviewed studies reported the applied instruments, methods of measurement and the sources of utilities employed. Inadequate reporting of utility measurement methods leads first to difficulties in comparing different assessments, given that discrepancies between these measures using different measurement instruments and methods were previously observed in other studies (82-84). In addition, in relation to the lack of reporting of sources of utility measures (populations used to derive these measures), if the ultimate objective of the evaluation is to influence the allocation of resources to decisions based on social interests, it would be important that health state evaluations were based on utility weights representative of the preferences of the general population (85).

CLINICS 2018;73:e385

Markov models for cervical cancer screening
Viscondi JY et al.

Specifically in relation to economic evaluations of cervical cancer screening, differences in utility values for CIN lesions, presence of cervical cancer and genital warts may partially explain the differences in the analysis results. In addition, considering the limited data available on the utility values associated with these states (7), it is fundamental that sensitivity analyses performed in future studies consider a wide range of variation, including all plausible utility values.

Uncertainty

Uncertainty is present in all HTA models (74). DMA researchers distinguish among parameter, structural and methodological uncertainties, all of which require assessment (27). Parameter uncertainty can be addressed by deterministic or probabilistic sensitivity analysis. Structural uncertainty can be managed through alternative model structures, which involves re-running the model under alternative structural assumptions and presenting the results of each scenario. Methodological uncertainty can be addressed with a similar method. Only one study simultaneously assessed methodological, structural, heterogeneity, and parameter uncertainties (61). Approximately half of the included studies failed to account for structural uncertainty, reflecting the gap between guidelines and applied research. This finding was also highlighted in a previous review (28), where many published models failed to account correctly for the major sources of uncertainty, particularly structural uncertainty. Most studies (92%) addressed only parameter uncertainty through deterministic sensitivity analysis. In addition to the standard considerations of uncertainty about parameter estimates, it is important to assess the implications of model uncertainty on results (28).

Most models (89%, n=34) simulated aggregate groups of women at risk of cervical cancer over time without accounting for other aspects of population heterogeneity in screening behavior. Heterogeneity (i.e., the extent to which variability between patients can be explained as a function of their characteristics) (86) reflects differences in outcomes that may in principle be explained by variations among subgroups of patients, including characteristics such as age, sex, level of risk and severity of the disease, or the relative effects of treatment (87). Given the natural history of cervical cancer, women less than 30 years of age have more HPV infections than older women, while older women may experience the progression of this virus 116-fold more frequently than younger women. Therefore, HPV-DNA screening after the age of 30 years seems to be more effective than before the age of 30 (88). Thus, not considering "heterogeneity" during the analysis, which could be performed by executing the model for different subgroups of patients, may lead to errors in the results obtained (89). To capture heterogeneity in screening and vaccination behavior, it would be ideal to use individual-based models (microsimulation).

Model consistency

Model consistency refers to the quality of the model overall. This parameter tests the internal logic of the modeling practice, changing model inputs and examining the direction of results (internal consistency). Model consistency also compares the model's result with the best available evidence or with the results of previously developed models (external consistency, also known as calibration). For instance, the model consistency of cervical precancerous lesions predicted by cytology can be compared with observed CIN-related outcomes. However, it is generally not clear whether these outcomes are predicted by cytological results or histologically confirmed lesions (8). Only 8 studies (21%) reported the use of some calibration method. This low value can be explained by the lack of standards in calibrating disease models in economic evaluation, especially cancer screening models (90,91). There is no consensus in the literature regarding an acceptable minimum specification for the fitting targets that should be reported (78). Another potential barrier to calibration is insufficient local data to estimate parameters associated with organized screening.

The Markov models for economic evaluation of screening strategies for cervical cancer varied in quality. Items that were generally well reported were the statement of the decision problem, the description of the strategies/comparators, the statement of time horizon, and informing disease states. One limitation of the present study is that most models did not adequately assess methodological, structural, heterogeneity, and parameter uncertainties. Moreover, the minority justified the model type and cycle length, assessed heterogeneity and the appropriateness of utilities, and evaluated external consistency. Future studies should evaluate the appropriateness of the different methods to account for uncertainty (through sensitivity analysis and alternative model structures), heterogeneity, consistency (through calibration and validation techniques), and the relevance of reporting guidelines for Markov models to improve their transparency.

■ AUTHOR CONTRIBUTIONS

Viscondi JY and De Soárez PC designed the research. Viscondi JY, Faustino CG and Itria A performed the research. Viscondi JY, Faustino CG, Campolina AG and De Soárez PC analyzed the data. Viscondi JY, Campolina AG and De Soárez PC wrote the manuscript.

■ REFERENCES

1. World Health Organization IARC. GLOBACAN 2012: Estimaded Cancer Incidence, Mortality and Prevalence Worldwide in 2012. 2012 [Access 2017 Sep 29]. Available from: http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx
2. Sankaranarayanan R, Budukh AM, Rajkumar R. Effective screening programmes for cervical cancer in low-and middle-income developing countries. Bull World Health Organ. 2001;79(10):954-62.
3. World Health Organization. WHO guidance note: comprehensive cervical cancer prevention and control: a healthier future for girls and women. WHO Press; 2013. 12 p.
4. Kim JJ, Brisson M, Edmunds WJ, Goldie SJ. Modeling cervical cancer prevention in developed countries. Vaccine. 2008;26 Suppl 10:K76-86, http://dx.doi.org/10.1016/j.vaccine.2008.06.009.
5. Jit M, Demarteau N, Elbasha E, Ginsberg G, Kim J, Praditsitthikorn N, et al. Human papillomavirus vaccine introduction in low-income and middle-income countries: guidance on the use of cost-effectiveness models. BMC Med. 2011;9:54, http://dx.doi.org/10.1186/1741-7015-9-54.
6. Canfell K, Chesson H, Kulasingam SL, Berkhof J, Diaz M, Kim JJ. Modeling preventative strategies against human papillomavirus-related disease in developed countries. Vaccine. 2012;30 Suppl 5:F157-67, http://dx.doi.org/10.1016/j.vaccine.2012.06.091.

Markov models for cervical cancer screening
Viscondi JY et al.

CLINICS 2018;73:e385

7. Fesenfeld M, Hutubessy R, Jit M. Cost-effectiveness of human papillo-mavirus vaccination in low and middle income countries: a systematic review. Vaccine. 2013;31(37):3786-804, http://dx.doi.org/10.1016/j.vaccine.2013.06.060.

8. Simonella L, Canfell K. Development of a quality framework for models of cervical screening and its application to evaluations of the cost-effectiveness of HPV vaccination in developed countries. Vaccine. 2015;33(1):34-51, http://dx.doi.org/10.1016/j.vaccine.2014.08.048.

9. Centre for Reviews and Dissemination. University of York. Systematic reviews: CRD's guidance for undertaking reviews in health care. York: York Publishing Services Ltd; 2009. Available from: https://www.york.ac.uk/media/crd/Systematic_Reviews.pdf

10. Olson BM, Armstrong EP, Grizzle AJ, Nichter MA. Industry's perception of presenting pharmacoeconomic models to managed care organizations. J Manag Care Pharm. 2003;9(2):159-67, http://dx.doi.org/10.18553/jmcp.2003.9.2.159.

11. Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6(7):e1000097, http://dx.doi.org/10.1371/journal.pmed.1000097.

12. Mendes D, Bains I, Vanni T, Jit M. Systematic review of model-based cervical screening evaluations. BMC Cancer. 2015;15:334, http://dx.doi.org/10.1186/s12885-015-1332-8.

13. Novaes HMD, Silva GA e, Ayres AR, Rama C, Padovan J, Sartori AM, et al. Avaliação tecnológica de vacinas para a prevenção de infecção por papilomavírus humano (HPV): estudo de custo-efetividade da incorpor-ação de vacina contra HPV no Programa Nacional de Imunizações/PNI do Brasil. Relatóriotécnico-científico. [Internet]. São Paulo: Ministério da Saúde; 2012. 154 p. Available from: ///C:/Users/Patrícia Soárez/Down-loads/AVE HPV (3).pdf

14. Colantonio L, Gomez J, Demarteau N, Standaert BA, Pichon-Riviere A, Augustovski F, et al. Cost-effectiveness analysis of a cervical cancer vac-cine in five Latin American countries. Value Health. 2009;12(7):A484, http://dx.doi.org/10.1016/S1098-3015(10)75286-4.

15. Eluf-Neto J, Booth M, Muñoz N, Bosch FX, Meijer CJ, Walboomers JM. Human papillomavirus and invasive cervical cancer in Brazil. Br J Cancer. 1994;69(1):114-9, http://dx.doi.org/10.1038/bjc.1994.18.

16. Schiffman M, Wentzensen N, Wacholder S, Kinney W, Gage JC, Castle PE. Human papillomavirus testing in the prevention of cervical cancer. J Natl Cancer Inst. 2011;103(5):368-83, http://dx.doi.org/10.1093/jnci/djq562.

17. Goldie SJ, Kim JJ, Kobus K, Goldhaber-Fiebert JD, Salomon J, O'shea MK, et al. Cost-effectiveness of HPV 16, 18 vaccination in Brazil. Vaccine. 2007;25(33):6257-70, http://dx.doi.org/10.1016/j.vaccine.2007.05.058.

18. Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, Cohen DJ, et al. State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--3. Value Health. 2012;15(6):812-20, http://dx.doi.org/10.1016/j.jval.2012.06.014.

19. Vanni T, Luz PM, Foss A, Mesa-Frias M, Legood R. Economic model-ling assessment of the HPV quadrivalent vaccine in Brazil: a dynamic individual-based approach. Vaccine. 2012;30(32):4866-71, http://dx.doi.org/10.1016/j.vaccine.2012.04.087.

20. Caetano R, Vianna CM, Thuler LC, Girianelli VR. Custo-efetividade no diagnóstico precoce do câncer de colo uterino no Brasil. Physis: Rev Saúde Coletiva. 2006;16(1):99-118.

21. Vanni T, Legood R, Franco EL, Villa LL, Luz PM, Schwartsmann G. Economic evaluation of strategies for managing women with equivocal cytological results in Brazil. Int J Cancer. 2011;129(3):671-9, http://dx.doi.org/10.1002/ijc.25708.

22. Moyer VA, U.S. Preventive Services Task Force. Screening for cervical cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2012;156(12):880-91, http://dx.doi.org/10.7326/0003-4819-156-12-201206190-00424.

23. Cox JT, Castle PE, Behrens CM, Sharma A, Wright TC Jr, Cuzick J, et al. Comparison of cervical cancer screening strategies incorporating different combinations of cytology, HPV testing, and genotyping for HPV 16/18: results from the ATHENA HPV study. Am J Obstet Gynecol. 2013;208(3):184.e1-184.e11, http://dx.doi.org/10.1016/j.ajog.2012.11.020.

24. CPI Inflation Calculator [Internet]. [Acced 2016 Jun 16]. Available from: https://www.bls.gov/data/inflation_calculator.htm

25. Inflation calculator and change of price between 2 dates [Internet]. [Acced 2016 Jun 16]. Available from: http://fxtop.com/en/inflation-calculator.php

26. PPP conversion factor, GDP (LCU per international $) [Internet]. [Acced 2016 Jun 16]. Available from: http://data.worldbank.org/indicator/PA.NUS.PPP

27. Philips Z, Bojke L, Sculpher M, Claxton K, Golder S. Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. Pharmacoeconomics. 2006;24(4):355-71, http://dx.doi.org/10.2165/00019053-200624040-00006.

28. Ramos MC, Barton P, Jowett S, Sutton AJ. A Systematic Review of Research Guidelines in Decision-Analytic Modeling. Value Health. 2015;18(4):512-29, http://dx.doi.org/10.1016/j.jval.2014.12.014.

29. Anderson R. Systematic reviews of economic evaluations: utility or futility? Health Econ. 2010;19(3):350-64, http://dx.doi.org/10.1002/hec.1486.

30. McCrory DC, Matchar DB, Bastian L, Datta S, Hasselblad V, Hickey J, et al. Evaluation of Cervical Cytology. Evidence Report/Technology Assessment No. 5. (Prepared by Duke University under Contract No. 290-97-0014.) [Internet]. AHCPR Publication No. 99-E010. Rockville, MD: Agency for Health Care Policy and Research; 1999. Available from: https://www.ncbi.nlm.nih.gov/books/NBK32961/

31. Hutchinson ML, Berger BM, Farber FL. Clinical and cost implications of new technologies for cervical cancer screening: the impact of test sensi-tivity. Am J Manag Care. 2000;6(7):766-80.

32. Myers ER, McCrory DC, Subramanian S, McCall N, Nanda K, Datta S, et al. Setting the target for a better cervical screening test: characteristics of a cost-effective test for cervical neoplasia screening. Obstet Gynecol. 2000;96(5 Pt 1):645-52, http://dx.doi.org/10.1016/S0029-7844(00)00979-0.

33. Taylor LA, Sorensen S V, Ray NF, Halpern MT, Harper DM. Cost-effectiveness of the conventional papanicolaou test with a new adjunct to cytological screening for squamous cell carcinoma of the uterine cervix and its precursors. Arch Fam Med [Internet]. 2000;9(8):713-21. Available from: http://www.ncbi.nlm.nih.gov/pubmed/10927709

34. Montz FJ, Farber FL, Bristow RE, Cornelison T. Impact of increasing Papanicolaou test sensitivity and compliance: a modeled cost and out-comes analysis. Obstet Gynecol. 2001;97(5 Pt 1):781-8, http://dx.doi.org/10.1016/S0029-7844(01)01322-9.

35. Mandelblatt JS, Lawrence WF, Womack SM, Jacobson D, Yi B, Hwang Y, et al. Benefits and costs of using HPV testing to screen for cervical cancer. JAMA. 2002;287(18):2372-81, http://dx.doi.org/10.1001/jama.287.18.2372.

36. Kulasingam SL, Myers ER. Potential health and economic impact of adding a human papillomavirus vaccine to screening programs. JAMA. 2003;290(6):781-9, http://dx.doi.org/10.1001/jama.290.6.781.

37. Goldie SJ, Kim JJ, Wright TC. Cost-effectiveness of human papillomavirus DNA testing for cervical cancer screening in women aged 30 years or more. Obstet Gynecol. 2004;103(4):619-31, http://dx.doi.org/10.1097/01.AOG.0000120143.50098.c7.

38. Goldie SJ, Kohli M, Grima D, Weinstein MC, Wright TC, Bosch FX, et al. Projected clinical benefits and cost-effectiveness of a human papilloma-virus 16/18 vaccine. J Natl Cancer Inst. 2004;96(8):604-15, http://dx.doi.org/10.1093/jnci/djh104.

39. Karnon J, Peters J, Platt J, Chilcott J, McGoogan E, Brewer N. Liquid-based cytology in cervical screening: an updated rapid and systematic review and economic analysis. Health Technol Assess. 2004;8(20):iii, 1-78, http://dx.doi.org/10.3310/hta8200.

40. Kim JJ, Leung GM, Woo PP, Goldie SJ. Cost-effectiveness of organized versus opportunistic cervical cytology screening in Hong Kong. J Public Health. 2004;26(2):130-7, http://dx.doi.org/10.1093/pubmed/fdh138.

41. Sherlaw-Johnson C, Philips Z. An evaluation of liquid-based cytology and human papillomavirus testing within the UK cervical cancer screening programme. Br J Cancer. 2004;91(1):84-91, http://dx.doi.org/10.1038/sj.bjc.6601884.

42. Kim JJ, Wright TC, Goldie SJ. Cost-effectiveness of human papillomavirus DNA testing in the United Kingdom, The Netherlands, France, and Italy. J Natl Cancer Inst. 2005;97(12):888-95, http://dx.doi.org/10.1093/jnci/dji162.

43. Anderson R, Haas M, Shanahan M. The cost-effectiveness of cervical screening in Australia: what is the impact of screening at different inter-vals or over a different age range? Aust N Z J Public Health. 2008;32(1):43-52, http://dx.doi.org/10.1111/j.1753-6405.2008.00165.x.

44. Andres-Gamboa O, Chicaiza L, Garcia-Molina M, Diaz J, Gonzalez M, Murillo R, et al. Cost-effectiveness of conventional cytology and HPV DNA testing for cervical cancer screening in Colombia. Salud Publica Mex. 2008;50(4):276-85, http://dx.doi.org/10.1590/S0036-36342008000400005.

45. Gutierrez-Delgado C, Baez-Mendoza C, Gonzalez-Pier E, de la Rosa AP, Witlen R. Generalized cost-effectiveness of preventive interventions against cervical cancer in Mexican women: results of a Markov model from the public sector perspective. Salud Publica Mex. 2008;50(2):107-18, http://dx.doi.org/10.1590/S0036-36342008000200004.

46. Rogoza RM, Ferko N, Bentley J, Meijer CJ, Berkhof J, Wang KL, et al. Optimization of primary and secondary cervical cancer prevention stra-tegies in an era of cervical cancer vaccination: a multi-regional health economic analysis. Vaccine. 2008;26 Suppl 5:F46-58, http://dx.doi.org/10.1016/j.vaccine.2008.02.039.

47. Coupe VM, de Melker HE, Snijders PJ, Meijer CJ, Berkhof J. How to screen for cervical cancer after HPV16/18 vaccination in The Netherlands. Vaccine. 2009;27(37):5111-9, http://dx.doi.org/10.1016/j.vaccine.2009.06.043.

48. Ginsberg GM, Edejer TT, Lauer JA, Sepulveda C. Screening, prevention and treatment of cervical cancer -- a global and regional generalized cost-effectiveness analysis. Vaccine. 2009;27(43):6060-79, http://dx.doi.org/10.1016/j.vaccine.2009.07.026.

49. Kulasingam SL, Rajan R, St Pierre Y, Atwood CV, Myers ER, Franco EL. Human papillomavirus testing with Pap triage for cervical cancer pre-vention in Canada: a cost-effectiveness analysis. BMC Med. 2009;7:69, http://dx.doi.org/10.1186/1741-7015-7-69.

50. Reynales-Shigematsu LM, Rodrigues ER, Lazcano-Ponce E. Cost-effectiveness analysis of a quadrivalent human papilloma virus vaccine in Mexico. Arch Med Re. 2009;40(6):503-13, http://dx.doi.org/10.1016/j.arcmed.2009.08.004.

CLINICS 2018;73:e385

Markov models for cervical cancer screening
Viscondi JY et al.

51. Balasubramanian A, Kulasingam SL, Baer A, Hughes JP, Myers ER, Mao C, et al. Accuracy and cost-effectiveness of cervical cancer screening by high-risk human papillomavirus DNA testing of self-collected vaginal samples. J Low Genit Tract Dis. 2010;14(3):185-95, http://dx.doi.org/10.1097/LGT.0b013e3181cd6d36.

52. Chuck A. Cost-effectiveness of 21 alternative cervical cancer screening strategies. Value Health. 2010;13(2):169-79, http://dx.doi.org/10.1111/j.1524-4733.2009.00611.x.

53. Creighton P, Lew JB, Clements M, Smith M, Howard K, Dyer S, et al. Cervical cancer screening in Australia: modelled evaluation of the impact of changing the recommended interval from two to three years. BMC Public Health. 2010;10:734, http://dx.doi.org/10.1186/1471-2458-10-734.

54. Sroczynski G, Schnell-Inderst P, Muhlberger N, Lang K, Aidelsburger P, Wasem J, et al. Decision-analytic modeling to evaluate the long-term effectiveness and cost-effectiveness of HPV-DNA testing in primary cervical cancer screening in Germany. GMS Health Technol Assess. 2010;6: Doc05, http://dx.doi.org/10.3205/hta000083.

55. Chen MK, Hung HF, Duffy S, Yen AM, Chen HH. Cost-effectiveness analysis for Pap smear screening and human papillomavirus DNA testing and vaccination. J Eval Clin Pract. 2011;17(6):1050-8, http://dx.doi.org/10.1111/j.1365-2753.2010.01453.x.

56. Kitchener HC, Blanks R, Cubie H, Desai M, Dunn G, Legood R, et al. MAVARIC - a comparison of automation-assisted and manual cervical screening: a randomised controlled trial. Health Technol Assess. 2011;15(3): iii-iv, ix-xi, 1-170, http://dx.doi.org/10.3310/hta15030.

57. Praditsitthikorn N, Teerawattananon Y, Tantivess S, Limwattananon S, Riewpaiboon A, Chichareon S, et al. Economic evaluation of policy options for prevention and control of cervical cancer in Thailand. Pharmacoeconomics. 2011;29(9):781-806, http://dx.doi.org/10.2165/11586560-000000000-00000.

58. Shi JF, Canfell K, Lew JB, Zhao FH, Legood R, Ning Y, et al. Evaluation of primary HPV-DNA testing in relation to visual inspection methods for cervical cancer screening in rural China: an epidemiologic and cost-effectiveness modelling study. BMC Cancer. 2011;11:239, http://dx.doi.org/10.1186/1471-2407-11-239.

59. Sopina E, Ashton T. Cost-effectiveness of a cervical screening program with human papillomavirus vaccine. Int J Technol Assess Health Care. 2011;27(4):290-7, http://dx.doi.org/10.1017/S0266462311000456.

60. Sroczynski G, Schnell-Inderst P, Muhlberger N, Lang K, Aidelsburger P, Wasem J, et al. Cost-effectiveness of primary HPV screening for cervical cancer in Germany--a decision analysis. Eur J Cancer. 2011;47(11):1633-46, http://dx.doi.org/10.1016/j.ejca.2011.03.006.

61. Vokó Z, Nagyjánosi L, Margitai B, Kövi R, Tóth Z, László D, et al. Modeling cost-effectiveness of cervical cancer screening in Hungary. Value Health. 2012;15(1):39-45, http://dx.doi.org/10.1016/j.jval.2011.10.003.

62. Yamamoto N, Mori R, Jacklin P, Osuga Y, Kawana K, Shibuya K, et al. Introducing HPV vaccine and scaling up screening procedures to prevent deaths from cervical cancer in Japan: a cost-effectiveness analysis. BJOG. 2012;119(2):177-86, http://dx.doi.org/10.1111/j.1471-0528.2011.03036.x.

63. Fonseca AJ, Ferreira LC, Neto GB. Cost-effectiveness of the vaccine against human papillomavirus in the Brazilian Amazon region. Rev Assoc Med Bras. 2013;59(5):442-51, http://dx.doi.org/10.1016/j.ramb.2013.03.004.

64. Ostensson E, Hellstrom AC, Hellman K, Gustavsson I, Gyllensten U, Wilander E, et al. Projected cost-effectiveness of repeat high-risk human papillomavirus testing using self-collected vaginal samples in the Swedish cervical cancer screening program. Acta Obstet Gynecol Scand. 2013;92(7):830-40, http://dx.doi.org/10.1111/aogs.12143.

65. Huh WK, Williams E, Huang J, Bramley T, Poulios N. Cost effectiveness of human papillomavirus-16/18 genotyping in cervical cancer screening. Appl Health Econ Health Policy. 2015;13(1):95-107, http://dx.doi.org/10.1007/s40258-014-0135-4.

66. Nghiem VT, Davies KR, Beck JR, Follen M, MacAulay C, Guillaud M, et al. Economic evaluation of DNA ploidy analysis vs liquid-based cytology for cervical screening. Br J Cancer. 2015;112(12):1951-7, http://dx.doi.org/10.1038/bjc.2015.95.

67. Myers ER, McCrory DC, Nanda K, Bastian L, Matchar DB. Mathematical model for the natural history of human papillomavirus infection and cervical carcinogenesis. Am J Epidemiol. 2000;151(12):1158-71, http://dx.doi.org/10.1093/oxfordjournals.aje.a010166.

68. Lauer JA, Röhrich K, Wirth H, Charette C, Gribble S, Murray CJ. PopMod: a longitudinal population model with two interacting disease states. Cost Eff Resour Alloc. 2003;1(1):6, http://dx.doi.org/10.1186/1478-7547-1-6.

69. Esselen KM, Feldman S. Cost-effectiveness of cervical cancer prevention. Clin Obstet Gynecol. 2013;56(1):55-64, http://dx.doi.org/10.1097/GRF.0b013e3182823797.

70. Jit M, Brisson M, Portnoy A, Hutubessy R. Cost-effectiveness of female human papillomavirus vaccination in 179 countries: a PRIME modelling study. Lancet Glob Health. 2014;2(7):e406-14, http://dx.doi.org/10.1016/S2214-109X(14)70237-2.

71. Beer H, Hibbitts S, Brophy S, Rahman MA, Waller J, Paranjothy S. Does the HPV vaccination programme have implications for cervical screening programmes in the UK? Vaccine. 2014;32(16):1828-33, http://dx.doi.org/10.1016/j.vaccine.2014.01.087.

72. Kim J, Bell C, Sun M, Kliewer G, Xu L, MclnerneyM, et al. Effect of human papillomavirus vaccination on cervical cancer screening in Alberta. CMAJ. 2016;188(12):E281-8, http://dx.doi.org/10.1503/cmaj.151528.

73. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. Med Decis Making. 1993;13(4):322-38, http://dx.doi.org/10.1177/0272989X9301300409.

74. Sculpher M, Fenwick E, Claxton K. Assessing quality in decision analytic cost-effectiveness models. A suggested framework and example of application. Pharmacoeconomics. 2000;17(5):461-77, http://dx.doi.org/10.2165/00019053-200017050-00005.

75. Castle PE, Rodríguez AC, Burk RD, Herrero R, Wacholder S, Alfaro M, et al. Short term persistence of human papillomavirus and risk of cervical precancer and cancer: population based cohort study. BMJ. 2009;339: b2569, http://dx.doi.org/10.1136/bmj.b2569.

76. Miller DK, Homan SM. Determining transition probabilities: confusion and suggestions. Med Decis Making. 1994;14(1):52-8, http://dx.doi.org/10.1177/0272989X9401400107.

77. Fleurence RL, Hollenbeak CS. Rates and probabilities in economic modelling: transformation, translation and appropriate application. Pharmacoeconomics. 2007;25(1):3-6, http://dx.doi.org/10.2165/00019053-200725010-00002.

78. Canfell K. Models of cervical screening in the era of human papillomavirus vaccination. Sex Health. 2010;7(3):359-67, http://dx.doi.org/10.1071/SH10016.

79. Olariu E, Cadwell K, Hancock E, Trueman D, Chevrou-Severac H. Current recommendations on the estimation of transition probabilities in Markov cohort models for use in health care decision-making: a targeted literature review. Clinicoecon Outcomes Res. 2017;9:537-546, http://dx.doi.org/10.2147/CEOR.S135445.

80. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. BMJ. 1996;313(7052):275-83, http://dx.doi.org/10.1136/bmj.313.7052.275.

81. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS)--explanation and elaboration: a report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. Value Health. 2013;16(2):231-50, http://dx.doi.org/10.1016/j.jval.2013.02.002.

82. Huang IC, Willke RJ, Atkinson MJ, Lenderking WR, Frangakis C, Wu AW. US and UK versions of the EQ-5D preference weights: does choice of preference weights make a difference? Qual Life Res. 2007;16(6):1065-72, http://dx.doi.org/10.1007/s11136-007-9206-4.

83. Johnson JA, Luo N, Shaw JW, Kind P, Coons SJ. Valuations of EQ-5D health states: are the United States and United Kingdom different? Med Care. 2005;43(3):221-8, http://dx.doi.org/10.1097/00005650-200503000-00004.

84. Badia X, Roset M, Herdman M, Kind P. A comparison of United Kingdom and Spanish general population time trade-off values for EQ-5D health states. Med Decis Making. 2001;21(1):7-16, http://dx.doi.org/10.1177/0272989X0102100102.

85. Ferko N, Postma M, Gallivan S, Kruzikas D, Drummond M. Evolution of the health economics of cervical cancer vaccination. Vaccine. 2008; 26 Suppl 5:F3-15, http://dx.doi.org/10.1016/j.vaccine.2008.02.004.

86. Briggs AH, Weinstein MC, Fenwick EA, Karnon J, Sculpher MJ, Paltiel AD, et al. Model parameter estimation and uncertainty: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force--6. Value Health. 2012;15(6):835-42, http://dx.doi.org/10.1016/j.jval.2012.04.014.

87. Petrou S, Gray A. Economic evaluation using decision analytical modelling: design, conduct, analysis, and reporting. BMJ. 2011;342:d1766.

88. Nahvijou A, Hadji M, Marnani AB, Tourang F, Bayat N, Weiderpass E, et al. A systematic review of economic aspects of cervical cancer screening strategies worldwide: discrepancy between economic analysis and policymaking. Asian Pac J Cancer Prev. 2014;15(19):8229-37, http://dx.doi.org/10.7314/APJCP.2014.15.19.8229.

89. Weinstein MC, O'Brien B, Hornberger J, Jackson J, Johannesson M, McCabe C, 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(1):9-17, http://dx.doi.org/10.1046/j.1524-4733.2003.00234.x.

90. Vanni T, Karnon J, Madan J, White RG, Edmunds WJ, Foss AM, et al. Calibrating models in economic evaluation: a seven-step approach. Pharmacoeconomics. 2011;29(1):35-49, http://dx.doi.org/10.2165/11584600-000000000-00000.

91. Stout NK, Knudsen AB, Kong CY, McMahon PM, Gazelle GS. Calibration methods used in cancer simulation models and suggested reporting guidelines. Pharmacoeconomics. 2009;27(7):533-45, http://dx.doi.org/10.2165/11314830-000000000-00000.

Markov models for cervical cancer screening
Viscondi JY et al.

CLINICS 2018;73:e385

Appendix 1 - Elctronic database search strategies.

Data base	Search strategy
Medline (via PubMed)	(((Model[All Fields] OR Modeling[All Fields] OR Modelling[All Fields]) AND (theoretical[All Fields] OR mathematical[All Fields] OR Markov[All Fields] OR (state[All Fields] AND (transition[All Fields] OR transitions[All Fields])))) OR "markov chains"[MeSH Terms] OR ("markov"[All Fields] AND "chains"[All Fields])) AND ("early detection of cancer"[MeSH Terms] OR (early[All Fields] AND detection[All Fields] AND cancer[All Fields]) OR "mass screening"[MeSH Terms] OR (mass[All Fields] AND screening[All Fields]) OR (cancer[All Fields] AND screening[All Fields])) AND ("human papillomavirus 16"[MeSH Terms] OR "human papillomavirus 18"[MeSH Terms] OR ((humans[MeSH Terms] OR humans[All Fields] OR human[All Fields]) AND (papillomaviridae[MeSH Terms] OR papillomaviridae[All Fields] OR papillomavirus[All Fields])) OR HPV[All Fields] OR "papillomavirus infections"[MeSH Terms] OR (papillomavirus[All Fields] AND (infection[All Fields] OR infections[All Fields]))) AND ("cervical intraepithelial neoplasia"[MeSH Terms] OR (cervical[All Fields] AND intraepithelial[All Fields] AND neoplasia[All Fields]) OR "squamous intraepithelial lesions of the cervix"[MeSH Terms] OR (squamous[All Fields] AND intraepithelial[All Fields] AND lesions[All Fields] AND cervix[All Fields]) OR "uterine cervical diseases"[MeSH Terms] OR "uterine cervical dysplasia"[MeSH Terms] OR "uterine cervical neoplasms"[MeSH Terms] OR (uterine[All Fields] AND cervical[All Fields] AND (diseases[All Fields] OR dysplasia[All Fields] OR neoplasms[All Fields])) OR (("cervix uteri"[MeSH Terms] OR (cervix[All Fields] AND uteri[All Fields]) OR cervix[All Fields]) AND (neoplasms[MeSH Terms] OR neoplasms[All Fields] OR cancer[All Fields] OR carcinoma[MeSH Terms] OR carcinoma[All Fields]) OR dysplasia[All Fields]) OR "cervical cancer"[All Fields] OR ("cervical"[All Fields] AND "cancer"[All Fields]) OR hsil[All Fields] OR lsil[All Fields]) AND ("cost-benefit analysis"[MeSH Terms] OR ("cost-benefit"[All Fields] OR "cost benefit"[All Fields] OR "cost-effectiveness"[All Fields] OR "cost effectiveness"[All Fields] OR "cost-utility"[All Fields] OR "cost utility"[All Fields] OR economic[All Fields]) AND (analysis[All Fields] OR analyses[All Fields] OR evaluation[All Fields]) OR "costs and cost analysis"[MeSH Terms] OR ((costs[All Fields] OR cost[All Fields]) AND (analysis[All Fields] OR analyses[All Fields])))
NHS EED (via CRD)*	(cervical cancer) AND (screening) AND (Markov)
Embase	(('model'/exp or model or 'modeling'/exp or modeling or 'modelling'/exp or modelling or 'mathematical model'/exp or 'mathematical model' or 'theoretical model'/exp or 'theoretical model') and ('hidden markov model'/exp or 'hidden markov model' or 'markov chains'/exp or 'markov chains' or 'state transition' or 'markov')) and ('cancer screening'/exp or 'cancer screening' or ('cancer'/exp or cancer and ('screening'/exp or screening)) or 'mass screening'/exp or 'mass screening' or ('mass'/exp or mass and ('screening'/exp or screening)) or 'early diagnosis'/exp or 'early diagnosis') and ('alphapapillomavirus'/exp or 'alphapapillomavirus' or 'human papillomavirus type 11'/exp or 'human papillomavirus type 11' or 'human papillomavirus type 16'/exp or 'human papillomavirus type 16' or 'human papillomavirus'/exp or 'human papillomavirus' or 'papillomavirus infection'/exp or 'papillomavirus infection' or 'virus carcinogenesis'/exp or 'virus carcinogenesis') and ('uterine cervix tumor'/exp or 'uterine cervix tumor' or 'uterine cervix cancer'/exp or 'uterine cervix cancer' or 'uterine cervix carcinoma'/exp or 'uterine cervix carcinoma' or 'uterine cervix carcinoma in situ'/exp or 'uterine cervix carcinoma in situ' or 'uterine cervix dysplasia'/exp or 'uterine cervix dysplasia' or 'squamous intraepithelial lesion of the cervix'/exp or 'squamous intraepithelial lesion of the cervix' or cin) and ('economic evaluation'/exp or 'economic evaluation' or (economic and ('evaluation'/exp or evaluation)) or 'cost benefit analysis'/exp or 'cost benefit analysis' or 'cost effectiveness analysis'/exp or 'cost effectiveness analysis' or 'cost minimization analysis'/exp or 'cost minimization analysis' or 'cost utility analysis'/exp or 'cost utility analysis' or (benefit or effectiveness or utility and ('analysis'/exp or analysis or analyses or 'evaluation'/exp or evaluation)))
Web of Science	#1 Tópico: (model) OR Tópico: (modeling) OR Tópico: (modelling) #2 Tópico: (markov) OR Tópico: (markov chains) OR Tópico: (state transition) OR Tópico: (static) #3 #2 AND #1 #4 Tópico: (early detection) OR Tópico: (screening) OR Tópico: (diagnosis) #5 Tópico: (human papillomavirus) OR Tópico: (hpv) #6 Tópico: (cervicalcancer) OR Tópico: (cervix uteri cancer) OR Tópico: (cervix uteri cancer) OR Tópico: (uterine cervical dysplasia) OR Tópico: (cervical intraepithelial neoplasia) OR Tópico: (cervicalintraepithelialneoplasm) OR Tópico: (squamous intraepithelial lesions) #7 Tópico: (cost) OR Tópico: (cost-benefit) OR Tópico: (cost-effectiveness) OR Tópico: (cost-utility) OR Tópico: (economic*) #8 #7 AND #6 AND #5 AND #4 AND #3

*UK National Health Services Economic Evaluation Database (NHS EED, via Centre for Review and Dissemination - CRD)

CLINICS 2018;73:e385

Markov models for cervical cancer screening
Viscondi JY et al.

Appendix 2 - Adapted checkliust for quality appraisal of decision-analytic modelling

Dimension of quality	Question(s) for critical appraisal
Structure (S)	
S1: Statement of decision problem/objective	Is there a clear statement of the decision problem?
	Is the objective of the evaluation and model specified and consistent with the stated decision problem?
	Is the primary decision maker specified?
S2: Statement of scope/perspective	Is the perspective of the model stated clearly?
	Are the model inputs consistent with the stated perspective?
	Has the scope of the model been stated and justified?
	Are the outcomes of the model consistent with the perspective, scope and overall objective of the model?
S3: Rationale for structure	Has the evidence regarding the model structure been described?
	Is the structure of the model consistent with a coherent theory of the health condition under evaluation?
	Have any competing theories regarding model structure been considered?
	Are the sources of data used to develop the structure of the model specified?
	Are the causal relationships described by the model structure justified appropriately?
S4: Structural assumptions	Are the structural assumptions transparent and justified?
	Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?
S5: Strategies/comparators	Is there a clear definition of the options under evaluation?
	Have all feasible and practical options been evaluated?
	Is there justification for the exclusion of feasible options?
S6: Model type	Is the chosen model type appropriate given the decision problem and specified causal relationships within the model?
S7: Time horizon	Is the time horizon of the model sufficient to reflect all important differences between options?
	Is the time horizon of the model, the duration of treatment and the duration of treatment effect described and justified?
	Has a lifetime horizon been used? If not, has a shorter time horizon been justified?
S8: Disease states/pathways	Do the disease states (state transition model) or the pathways (decision tree model) reflect the underlying biological process of the disease in question and the impact of interventions?
S9: Cycle length	Is the cycle length defined?
	Is the cycle length justified in terms of the natural history of disease?
Data (D)	
D2a: Baseline data	Is the choice of baseline data described and justified?
	Are transition probabilities calculated appropriately?
	Has a half cycle correction been applied to both cost and outcome? If not, has this omission been justified?
D2c: Quality-of-life weights (utilities)	Are the utilities incorporated into the model appropriate?
	Is the source for the utility weights referenced?
	Are the methods of derivation for the utility weights justified?
D4: Assessment of uncertainty	Have the four principal types of uncertainty been addressed?
	If not, has the omission of particular forms of uncertainty been justified?
D4a: Methodological	Have methodological uncertainties been addressed by running alternative versions of the model with different methodological assumptions?
D4b: Structural	Is there evidence that structural uncertainties have been addressed via sensitivity analysis?
D4c: Heterogeneity	Has heterogeneity been dealt with by running the model separately for different sub-groups?
D4d: Parameter	Are the methods of assessment of parameter uncertainty appropriate?
	Has probabilistic sensitivity analysis been done, if not has this been justified?
	If data are incorporated as point estimates, are the ranges used for sensitivity analysis stated clearly and justified?
Consistency (C)	
C1: Internal consistency	Is there evidence that the mathematical logic of the model has been tested thoroughly before use?
C2: External consistency	Are the conclusions valid given the data presented?
	Are any counterintuitive results from the model explained and justified?
	If the model has been calibrated against independent data, have any differences been explained and justified?
	Have the results of the model been compared with those of previous models and any differences in results explained?

Markov models for cervical cancer screening
Viscondi JY et al.

CLINICS 2018;73:e385

Appendix 3 - Critical evaluation of the quality of the Morkov

Study	Structure (S)									Data (D)							Consistency (C)	
	S1	S2	S3	S4	S5	S6	S7	S8	S9	D1	D2	D3	D3a	D3b	D3c	D3d	C1	C2
(McCrory et al. 1999)	✓	✓	✓	✓	✓	✗	✗	✓	✗	✓	na	✗	✗	✓	✗	✗	✓	✓
(Hutchinson et al. 2000)	✓	✓	✓	✓	✓	✗	✓	✗	✗	✗	na	✗	✗	✓	✗	✗	✓	✗
(Myers et al. 2000)	✓	✓	✗	✓	✓	✗	✓	✓	✗	✗	na	✗	✗	✓	✗	✗	✓	✗
(Taylor et al. 2000)	✓	✗	✓	✓	✓	✗	✓	✓	✗	✗	na	✗	✗	✓	✓	✗	✗	✗
(Montz et al. 2001)	✓	✗	✓	✓	✓	✗	✓	✓	✗	✗	na	✗	✗	✓	✓	✗	✗	✗
(Mandelblatt et al. 2002)	✓	✓	✓	✓	✓	✗	✓	✓	✗	✓	✗	✗	✗	✓	✗	✗	✓	✗
(Kulasingam and Myers 2003)	✓	✓	✗	✓	✓	✗	✓	✓	✓	✗	na	✗	✗	✓	✗	✗	✗	✗
(Goldie et al. 2004a)	✓	✓	✗	✓	✓	✓	✗	✓	✓	✗	✗	✗	✗	✓	✗	✗	✓	✗
(Goldie et al. 2004b)	✓	✓	✓	✓	✓	✗	✓	✓	✓	✗	✓	✗	✗	✓	✗	✗	✓	✓
(Karnon et al. 2004)	✓	✓	✓	✓	✓	✗	✓	✓	✗	✓	✗	✗	✗	✓	✗	✓	✓	✗
(Kim et al. 2004)	✓	✓	✓	✓	✓	✗	✓	✓	✗	✓	✗	✗	✗	✓	✗	✗	✗	✗
(Sherlaw-Johnson and Philips 2004)	✓	✓	✓	✓	✓	✗	✓	✓	✗	✗	na	✗	✗	✓	✗	✗	✗	✗
(Kim et al. 2005)	✓	✓	✓	✓	✓	✗	✓	✓	✗	✓	✗	✗	✗	✓	✗	✗	✓	✗
(Anderson et al. 2008)	✓	✗	✓	✓	✓	✓	✓	✓	✗	✓	na	✗	✗	✓	✗	✗	✗	✗
(Andres-Gamboa et al. 2008)	✓	✓	✓	✓	✓	✗	✓	✓	✗	✓	na	✗	✗	✓	✗	✗	✗	✗
(Bistoletti et al. 2008)	✓	✗	✓	✗	✓	✗	✓	✓	✓	✗	na	✗	✗	✓	✗	✗	✗	✗
(Gutierrez-Delgado et al. 2008)	✓	✓	✓	✓	✓	✗	✓	✓	✗	✗	✗	✗	✗	✓	✗	✗	✗	✗
(Rogoza et al. 2008)	✓	✓	✗	✓	✓	✗	✓	✓	✗	✗	✗	✗	✗	✓	✗	✗	✗	✗
(Coupe et al. 2009)	✓	✓	✓	✓	✓	✗	✓	✓	✗	✓	✗	✗	✗	✓	✗	✗	✗	✗
(Ginsberg et al. 2009)	✓	✗	✗	✗	✓	✗	✓	✗	✗	✓	✗	✗	✗	✗	✗	✓	✗	✗
(Kulasingam et al. 2009)	✓	✓	✓	✓	✓	✗	✓	✓	✗	✗	✗	✗	✗	✓	✓	✓	✓	✗
(Reynales-Shigematsu et al. 2009)	✓	✓	✗	✗	✓	✗	✓	✗	✗	✗	na	✗	✗	✓	✗	✗	✓	✓
(Balasubramanian et al. 2010)	✓	✗	✗	✗	✓	✗	✓	✗	✗	✗	✗	✗	✗	✓	✗	✗	✗	✗
(Chuck 2010)	✓	✓	✓	✓	✓	✗	✓	✓	✗	✓	✗	✗	✗	✗	✗	✓	✗	✗
(Creighton et al. 2010)	✓	✗	✗	✗	✓	✗	✓	✓	✗	✗	na	✗	✗	✗	✗	✗	✗	✗
(Sroczynski et al. 2010)	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	na	✗	✗	✓	✗	✗	✓	✓
(Chen et al. 2011)	✓	✓	✓	✓	✓	✗	✓	✓	✗	✓	✗	✗	✗	✗	✗	✓	✗	✗
(Kitchener et al. 2011)	✓	✓	✓	✓	✓	✗	✓	✓	✓	✓	✓	✗	✗	✓	✗	✗	✗	✓
(Praditsitthikorn et al. 2011)	✓	✓	✓	✗	✓	✗	✓	✓	✗	✗	✓	✗	✓	✓	✗	✓	✗	✗
(Shi et al. 2011)	✓	✓	✓	✗	✓	✗	✓	✓	✗	✗	✗	✗	✗	✓	✗	✗	✗	✗
(Sopina and Ashton 2011)	✓	✗	✗	✓	✓	✗	✓	✓	✗	✗	✗	✗	✗	✓	✗	✗	✓	✗
(Sroczynski et al. 2011)	✓	✓	✓	✓	✓	✗	✓	✓	✗	✓	na	✗	✗	✓	✗	✗	✓	✓
(Vokó et al. 2012)	✓	✓	✓	✓	✓	✗	✗	✓	✗	✗	✓	✗	✗	✓	✓	✓	✓	✓
(Yamamoto et al. 2012)	✓	✓	✓	✓	✓	✗	✓	✓	✗	✓	✗	✗	✗	✗	✗	✗	✓	✗
(Fonseca et al. 2013)	✓	✓	✗	✓	✓	✗	✓	✗	✓	✗	✗	✗	✗	✓	✗	✗	✗	✗
(Ostensson et al. 2013)	✓	✓	✓	✓	✓	✗	✓	✓	✗	✓	na	✗	✗	✓	✗	✗	✗	✗
(Huh et al. 2015)	✓	✓	✓	✓	✓	✗	✓	✓	✗	✓	✓	✗	✗	✓	✗	✓	✓	✓
(Nghiem et al. 2015)	✓	✓	✓	✓	✓	✗	✓	✓	✗	✗	✗	✗	✗	✓	✗	✓	✗	✗