Vaccine xxx (2014) xxx–xxx

Contents lists available at ScienceDirect

Vaccine

journal homepage: www.elsevier.com/locate/vaccine

Review

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

Q1 Leonardo Simonella [a,1,2], Karen Canfell [b,*,1,2]

[a] Saw Swee Hock School of Public Health, National University of Singapore, Singapore
[b] Lowy Cancer Research Centre, Prince of Wales Clinical School, The University of NSW, Sydney, Australia

ARTICLE INFO

Article history:
Received 18 September 2013
Received in revised form 1 July 2014
Accepted 15 August 2014
Available online xxx

Keywords:
HPV vaccination
Cervical screening
Systematic review
Quality framework

ABSTRACT

Background: HPV vaccination has now been introduced in most developed countries, but this has occurred in the context of established cervical cancer screening mechanisms which provide population-level protection against the most common HPV-related cancer. Therefore, estimating the cost-effectiveness of HPV vaccination to further reduce HPV-related disease depends in large part on the estimation of the effectiveness of the cervical screening 'background'. The aim of this study was to systematically review and assess methods for simulating cervical screening in decision analytic models used for evaluation of HPV vaccination.
Methods: Existing quality frameworks for economic models were extended to develop a specific quality framework for models of cervical screening. This involved domains for model structure, parameterisation (data sources) and validation (consistency). A systematic review of economic evaluations of HPV vaccination was then conducted, and assessment of cervical screening model components was then performed via application of the new quality framework.
Results: Generally, models took into account population-level cervical screening participation, but were inconsistent in their approach modelling abnormal smear management, diagnostic evaluation and treatment of precancerous disease. There was also considerable variability in the accuracy of modelling clinical pathways and the scope of validation performed for screening-related outcomes, with focus directed towards cervical cancer targets. Only a few models comprehensively validated against observed pre-cancerous abnormalities; however a number of them did not consider the effect of screening test occurrence and/or accuracy on detection of precancerous abnormalities.
Conclusion: Models of HPV vaccination in developed countries can be improved by further attention to the 'background' modelling of secondary protection via cervical screening. The quality framework developed for this review can be used to inform future HPV vaccination evaluations, including evaluations of the cost-effectiveness of male vaccination and next generation HPV vaccines, and to assess models used to evaluate new cervical screening technologies and recommendations.

1. Background

Prophylactic HPV vaccines protect against infection with HPV types 16 and 18, which are together implicated in 70–80% of invasive cervical cancers. Vaccination of preadolescent females against HPV has been evaluated and implemented in the majority of developed countries, including Australia, New Zealand, Canada, the United States, and most of the European Union Countries [1,2]. The decision to introduce HPV vaccination in these countries has been supported by favourable cost-effectiveness evaluations produced by various decision analytic models [3–9]. In general, the purpose

* Corresponding author at: Level 4, Lowy Cancer Research Centre, The University of New South Wales, Kensington, NSW 2052, Australia. Tel.: +61 2 9385 3817; fax: +61 2 9385 1430.
E-mail addresses: k.canfell@unsw.edu.au, l.testa@unsw.edu.au (K. Canfell).
[1] Past affiliation: Sydney School of Public Heath, University of Sydney, Sydney, NSW, Australia.
[2] Past affiliation: Cancer Research Division, Cancer Council NSW, Woolloomooloo, Australia.

http://dx.doi.org/10.1016/j.vaccine.2014.08.048
0264-410X/

of these models has been to assess the additional 'value add', or increase in effectiveness, that an HPV vaccination programme in females can provide in the context of existing well-established cervical screening mechanisms; this 'value add' includes some level of protection against other HPV-related cancers in females, protection against HPV-related cancers in males (via the effects of herd immunity), the protection provided against anogenital warts for both sexes (in the case of the quadrivalent HPV 1,618,6,11 vaccine (Gardasil, Merck)), and the additional protection, beyond what is provided by cervical screening, against cervical cancer in females. Cervical cancer is associated with the highest burden of disease for any HPV-related cancer, and so the cost-effectiveness of HPV vaccination in females is highly dependent on its effectiveness of further reducing rates of cervical cancer beyond that achieved by cervical screening.

A number of literature reviews have summarised the methods and findings of decision analytic models that have been used to assess the cost-effectiveness of HPV vaccination [3–8,10–13] and a recent systematic review [14] identified over 40 published analyses of the cost-effectiveness of HPV vaccination in females in developed countries. However, despite the fact that cervical screening is a major determinant of the burden of cervical cancer in a particular population, and thus a key factor in the assessment of HPV vaccination cost-effectiveness, none of the reviews performed to date have comprehensively assessed the quality of vaccination cost-effectiveness analyses in relation to their modelling of the cervical screening 'background', nor has any review specifically assessed the screening-related outcomes used to validate these models.

The concept of 'quality' in decision analytic modelling has been previously described [15,16]. The overall intent of a quality framework is to enable a set of standards that ensure models contribute to an optimal decision related to the introduction of a new intervention or health programme [16,17]. A quality framework attempts to address poorly designed models, over-simplification, misinterpretation and failure to appreciate the degree of uncertainty in a decision process [15]. Although generic guidelines for health economic models are available, and a broad consensus statement has been published on model aspects that need to be considered when modelling HPV infections [18–20], neither of these address specific issues in modelling cervical screening and in general terms no guidelines or consensus statements have been published specifically for cervical screening models. Therefore, the aims of the current project were to: (i) adapt existing, broader frameworks to construct a specific quality framework for models of existing cervical screening processes; and (ii) to perform a systematic review of HPV vaccination cost-effectiveness assessments in developed countries and to apply the new quality framework to the assessment of the cervical screening models.

2. Methods

2.1. Quality framework development

A broad quality framework for decision analytic modelling has previously been proposed [17]. This framework incorporates the following criteria: model structure (aspects related to disease state models, treatment options, time horizon, cycle length); quality of data (the choice of input parameters and method of translation as applied into the model); and model consistency (this refers to assessing the internal logic of the modelling exercise, such as modifying model input parameters and noting observed changes with outputs, and external consistency of intermediate and final outputs with observed targets). However, it was acknowledged that the disease area and model design will affect how the quality criteria are applied and that the expectation of universal criteria to evaluate all decision analytic models is impractical [17].

Given the lack of existing specific guidelines on how cervical screening should be modelled and validated, we developed a new quality framework to assess models of cervical screening. The framework was informed by two sources. The first was the International Agency for Research on Cancer (IARC) recommended set of programme indicators to monitor cervical screening effectiveness [21]. The second was a series of guidelines on the recommended management of abnormal cervical disease developed for Australia, New Zealand and England [22–24]. The design of the framework was constrained such that its elements reflected aspects of cervical screening (health outcomes and participation) that were likely to affect a program's effectiveness and/or its cost. The cervical screening programme indicators developed by IARC generally measure population-based behaviour related to cervical screening participation, as well as cervical precancerous and cervical cancer outcomes over the short term. The indicators broadly include: overall and age-specific participation within recommended screening intervals; compliance with follow-up of abnormal cytology and histologically confirmed outcomes; follow-up of women with unsatisfactory smears; compliance with recommended treatment; the number of interval cancers; the distribution of stage of diagnosed cancers; and level of over-screening [21]. The management guidelines for Australia, New Zealand and England provide a series of recommendations for management of cervical precancerous lesions (predicted by cytology and histologically confirmed). These guidelines were used to inform the differential management processes for low and high grade cervical precancerous lesions predicted by cytology, diagnosis and follow-up of abnormal cytology, and management of histologically confirmed disease (treated [CIN 2/3] and untreated [CIN 1]). While there are other national guidelines on how precancerous lesions should be managed, the guidelines selected were considered, in combination, as likely to be reflective of the evidence for best practice for organised cytological screening programmes. The conceptual relationship between the IARC screening performance indicators, and management guidelines are presented in Fig. 1. Based on these sources, four key aspects of organised screening were identified as necessary for comprehensive modelling of the observed effectiveness of an organised programme. These were:

2.1.1. Routine screening behaviour

A model should aim to reflect screening patterns as observed in the 'real world'. This would therefore include women who regularly attended screening at the recommended interval. However, there are women who also over-screen (or re-screen early) and those who never-attend screening. The latter group are most likely to reflect women at greatest risk of developing cervical cancer [25,26]. In addition, unsatisfactory screening tests should be incorporated as these reflect the quality of smear taking and the increased costs associated with rescreening.

2.1.2. Abnormal smear management processes

A model should also accurately reflect the test characteristics of cytology and the processes of referral for follow-up or diagnosis after an abnormal smear result (abnormal cytology management). A model needs to be able to simulate, in a way that reflects the local performance of cytology, the probability of detecting each stage of the underlying cervical precancerous spectrum with cytology. Additionally, the model structure needs to differentiate how each grade of abnormal smear is managed.

2.1.3. Diagnostic processes

A model should have a structure that takes into account diagnostic processes including colposcopy and biopsy. Moreover, given

Please cite this article in press as: 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 (2014), http://dx.doi.org/10.1016/j.vaccine.2014.08.048

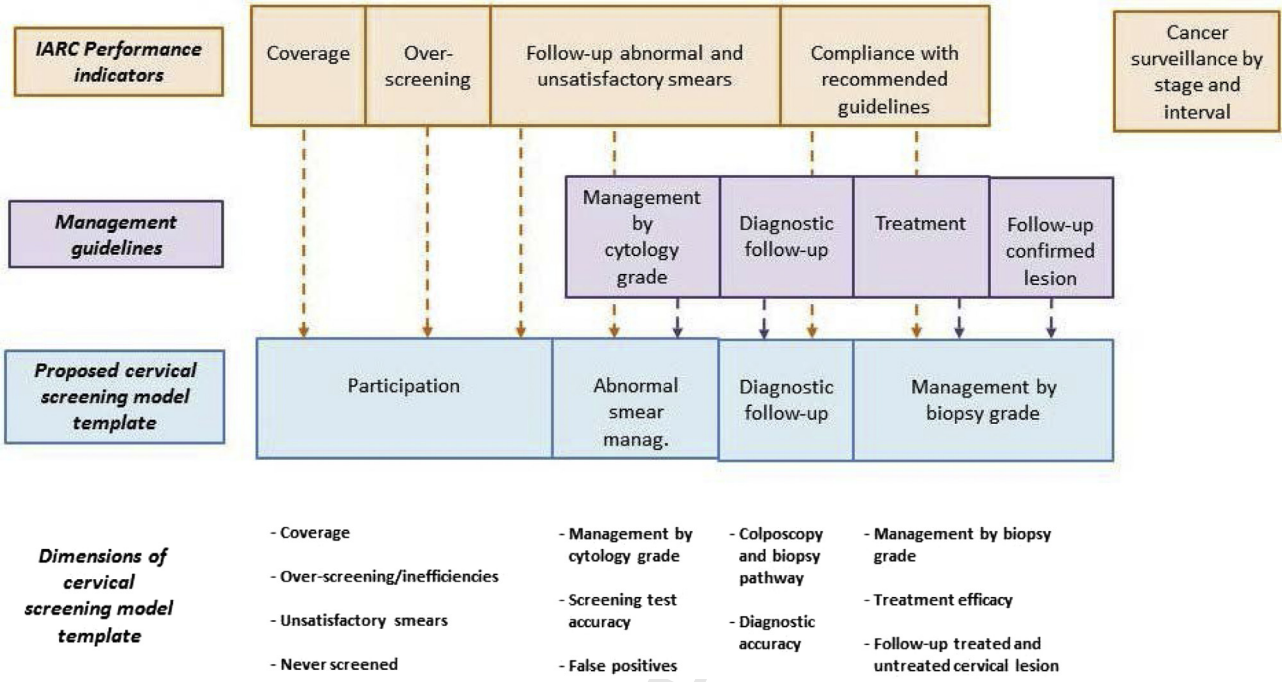

Fig. 1. Conceptual development of framework for the evaluation.

colposcopy is an imperfect diagnostic procedure, a model should reflect the test-characteristics of colposcopy.

2.1.4. Management of abnormalities

The final area a model should reflect is the different management strategies for the grades of cervical intraepithelial neoplasia (CIN) that are histologically confirmed. Thus CIN 1 is unlikely to be treated in most cases, and require follow-up in 6–12 months in many settings. By contrast women with CIN 2/3 will generally be referred for treatment. In addition, women previously treated with CIN 2/3 are at increased risk of developing cervical cancer compared with the rest of the screening population [27]. Therefore a model should reflect follow-up of treated cases as well as separate natural history for the possibility of recurrent disease after treatment for precancer.

Within each of these areas of the quality framework, a specific set of dimensions were defined for quality assessment, defined by the quality criteria (structure, data, consistency/validation) and screening framework (participation, abnormal smear management, diagnostic follow-up of abnormal smear, management by CIN grade) (Table 1). To apply the quality framework, the criteria were used to identify the how each component of the framework was modelled. For the model consistency assessment, real-world outcomes, or 'validation targets' were broadly categorised as: cervical screening-related targets, which reflects intermediate outcomes including screen-detected CIN; cervical cancer targets, and HPV targets, which relate to the assessment of HPV infection rates in either cervical screening or cervical cancer outcomes. Validation against each of these domains should be attempted to ensure the most accurate model to predict the impact of HPV vaccination in the context of HPV vaccination.

2.2. Systematic review of the literature for HPV vaccination cost-effectiveness evaluations

Published scientific papers were included for review if they met the following criteria. Firstly, models had to assess the incremental cost-effectiveness ratio (ICER) of HPV vaccination with cervical screening compared to cervical screening alone. Secondly, the evaluation had to be performed in a setting with an organised screening programme; either a national or regional programme (including pilot organised programmes). Thirdly, models needed to take into account screening behaviour in a 'real world' situation in which women are not perfectly compliant to recommendations; studies were excluded if they modelled prescribed or normative screening protocols only and did not take into account any variability in screening participation. (Note that some studies predicted outcomes only in women who were screened and/or vaccinated, but such evaluations are not directly informative to population based decision-making and therefore these studies were excluded from the current evaluation on a priori grounds.)

Countries with organised screening programmes were identified from IARC Cervix Cancer Screening Handbook: Volume 10, and the European Cancer Observatory website [21,28]. A total of 20 countries were identified as having an established, or rolling out, an organised cervical screening programme at either a national or regional level (Australia, Canada, Denmark, Estonia, Finland, Hungary, Iceland, Ireland, Italy, Netherlands, New Zealand, Norway, Poland, Portugal, Slovenia, Spain, Sweden, United Kingdom,), or were in the process of piloting an organised programme in specific provinces (France, Romania). Articles published up to December 31, 2010 were included in PUBMED and Medline searches. Key terms used to identify relevant articles included a combined search of: 'vaccine*', 'vaccination', 'human papilloma*', 'HPV', 'cervical cancer', 'cervix', and 'papilloma'. Results based on these search terms were combined with the 'AND' term with the names of countries identified as having an organised screening programme.

3. Results

3.1. Systematic review of the literature

A total of 742 potential citations were identified from the initial search. Of these, 67 evaluated the impact of HPV vaccination. Of the 67 studies, 32 publications were identified as evaluating the ICER (using either quality-adjusted life-years or life years saved)

Please cite this article in press as: 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 (2014), http://dx.doi.org/10.1016/j.vaccine.2014.08.048

Country	Number of hits (n)	Citations assessing HPV vaccination effects (n)	Included for review (n)
Hungary	6	2	1
Poland	17	1	1
Romania	3	0	0
Denmark	20	1	1
Estonia	3	0	0
Finland	32	4	2
Iceland	12	2	1
Ireland	15	4	4
Norway	12	1	1
Sweden	27	1	0
United Kingdom	186	9	4
Italy	78	9	4
Portugal	5	0	-
Slovenia	8	1	1
Spain	33	3	1
France	48	4	3
The Netherlands	38	11	4
Australia	101	7	1
New Zealand	24	0	-
Canada	74	7	3
Totals	**742**	**67**	**32**

Main reason for exclusion*	n
Commentary	3
No modelling	3
No screening modelled	8
New normative screening protocols modelled	6
Review	5
Not in English	5
Vaccine cost /uptake only	2
Cost-effectiveness not assessed	3

*Studies could have multiple reasons for exclusions. These were the main ones identified.

Fig. 2. Flowchart describing PUBMED search results.

for HPV vaccination in the context of organised screening. The reasons for excluding publications were: papers were commentary/editorials ($n = 3$); no modelling performed ($n = 3$); no screening modelling performed ($n = 8$); modelled normative screening protocols only ($n = 6$); review studies ($n = 5$); not published in English ($n = 5$); assessed vaccination costs or estimated uptake only ($n = 2$); or did not assess the cost-effectiveness of HPV vaccination (even though a simulation of HPV vaccination in the context of an organised screening programme was performed) ($n = 3$) (Fig. 2).

The 32 published HPV cost-effectiveness evaluations identified for review were undertaken within the context of 15 countries with organised cervical screening; as follows: Australia [29], Canada (3 published cost-effectiveness evaluations) [30–32], Denmark [33], Finland (2 evaluations) [34,35], France (3 evaluations) [36–38], Hungary [39], Iceland [40], Ireland (4 evaluations) [34,38,41,42], Italy (4 evaluations) [32,38,43,44], the Netherlands (4 evaluations) [45–48], Norway [49], Poland [34], Slovenia [50], Spain [51], and the United Kingdom (4 evaluations) [32,52–55]. Of the 32 identified published studies (Table 2), 'duplicate' (defined here as the same or similar organised screening model) cost-effectiveness evaluations of HPV vaccination in the context of organised screening were identified for France (2 evaluations) (Demarteau et al. [37,38]) and Ireland (2 evaluations) (Suarez et al. [34] and Demarteau et al. [38]). These models were combined into one, leaving a total of 30 models that were evaluated using the predefined quality criteria and screening framework.

Of the 30 models identified for the review, 4 studies referenced complementary papers to describe their screening or natural history models. Brisson et al. [30] referenced Van de Velde et al. [56] for the HPV cost-effectiveness evaluation done in Canada; Jit et al. [54] referenced Choi et al. [53] for their evaluation done for the United Kingdom; Debicki et al. [32] referenced Kohli et al. [55] for the evaluation done in United Kingdom; while Boot et al. [48] referenced Berkhof et al. [57] for the evaluation done in the Netherlands. The complementary publications were also used to inform the assessment of reviewed organised screening models.

A number of the published models included in this review were adapted from previously developed models, used either for a prior HPV vaccination cost-effectiveness evaluation, or originally developed for cervical screening evaluations. The specific mechanisms for how these models are adapted are not reported in detail. In general, for example, Kulasingam et al. evaluated the cost-effectiveness of HPV vaccination in the context of organised screening in Australia [29], and the United Kingdom [52]. Her model was adapted from a Markov model of cervical screening and natural history developed by Myers et al. [58]. The same model developed by Myers et al. was adapted for a cost-effectiveness evaluation of HPV vaccination by Bergeron et al. for France [36], Obradovic et al. for Slovenia [50], and Mennini et al. for Italy [44]. Similarly, a Markov model developed by Goldie et al. [59] in the United States was adapted for a cost-effectiveness evaluation of HPV vaccination for Italy by La Torre et al. [43], and for Spain

Please cite this article in press as: 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 (2014), http://dx.doi.org/10.1016/j.vaccine.2014.08.048

Table 1
Q7 Data items extracted for assessment of cervical screening models.

Structural item	Collected data
Routine screening behaviour	- Screening age-group
	- Modelled interval - Proportion attending (or rate) screening for each age group at specified interval cycles - Proportion attending (or rate) screening beyond screening intervals (over-screening) - Proportion unsatisfactory smears
Abnormal smear management	- Distinction of cytology grade
	- Separate management of cytology grade - Sensitivity and specificity values of cytology
Diagnostic follow-up of abnormal cytology	- Indication of diagnostic follow-up
	- Distinct colposcopy and biopsy health states - Sensitivity and specificity values of colposcopy
Management by CIN grade	- Indication of CIN management
	- Distinct CIN management approaches - Treatment efficacy of treated CIN - Follow-up of treated disease
Data source	Collected data
Sources of information for screening structure and parameterisation[a]	- Guidelines
	- Survey or surveillance - Expert opinion/Personal communication - Assumption
Consistency	Collected data
Cervical cancer	- Cervical cancer incidence (age-specific or overall) - Cervical cancer mortality (age-specific overall)
Cervical pre-cancers and screening related activity	- Age-specific or overall histologically confirmed CIN prevalence
	- Age-specific or overall abnormal cytology grades (low grade or high grade) rates - Age-specific or overall smears
HPV prevalence	- Age-specific or overall prevalence/incidence of oncogenic HPV in: Normal cytology Abnormal cytology or histologically confirmed CIN Cervical cancer

[a] More than one item can be recorded for a study.

by Gauthier et al. [51]. The most extensively adapted model for a cost-effectiveness evaluation of HPV vaccination identified in this review was developed by Debicki et al. [32] Debicki et al. originally adapted a Markov model that was developed by Kohli et al. [55]. This model was used to assess the predicted health outcomes (not cost-effectiveness) associated with the introduction of HPV vaccination in the United Kingdom, in the context of an organised screening programme. Debicki et al. used this model for a cost-effectiveness evaluation of HPV vaccination within the context of the United Kingdom, Canada and Italy [32]. The same model was adapted by Debicki et al. to develop a simplified/succinct screening and natural history model. The simplified model developed by Debicki et al. was used by Anonychuck et al. for a Canadian evaluation [31], Suarez et al. for evaluations in Finland, Ireland and Poland [34], Torvinen et al. for an evaluation in Finland (Torvinen modified the screening model parameters) [35], and Demarteau et al. for an evaluation in France and Italy [38].

In addition to these, a dynamic transmission model included in this review was developed by Dasbach et al. for a health economic evaluation of HPV vaccination in the United States [60–62]. This model was applied to cost-effectiveness evaluations of HPV vaccination in Norway [49], Hungry [39], and the United Kingdom [63]. Another dynamic transmission model developed by Garnett et al. [64] was adapted by Jit et al. and Choi et al. for a cost-effectiveness evaluation of HPV vaccination in the United Kingdom [53,54].

3.2. Quality assessment for routine screening

Table 3 summarises how each of the reviewed models addressed the four pre-defined dimensions associated with routine screening behaviour.

3.2.1. Structure

Of the 30 models reviewed, none modelled all four dimensions. Twenty-eight models specified screening coverage, while 2 models did not (Table 3) [40,43]. Of the 28, 13 reported coverage for all targeted participants over a screening interval period (3 or 5 years) [31,33–35,37,41,42,45,46,48,65] 9 reported age-specific proportions over a screening interval period [29,32,34,36,44,50–52] while 6 described an age-specific yearly rate or annual proportion [30,39,49,53,63] (see supplementary material for details). Of the 30 country specific models, 15 indicated a proportion of women who never attended screening [29,31,32,35,39,42,44–46,48,51,63,65] while over-screening was outline by 6 models [29,31,34,35] and unsatisfactory smears were modelled by 2 [29,36] (see supplementary material for details).

3.2.2. Data

A total of 20 of the 30 models indicated the use of survey data or surveillance data to inform their coverage estimates [30–32,34–37,39,42,44,48,50–53,63]. Countries that were piloting an organised programme (Ireland, Poland) utilised expert opinion [34]. Data sources for never-screened women were largely local population based study data, while those who modelled screening inefficiency either based models on both surveillance and expert opinion.

3.2.3. Consistency

Models for Finland's organised screening programme developed by Suarez et al. [34] and Torvinen et al. [35] assessed consistency of routine screening patterns through calibrating to total number of smears (Table 8). No other identified decision analytic model assessed whether routine screening patterns were consistent with real world behaviour. The prevalence of HPV in normal cytology was assessed for consistency in 11 models (Table 8) [30,32,44–48,51,52]. The majority of these model validation targets were age-specific and reported the prevalence of HPV for any virus type (oncogenic and non-oncogenic types).

3.3. Quality assessment for management of abnormal cytology

Table 4 is a summary table that describes how a cervical precancerous lesion, predicted by cytology, was modelled, while Table 5 presents a detailed description of modelled cytology test characteristics for each of the reviewed models.

3.3.1. Structure

Of the 30 models, 16 differentiated the cytology prediction of low and high grade cervical precancerous lesion, along with management strategies for each (Table 5) [29,32,36,40,41,43–46,48,50–52,65]. In these models, cytology predictive of low grade disease was triaged to either follow-up with another pap (or liquid based cytology), HPV testing, or

Please cite this article in press as: 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 (2014), http://dx.doi.org/10.1016/j.vaccine.2014.08.048

Table 2
Q8 Summary of included studies in review.

Country	Cervical screening programme			Number of models reviewed	Authors	Model type
	Organisation	Target population	Screening interval			
Australia	Population based-nationwide	18/21–70 years	2 yearly	1	Kulasingam et al. (28)	Markov
Canada	Population based – Region specific	18–69	1–3 yearly (province specific)	3	Debicki et al. (31) Anonychuk et al. (30) Brisson et al. (29,55)[a]	Markov Markov Markov
Denmark	Population based – Nationwide	23–59	3 yearly	1	Olsen et al. (32)	Hybrid
Finland	Population based – Nationwide	30–60	5 yearly	2	Suarez et al. (33) Torvinen et al. (34)	Markov Markov
France	Population based – Regional pilot	25–65	3 yearly	3	Bergeron et al. (35) Demarteau et al. (36)[b] Demarteau et al. (37)[b]	Markov Markov Markov
Hungary	Population based – Nationwide	25–65	3 yearly	1	Dasbach et al. (38)	Dynamic
Iceland	Population based – Nationwide	20–69	2–3 yearly	1	Oddsson (39)	Markov
Ireland	Population based – Nationwide	25–60	25–44 years 3 yearly 45–60 years 5 yearly	4	Suarez et al. (33)[c] Demarteau et al. (37)[c] Dee et al. (41) Usher et al. (40)	Markov Markov Markov Hybrid
Italy	Population based – Nationwide (rolling out)	25–64	3 yearly	4	Debicki et al. (31) Demarteau et al. (37) La Torre et al. (42) Mennini et al. (43)	Markov Markov Markov Markov
Netherlands	Population based – Nationwide	30–60	5 yearly	4	Coupe et al. (45) de Kok et al. (44) Rogoza et al. (46) Boot et al. (47,56)[d]	Markov Micro Markov Markov
Norway	Population based – Nationwide	25–69	3 yearly	1	Dasbach et al. (48)	Dynamic
Poland	Population based – Nationwide	25–59	3 yearly	1	Suarez et al. (33)	Markov
Slovenia	Population based – Nationwide	20–64	3 yearly	1	Obradovic et al. (49)	Markov
Spain	Population based – Region specific	25–65	3 yearly	1	Gauthier et al. (50)	Markov
United Kingdom	Population based – Nationwide	25–64	25–49 3 yearly 50–64 5 yearly	4	Kulasingam et al. (51) Jit et al. (52,53)[e] Dasbach et al. (62) Debicki et al. (31,54)[f]	Markov Hybrid Dynamic Markov

[a] Includes the complementary publication by Van de Velde et al. (57).
[b] Demarteau et al. models broadly adopt the same screening model and cost for France. These are combined into one analysis: Demarteau et al. – France.
[c] Suarez et al. and Demarteau models broadly adopt the same screening model and costs for Ireland. These are combined into one analysis: Suarez et al. – Ireland.
[d] Includes the complementary publication by Berkhof et al. (58).
[e] Includes the complementary publication by Choi et al. (54).
[f] Includes the complementary publication by Kohli et al. (56).

immediate colposcopy. In contrast, cytology predictive of high grade disease was modelled as immediate referral to colposcopy. However, models developed by Brisson et al. for Canada [30,56], Olsen for Denmark [33], Dasbach et al. for Hungary [39], Norway [49], the United Kingdom [63], and Jit et al. for the United Kingdom [53,54], alluded to differentiating the detection of cervical precancerous lesions by cytology, however the management strategies were not clearly specified. By contrast, 8 models modelled a cytology prediction of low or high grade precancerous lesion as one cytology grade [31,34,35,37,42]. In these models, an abnormal cytology result was referred to colposcopy.

The reporting of cytology test characteristics was varied. Of the 16 models that separated the detection of precancerous lesions predicted by cytology into low and high grade disease, 3 reported a test sensitivity value for a cytology threshold (Table 6). This was done by Kulasingam et al. for Australia [29], Boot et al. for the Netherlands [48,57], and Obradovic et al. for Slovenia [50]. In the remaining 12 of the 16 models that specified a separate structural component for the management of low and high grade disease, a sensitivity value for grades of precancerous lesion detected by cytology was not reported.

The modelled (or assumed) cytology sensitivity (defined as normal/abnormal or grade of cytological abnormality) for the detection of underlying cervical precancerous lesions, by CIN grade, was quoted in 22 of the 30 models (Table 6) [29–32,34–36,40,42–46,48,50–52,56,57,63,65]. The assumed

Please cite this article in press as: 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 (2014), http://dx.doi.org/10.1016/j.vaccine.2014.08.048

Table 3
Modelling approach to routine screening patterns.

Author	Country	Reported model coverage?	Modelled never screened women?	Modelled screening inefficiency/over-screening?	Modelled unsatisfactory smears?
Kulasingam et al.	Australia	Yes	Yes	Yes	NI
Debicki et al.	Canada	Yes	Yes	NI	NI
Anonychuk et al.	Canada	Yes	Yes	Yes	NI
Brisson et al.[a]	Canada	Yes	NI	NI	NI
Olsen et al.	Denmark	Yes	NI	NI	NI
Suarez et al.	Finland	Yes	NI	Yes	NI
Torvinen et al.	Finland	Yes	Yes	Yes	NI
Bergeron et al.	France	Yes	NI	NI	Yes
Demarteau et al.	France	Yes	NI	NI	NI
Dasbach et al.	Hungary	Yes	Yes	NI	NI
Oddsson	Iceland	NI	NI	NI	NI
Suarez et al.	Ireland	Yes	NI	Yes	NI
Usher et al.	Ireland	Yes	NI	NI	NI
Dee et al.	Ireland	Yes	Yes	NI	NI
Debicki et al.	Italy	Yes	Yes	NI	NI
Demarteau et al.	Italy	Yes	NA	NI	NI
La Torre et al.	Italy	NI	NI	NI	NI
Mennini et al.	Italy	Yes	Yes	NI	NI
Coupe et al.	Netherlands	Yes	Yes	NI	NI
de Kok et al.	Netherlands	Yes	Yes	NI	NI
Rogoza et al.	Netherlands	Yes	Yes	No	NI
Boot et al.[b]	Netherlands	Yes	Yes	NI	NI
Dasbach et al.	Norway	Yes	NI	NI	NI
Suarez et al.	Poland	Yes	NI	Yes	NI
Obradovic et al.[c]	Slovenia	Yes	NI	NI	NI
Gauthier et al.[d]	Spain	Yes	Yes	NI	NI
Kulasingam et al.	United Kingdom	Yes	NI	NI	Yes
Jit et al.[e]	United Kingdom	Yes	NI	NI	NI
Dasbach et al.	United Kingdom	Yes	Yes	NI	NI
Debicki et al.[f]	United Kingdom	Yes	Yes	NI	NI

NI, not indicated.
[a] Reviewed Van de Velde et al. (55) in conjunction with Brisson et al. to assist with model interpretation.
[b] Reviewed Berkhof et al. (56) in conjunction with Boot et al. to assist with model interpretation.
[c] Reviewed Myers et al. (57) in conjunction with Obradovic et al. to assist with model interpretation.
[d] Reviewed Goldie et al. (58) in conjunction with Gauthier et al. to assist with model interpretation.
[e] Reviewed Choi et al. (52) in conjunction with Jit et al. to assist with model interpretation.
[f] Reviewed Kohli et al. (54) in conjunction with Debicki et al. to assist with model interpretation.

cytology sensitivity for the detection of underlying CIN 1 disease ranged from 29%, as reported by Dasbach et al. for the United Kingdom [63], to 68% as reported by Debicki et al. for Canada [32]. For the detection of underlying CIN 2/3 disease, assumed cytology sensitivity ranged from 59%, as reported by Dasbach et al. for the United Kingdom [63], to 80% for the Netherlands, as reported by Coupe et al. [46], de Kok et al. [45] and Boot et al. [48]. In terms of cytology test specificity, 13 of the 31 country-specific models reported any value (Table 6). These ranged from 94% by Dasbach for the United Kingdom [63], to 98% for the Netherlands as reported by Coupe et al. [46], de Kok et al. [45] and Boot et al. [48].

3.3.2. Data

For each of the reviewed models, guidelines were cited from each country to inform how participants with a cervical precancerous lesion detected by cytology would be managed. To inform the proportion of low grade cervical precancerous lesions detected by cytology to triage (either HPV, repeat smear or colposcopy), study data was used by Bergeron et al. for France [36], Usher et al. for Ireland [41], Debicki et al. for Italy [32], Rogoza et al. for the Netherlands [65], Gauthier et al. for Spain [51], and Kohli et al. for the United Kingdom [55].

For most of the models, data sources used modelled test characteristics of cytology derived from two systematic reviews. These were Nanda et al. [66] and Fahey et al. [67] Where HPV testing was applied for triage of low grade smears, Cuzick et al.'s [68] review was used as the main source. The exception was models that assessed the cost-effectiveness of HPV vaccination in the context of the Netherlands used local data from trials (directly or indirectly) to model assumed cytology test sensitivity and specificity [69,70].

3.3.3. Consistency

For the assessment of model consistency with cervical precancerous lesions predicted by cytology, of the 30 models, only two – Boot et al. for the Netherlands [48,57], and Debicki et al. for the United Kingdom [32,55], included cytological predictions of precancerous lesions as a validation target (Table 8). Of the remaining 28 models, 11 assessed model consistency of detecting cervical precancerous lesions based on comparisons with observed CIN (Table) [29–33,41,45,46,48,49,51,55]. However, in most cases it was not possible to determine from the reports whether the reported results for cervical precancerous lesions were predictions of cytology results or alternatively, represented histologically confirmed lesions; nor was it generally clear whether or not an appropriate screen-detection and diagnostic overlay was in place to simulate histological confirmed lesions. Overall type-specific HPV prevalence was reported in LSIL and HSIL as a model validation target by Debicki et al. for the United Kingdom [32,55].

3.4. Quality assessment for diagnostic processes

Table 5 reports on the diagnostic accuracy of colposcopy published by the reviewed models and Table 6 is a summary table describing how precancerous lesions detected by cytology were managed at colposcopy.

Please cite this article in press as: 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 (2014), http://dx.doi.org/10.1016/j.vaccine.2014.08.048

Table 4
Modelling the management of abnormal cytology.

Author	Country	Model structure differentiates between low and high grade cytology	Low grade smear Triage
Kulasingam et al.	Australia	Yes	Uncertain
Debicki et al.	Canada	Yes	Yes
Anonychuk et al.	Canada	No	No
Brisson et al.[a]	Canada	NI	NI
Olsen et al.	Denmark	NI	NI
Suarez et al.	Finland	No	No
Torvinen et al.	Finland	No	No
Bergeron et al.	France	Yes	Yes
Demarteau et al.	France	No	No
Dasbach et al.	Hungary	NI	NI
Oddsson	Iceland	Yes	Yes
Suarez et al.	Ireland	No	No
Usher et al.	Ireland	Yes	Yes
Dee et al.	Ireland	No	No
Debicki et al.	Italy	Yes	Yes
Demarteau et al.	Italy	No	No
La Torre et al.	Italy	Yes	Yes
Mennini et al.	Italy	Yes	Yes
Coupe et al.	Netherlands	Yes	Yes
de Kok et al.	Netherlands	Yes	Yes
Rogoza et al.	Netherlands	Yes	Yes
Boot et al.[b]	Netherlands	Yes	Yes
Dasbach et al.	Norway	NI	NI
Suarez et al.	Poland	No	No
Obradovic et al.[c]	Slovenia	Yes	Yes
Gauthier et al.[d]	Spain	Yes	Yes
Kulasingam et al.	United Kingdom	Yes	Yes
Jit et al.[e]	United Kingdom	NI	NI
Dasbach et al.	United Kingdom	NI	NI
Debickiet al.[f]	United Kingdom	Yes	Yes

NI not indicated.
[a] Reviewed Van de Velde et al. (55) in conjunction with Brisson et al. to assist with model interpretation.
[b] Reviewed Berkhof et al. (56) in conjunction with Boot et al. to assist with model interpretation.
[c] Reviewed Myers et al. (57) in conjunction with Obradovic et al. to assist with model interpretation.
[d] Reviewed Goldie et al. (58) in conjunction with Gauthier et al. to assist with model interpretation.
[e] Reviewed Choi et al. (52) in conjunction with Jit et al. to assist with model interpretation.
[f] Reviewed Kohli et al. (54) in conjunction with Debicki et al. to assist with model interpretation.

3.4.1. Structure

Of the 30 models, 25 specified diagnostic follow-up of a referred cervical precancerous lesions detected by cytology (Table 7). Of these, 8 models differentiated colposcopy and biopsy into separate health states [32,36,50,51,53,55,65], while 16 combined colposcopy investigation and biopsy sampling into one [31,34,35,37,39,41–46,52]. Colposcopy was modelled in 8 models as 100% accurate by Kulasingam et al. for Australia [29], Anonychuck et al. for Canada [31], Suarez et al. and Torvinen et al. for Finland [34,35], Demarteau et al. for France and Italy [37], and Suarez et al. for Ireland and Poland [34]. Of the 9 models that modelled less than prefect colposcopy accuracy, these ranged from 82% as modelled by Choi et al. for the United Kingdom [53], to 96% as reported by Rogoza et al. for the Netherlands [65], Gauthier et al. for Spain [51], Dasbach et al. and Kohli et al. for the United Kingdom [55,63]. Colposcopy specificity was reported by 8 models, with values that were either 48% for 5 models [44,51,55,63,65], or 100% for 3 models [34,36,52].

3.4.2. Data

No data source was indicated to support the combination of colposcopy and biopsy into one health state. The proportion of abnormal cytology cases referred for colposcopy was reported by 4 of the 25 models that reported diagnostic follow-up. This was done by Bergeron et al. for France [36], Usher et al. for Ireland [41,65], Gauthier et al. for Spain [51], and Choi et al. for the United Kingdom [53] (data no reported). Otherwise, perfect compliance was assumed with a referral to colposcopy. For models that reported imperfect colposcopy sensitivity and specificity (see Table 6) the main source for the test-characteristics of colposcopy were the meta-analyses published by Michell et al. [71] and Hopman et al. [72].

3.4.3. Consistency

The choice of model validation for cervical cancerous lesions detected by cytology and referred for colposcopy follow-up was limited to the broader measure of observed CIN (age-specific and overall). Of the 30 models reviewed, 13 assessed model validation with observed CIN (Table 8) [29–33,41,45,46,48,49,51,55]. However as pointed out in Section 3.3.4, it was difficult to determine whether the CIN terminology referred to screen detected CIN or histologically confirmed CIN. For example, The model adapted by Debicki et al., Kohli et al. [55] reported CIN 1 validation which referred to cytology detected, while CIN 2/3 was referred to as histologically confirmed. Model predicted HPV prevalence in CIN (either cytology or histologically confirmed) was assess by Brisson et al. for Canada [30], Gauthier et al. for Spain [51], and Kohli et al. for the United Kingdom [55] (Table 8). This was predominantly in the form of overall type-specific HPV prevalence in CIN 1, 2 or 3. No studies reported age specific infection patterns of HPV (any or type-specific) in CIN.

3.5. Quality assessment for management of CIN

Table 7 summarises the how detected CIN was managed in the reviewed models of organised screening, and Table 8 describes consistency/validation assessment.

3.5.1. Structure

Of the 30 country-specific models, 22 were identified as modelling separate management strategies for detected CIN 1 and CIN 2/3 (Table 7) [30,32–37,39,41,43,47,49–55,63]. It was assumed, however, that all 30 models treated histologically confirmed CIN 2/3. Of those who modelled treatment for some histologically confirmed CIN 1, there was variability within countries on the proportion treated (data not reported). For Finland, this ranged from 0% by Suarez et al. [34] to 70% by Torvinen et al. [35]; for France this ranged from 50% by Demarteau et al. [37] to 75% by Bergeron et al. [36]; for Ireland the range was 0% by Suarez et al. [34] to 50% by Demarateau et al. [37]; and in the United Kingdom, 20% by Kulasingam et al. [52] to 50% by Kohli et al. [55]. In terms of treatment efficacy of CIN (any grade), 6 modelled less than 100% efficacy [30,34,37,41,55], while 7 modelled perfect treatment [32,33,36,37,43,50,52–54]. Finally, follow-up of histologically confirmed CIN was indicated for 9 of the reviewed models [32,34,35,37,40,41,50]. However, the detail on how this was done was not discernible for any of the reviewed models.

3.5.2. Data

Of the 22 models identified as modelling distinct strategies for the management of CIN, 6 utilised survey data to inform the proportion of treated CIN 1 [32,36,37,47,50,51]. With respect to treatment efficacy and follow-up, the model by Obradovic et al. utilised survey data for Slovenia [50]. 12 of the reviewed models based the model

Please cite this article in press as: 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 (2014), http://dx.doi.org/10.1016/j.vaccine.2014.08.048

Table 5
Screening and colposcopy test characteristics for each reviewed model.

Author	Country	Pap test						Colposcopy					
		Sensitivity			Specificity			Sensitivity			Specificity		
		Cytology result	Histology	%	Cytology result	Histology	%	Colposcopy result	Histology	%	Colposcopy result	Histology	%
Kulasingam et al.	Australia	≥ASCUS	≥CIN 1	80	≥ASCUS	≥CIN 1	95	Positive	Any CIN	100	NA[b]	–	–
Debicki et al.	Canada	NS[c]	NS[c]	68–81	NS	–	–	NS[d]	–	–	NS[d]	–	–
Anonychuk et al.	Canada	Abnormal	≥CIN 1	55.4	Abnormal	≥CIN 1	96.8	Positive	Any CIN	100	NA[b]		
Brisson et al.‡	Canada	NS[e]	CIN 1 CIN 2/3	60 78	NS	–	–	NS	–	–	NS	–	–
Olsen et al.	Denmark	NS	–	–	NS	–	–	NS	–	–	NS	–	–
Suarez et al.	Finland	Abnormal[g]	CIN 1 CIN 2/3	65 81	NS	–	–	Positive[f]	Any CIN	100	NA[b]	–	–
Torvinen et al.	Finland	NS[e]	CIN 1 CIN 2/3	65 81	NS	–	–	Positive[f]	Any CIN	100	NA[b]	–	–
Bergeron et al.	France	NS[e]	CIN 1 CIN 2/3	63 61	NS	>CIN 1	95.7	Positive[f]	Any CIN	90	NA[b]	Any CIN	100
Demarteau et al.	France	NS[h]	NS	58–61	NS[h]	–	–	Positive[f]	Any CIN	100	NA[b]	–	–
Dasbach et al.	Hungary	NS[i]	–	–	NS[i]	–	–	NS[i]	–	–	NS[i]	–	–
Oddsson	Iceland	NS[i]	–	–	NS[i]	–	–	NS[i]	–	–	NS[i]	–	–
Suarez et al.	Ireland	Abnormal	CIN 1 CIN 2/3	58 61	NS	–	–	Positive[f]	Any CIN	100	NA[b]	–	–
Usher et al.	Ireland	NS[j]	–	–	NS[j]	–	–	NS[i]	–	–	NS[i]	–	–
Dee et al.	Ireland	NS[c]	NS[c]	41–67	NS	–	–	NS[d]	–	–	NS[d]	–	–
Debicki et al.	Italy	NS[e]	CIN 1 CIN 2/3	58 61	NS	–	–	NS[d]	–	–	NS[d]	–	–
Demarteau et al.	Italy	NS[h]	NS[h]	NS	58–61	NS[h]	–	Positive[f]	Any CIN	100	NA[b]	–	–
La Torre et al.	Italy	NS[i]	–	–	NS[i]	–	–	NS[i]	–	–	NS[i]	–	–
Mennini et al.	Italy	NS	NS	61	NS	NS	95.7	Positive	Any CIN	90	Negative	Any CIN	48
Coupe et al.	Netherlands	NS	CIN 0[j] CIN 1[j] CIN 2[j] CIN 3[j]	1.5 40 50 75	NS	–	–	NS	–	–	NS	–	–
de Kok et al.	Netherlands	NS[e]	CIN 1 CIN 2 CIN 3 IA/B II+	50 65 80 85 90	NS	NS	98.5	NS	–	–	NS	–	–
Boot et al.[a]	Netherlands	Borderline or mild dyskarosis	CIN 1 CIN 2 CIN 3	50 65 80	Anyabnormal	NS	98.5	NS	–	–	NS	–	–
Rogoza et al.	Netherlands	NS[c]	NS[c]	40–80	Anyabnormal	NS	98.5	Positive	Any CIN	96	Negative	Any CIN	48
Dasbach et al.	Norway	NS	–	–	NS	–	–	NS	–	–	NS	–	–
Suarez et al.	Poland	Abnormal	CIN 1 CIN 2/3	58 61	NS	–	–	Positive[f]	Any CIN	100	NA[b]		
Obradovic et al.	Slovenia	LG HG	CIN 1 CIN 2/3+	61 65	NS	NS	95.7	Positive	Any CIN	90	Negative	Any CIN	100

Please cite this article in press as: 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 (2014), http://dx.doi.org/10.1016/j.vaccine.2014.08.048

Table 5 (*Continued*)

Author	Country	Pap test						Colposcopy					
		Sensitivity			Specificity			Sensitivity			Specificity		
		Cytology result	Histology	%	Cytology result	Histology	%	Colposcopy result	Histology	%	Colposcopy result	Histology	%
Gauthier et al.	Spain	NS	NS	41–67	NS	NS	96.6	Positive	Any CIN	96	Negative	Any CIN	48
Kulasingam et al.	United Kingdom	NS[e]	CIN 1	61	NS	NS	95.7	Positive	Any CIN	95.8	Negative	Any CIN	100
			CIN 2/3	65									
Jit et al.	United Kingdom	NS	CIN 1	69	NS	CIN 1CIN 2/3	8192	Positive	Any CIN	82	NS	–	–
			CIN 2/3	83									
Dasbach et al.	United Kingdom	NS[k]	CIN 1	29	NS	NS	94	Positive	Any CIN	96	Negative	Any CIN	48
			CIN 2/3	59									
Debicki et al.	United Kingdom	NS	NS	41–67	NS	NS	96.6	Positive	Any CIN	96	Negative	Any CIN	48

NS, not specified; NA, not applicable.
[a] Information obtained from Berkhof et al. (56).
[b] NA, not applicable as test is perfectly accurate in picking true disease, which implies perfect accuracy in detecting no disease.
[c] NS, Not specified – Debicki paper states detail model is able to distinguish between cytology grades for various grades of histology.
[d] NS, Not specified – Debicki paper states detail model is capable of parameterising colposcopy and biopsy inaccuracies, however specific values are not presented.
[e] NS, Not specified – The grade of abnormality for cytology is not specified.
[f] Model structure is based on Debicki et al. succinct model. In this model colposcopy is assumed to be 100% sensitive.
[g] Model is based on Debicki succinct model. Torvinen does not specify cytology result, however, Debicki's paper outlines that the succinct model classifies cytology as abnormal or normal.
[h] NR, not reported – data on sensitivity and specificity for CIN are not published.
[i] NR, not reported – data on sensitivity and specificity for CIN are not published.
[j] Sensitivity is reported as positive smear rate for each grade of histology.
[k] Sensitivity and specificity test results for liquid based cytology.

structure of follow-up of histologically confirmed CIN on guidelines, assumption or cost data [29,32,34,35,37,40,43–46,53,54]. Otherwise, models did not quote the source, or did not model follow-up.

3.5.3. Consistency

No specific validation or calibration target was indicated to specify the accuracy of how CIN was managed (treated/untreated). Similar to the reported model validation targets in Sections 3.3.4 and 3.3.5, observed CIN (age-specific or overall) was used as a broad target to assess model predicted CIN (age-specific or overall). However, it was difficult to determine whether detected CIN referred to screen detected or histologically confirmed cervical precancerous disease.

4. Discussion

The aim of this systematic review was to develop a suitable quality framework and to assess the quality of models of organised screening programmes that have been used to evaluate the cost-effectiveness of HPV vaccination. The quality framework constituted four broad areas: (1) routine screening behaviour; (2) abnormal smear management; (3) diagnostic follow-up of abnormal smears; and (4) management of CIN. Each area was assessed according to model structure, data sources, and model validation. Using this framework, we identified a total of 30 decision analytic models suitable for review that assessed the cost-effectiveness of HPV vaccination in developed countries. In reviewing these studies, ewe found that most evaluations took into account population-level screening participation. However, abnormal smear management, diagnostic follow-up and management of CIN were modelled inconsistently. In some cases, modelling of these aspects was highly simplified. In terms of model validation, a number of models relied on using a small number of targets, while there was also considerable variability in the choice of model validation targets; especially for precancerous abnormalities.

To evaluate the cost-effectiveness of HPV vaccination, consideration needs to be given to the pre-existing levels of disease and the existing costs of cervical cancer prevention in the population. However, these depend on cervical screening recommendations, compliance with those recommendations, the screening test used and follow-up processes. A simple proportionate model would potentially be adequate to inform reports of some intermediate outcomes if care is taken to calibrate all outcomes on which estimates will be reported. A proportionate outcome model can also potentially model reductions in the occurrence of abnormal cytology and pre-cancerous lesions; the main requirement is that a high-quality (preferably nationwide) database of cytology and histology records is used to inform assumptions and that future screening and compliance assumptions are well-documented. However, care should be taken in reporting other intermediate outcomes which may not necessarily be directly proportionate to cervical cancer rates–non-proportionate outcomes might include, for example, numbers of test-of-cure follow-up tests after CIN treatment. In this paper we do not argue against the use of simplified screening models per se, but we propose a framework which allows a rational and structured approach to the evaluation of screening models and any simplifications used, and thus a better and more informed assessment of the validity of the model for the research question at hand.

4.1. Modelled participation

The majority of the models included in the review reported that they had modelled some form of age-specific participation within the target age group for screening. However, the dimensions of

Please cite this article in press as: 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 (2014), http://dx.doi.org/10.1016/j.vaccine.2014.08.048

Table 6
Modelling diagnostic follow-up of colposcopy for referred abnormal cytology.

Author	Country	Colposcopy and biopsy assessment combined or separated?	Colposcopy accuracy	
			Sensitivity	Specificity
Kulasingam et al.	Australia	Uncertain	100	NI
Debicki et al.	Canada	Separate	NI	
Anonychuk et al.	Canada	Combined	100	NI
Brisson et al.[a]	Canada	Uncertain	NI	
Olsen et al.	Denmark	Not modelled	NI	
Suarez et al.	Finland	Combined	100	NI
Torvinen et al.	Finland	Combined	100	NI
Bergeron et al.	France	Separate	90	100
Demarteau et al.	France	Combined	100	NI
Dasbach et al.	Hungary	Combined	NI	
Oddsson	Iceland	Combined	NI	
Suarez et al.	Ireland	Combined	100	
Usher et al.	Ireland	Combined	NI	
Dee et al.	Ireland	Combined	NI	
Debicki et al.	Italy	Separate	NI	
Demarteau et al.	Italy	Combined	100	NI
La Torre et al.	Italy	Combined	NI	
Mennini et al.	Italy	Combined	90	48
Coupe et al.	Netherlands	Combined	NI	
de Kok et al.	Netherlands	Combined	NI	
Rogoza et al.	Netherlands	Separate	96	48
Boot et al.[b]	Netherlands	Uncertain	NI	
Dasbach et al.	Norway	Uncertain	NI	
Suarez et al.	Poland	Combined	100	NI
Obradovic et al.[c]	Slovenia	Separate	90	100
Gauthier et al.[d]	Spain	Separate	96	48
Kulasingam et al.	United Kingdom	Combined	95.8	100
Jit et al.[e]	United Kingdom	Separate	82	NI
Dasbach et al.	United Kingdom	Uncertain	96	48
Debickiet al.[f]	United Kingdom	Separate	96	48

NI, not indicated.
[a] Reviewed Van de Velde et al. (55) in conjunction with Brisson et al. to assist with model interpretation.
[b] Reviewed Berkhof et al. (56) in conjunction with Boot et al. to assist with model interpretation.
[c] Reviewed Myers et al. (57) in conjunction with Obradovic et al. to assist with model interpretation.
[d] Reviewed Goldie et al. (58) in conjunction with Gauthier et al. to assist with model interpretation.
[e] Reviewed Choi et al. (52) in conjunction with Jit et al. to assist with model interpretation.
[f] Reviewed Kohli et al. (54) in conjunction with Debicki et al. to assist with model interpretation.

never-screening, over-screening and repeat testing for unsatisfactory smears were mostly not well developed (or not well reported). Women who are never-screened are known to be at substantially increased risk of cervical cancer compared to screened women [25,26]; therefore adequate modelling of never and under-screened women is an important aspect of the accurate simulation of the effectiveness (and costs) of organised screening. If a model does not consider never or under-screened women, it is difficult to see how accurate simulation of the mechanisms by which cancers arise can be performed. None of the publications identified in the systematic review documented the simulated or observed proportion of cancer cases that arose in under-screened or never-screened women. One example of why this might be important is that in order to calibrate the model output to observed age-specific cervical cancer incidence, models that do not consider inadequately screened women are likely to compensate by increasing the modelled aggressiveness of the disease process itself; this would have a flow-on effect to the assessment of the effectiveness and thus cost-effectiveness of HPV vaccination.

A number of the models also did not report whether they had considered 'over-screening', or screening more frequently than the optimal interval. Over-screening can lead to unnecessary quality of life decrements as well as additional costs; but by definition does not increase the effectiveness of screening. A longer screening interval allows for the regression of the majority of low cervical precancerous lesions (CIN 1). If the interval is shortened, or women attend screening more frequently than recommended, the regressing cervical precancerous lesion has a chance of being detected, which is likely to have associated costs and quality adjusted life years lost. Inadequate modelling of over-screening will result in inaccurate estimation of the baseline costs of cervical cancer prevention prior to the introduction of HPV vaccination and will thus have a potential flow-on effect to the potential cost-effectiveness of HPV vaccination.

4.2. Modelled abnormal smear management

Abnormal smear management was defined according to two broad areas: the structural approach to modelling the management of abnormal cytology by grade of disease (low or high grade); and secondly, screening test sensitivity and specificity. Of the 30 reviewed models, approximately half separately modelled cytology detection of low and high grade precancerous disease (as opposed to modelling cytology results only as 'normal' or 'abnormal'). In these models, low grade cytology smears were generally modelled as being managed with follow-up triage at 6 or 12 months, depending on the women's age and the screening programme, while high grade smears were referred to colposcopy. The structural approach to dividing the management of cytology into high and low grade smears was broadly based on country specific guidelines. However, 10 of the 30 models combined the management of low and high grade abnormal smears into one health state, and assumed that all abnormal smears are referred to colposcopy immediately. Unless specified as a policy recommendation (which is not the case in the very large majority of settings), such an approach to modelling abnormal smears could potentially substantially overestimate the

Please cite this article in press as: 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 (2014), http://dx.doi.org/10.1016/j.vaccine.2014.08.048

Table 7
Modelling approach to management of CIN.

Author	Country	Variation in management by CIN grade	Treatment efficacy less than 100%	Modelled follow-up for treated CIN
Kulasingam et al.	Australia	Uncertain	Uncertain	Uncertain
Debicki et al.	Canada	Yes	NI	NI
Anonychuk et al.	Canada	No	NI	NI
Brisson et al.[a]	Canada	Yes	Yes	NI
Olsen et al.	Denmark	Yes	No	NI
Suarez et al.	Finland	Yes	NI	Yes
Torvinen et al.	Finland	Yes	NI	Yes
Bergeron et al.	France	Yes	No	NI
Demarteau et al.	France	Yes	Yes	Yes
Dasbach et al.	Hungary	Yes	NI	NI
Oddsson	Iceland	Uncertain	NI	Yes
Suarez et al.	Ireland	Yes	Yes	Yes
Usher et al.	Ireland	Yes	Yes	Yes
Dee et al.	Ireland	Uncertain	Uncertain	No
Debicki et al.	Italy	Yes	No	Yes
Demarteau et al.	Italy	Yes	Yes	Yes
La Torre et al.	Italy	Yes	No	No
Mennini et al.	Italy	Uncertain	Uncertain	Uncertain
Coupe et al.	Netherlands	Uncertain	Uncertain	Uncertain
de Kok et al.	Netherlands	Uncertain	Uncertain	Uncertain
Rogoza et al.	Netherlands	Yes	Uncertain	Uncertain
Boot et al.[b]	Netherlands	Uncertain	Uncertain	Uncertain
Dasbach et al.	Norway	Yes	NI	NI
Suarez et al.	Poland	No	NI	NI
Obradovic et al.[c]	Slovenia	Yes	No	Yes
Gauthier et al.[d]	Spain	Yes	NI	NI
Kulasingam et al.	United Kingdom	Yes	No	NI
Jit et al.[e]	United Kingdom	Yes	No	NI
Dasbach et al.	United Kingdom	Yes	NI	NI
Debickiet al.[f]	United Kingdom	Yes	Yes	NI

NI, not indicated.
[a] Reviewed Van de Velde et al. (55) in conjunction with Brisson et al. to assist with model interpretation.
[b] Reviewed Berkhof et al. (56) in conjunction with Boot et al. to assist with model interpretation.
[c] Reviewed Myers et al. (57) in conjunction with Obradovic et al. to assist with model interpretation.
[d] Reviewed Goldie et al. (58) in conjunction with Gauthier et al. to assist with model interpretation.
[e] Reviewed Choi et al. (52) in conjunction with Jit et al. to assist with model interpretation.
[f] Reviewed Kohli et al. (54) in conjunction with Debicki et al. to assist with model interpretation.

incremental cost-effectiveness of HPV vaccination. Under this scenario the HPV vaccine will avert substantial colposcopy costs and morbidity (quality of life decrement) associated with overly aggressive modelled management of low grade cytology, and flow-on biopsy and treatment costs.

4.3. Modelled screening sensitivity and specificity

Cervical screening cytology test performance characteristics (sensitivity and specificity) were reported in one of two main ways. The first, which was most often used, was to report cytology only as normal or abnormal, then if cytology was abnormal, specified probabilities were defined for the detection of underlying CIN (grades 1, 2/3). However, it was not clear, in such studies, how the relative proportions of women with low and high grade cytological predictions were assigned to follow-on branches for the model. In practice women with high grade smears are usually referred directly to colposcopy, whereas women with low grade smears are often referred for follow-up cytology testing a 6 or 12 months; the relative proportions managed in these different ways thus has important implications for the costs and effects of screening.

The second method of reporting the modelling of the test characteristics of cytology was to construct a matrix describing the relationship between a specified cytology grade (for example, borderline or mild dyskaryosis) with a prediction of a specific grade of CIN threshold. In practice, each cytology grade does have a certain probability of detecting each underlying grade of CIN [73,74]. Because each country is likely to have a localised set of performance characteristics associated with cytology, to assess the accuracy of modelled test performance characteristic, models should be validated against the 'call rate' of abnormal smears for the relative proportions of each grade of cytology smear (normal, low grade, high grade) [74].

4.4. Modelled diagnosis: colposcopy and biopsy

We assessed whether models separated colposcopy and biopsy, or combined these into one intervention and assessed the parameter values assigned to the diagnostic accuracy of colposcopy. Of the reviewed models, a number combined colposcopy and biopsy into a single health state; this is a considerable simplification since not all colposcopies involve biopsy [71] and thus such models may over-estimate the diagnostic costs. If no sensitivity or specificity is indicated for colposcopy and no other qualifying information is given, then it is reasonable to assume that colposcopy was modelled as being 100% accurate. The accuracy of any direct adjustments of costs and QALYs to account for false negative colposcopies (leading to missed disease) or false positive colposcopies (leading to additional biopsies), can only be considered in the light of whether such adjustments reflect reasonable assumptions for colposcopy accuracy. Moreover, tot all models reported on how the diagnostic accuracy of colposcopy was modelled. Of those that did, a number of models indicated either perfect accuracy [29,31,34,35,37], or imperfect accuracy using test characteristics obtained from published meta-analyses [71,72]. One of the concerns with modelling perfect accuracy is that such models assume there are no lesions that are treated unnecessarily. In the models that indicated imperfect colposcopy, some modelled the sensitivity of colposcopy to be

Please cite this article in press as: 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 (2014), http://dx.doi.org/10.1016/j.vaccine.2014.08.048

Table 8
Model validation/calibration and associated targets.

Author	Country	Country specific model validation?	Number of targets	Cervical cancer	Screening related[a]	HPV related
Kulasingam et al.	Australia	Yes	3	1. Age-specific incidence	Age-specific: 1. CIN 1 2. CIN 2/3	–
Debicki et al.	Canada	Yes	6	Age-specific: 1. incidence 2. mortality	Age-specific: 3. CIN 1 4. CIN 2/3	5. Age-specific HPV prevalence (any type) in normal cytology Type-specific HPV (16/18) prevalence in cervical cancer overall
Anonychuk et al.	Canada	Yes	3	Age-specific: 1. incidence 2. mortality	3. Total CIN	–
Brisson et al.[b]	Canada	Yes	10	Age-specific: 1. incidence 2. mortality 3. Prevalence	Age-specific prevalence: 4. CIN 1 5. CIN 2/3 6. CIN 1/2/3	7. Age-specific prevalence of type-specific HPV (16,18, all HR, all LR) in normal cytology 8. Ttype-specific HPV (16,18, all HR, all LR) incidence in 15–24 years 9. Type-specific HPV distribution (16,18, all HR, all LR) in CIN 1 overall Type-specific HPV distribution (16,18, all HR, all LR) in CIN 2/3 overall
Olsen et al.	Denmark	Yes	5	1. Incidence[c]	2. Overall CIN 1[c] 3. Overall CIN 2[c] 4. Overall CIN 3[c]	5. HPV Prevalence (any HPV type) in unspecified cervical disease grade[c]
Suarez et al.	Finland	Yes	1	–	1. Total number of smears	–
Torvinen et al.	Finland	Yes	3	Age-specific: 1. incidence 2. mortality	3. Total number of smears	–

Author	Country	Country specific model validation?	Number of targets	Cervical cancer	Screening related[d]	HPV related
Bergeron et al.	France	Yes	2	Age-specific: 1. incidence 2. mortality	–	–
Demarteau et al.	France	Yes	1	1. Age-specific incidence	–	–
Dasbach et al.	Hungary	Yes	1	–	–	1. Type-specific HPV (16/18) prevalence in cervical cancer overall
Oddsson	Iceland	Not reported	0	–		
Suarez et al.[d]	Ireland	Not reported	0	No report on validation or calibration. However, quotes Debicki et al. as proof of model validation		
Usher et al.	Ireland	Yes	4	1. Incidence (undefined)	2. Overall CIN 1 3. Overall CIN 2 Overall CIN 3	–
Dee et al.	Ireland	Yes	2	Overall: 1. incidence 2. mortality	–	–
Debicki et al.	Italy	Yes	6	Age-specific: 1. incidence 2. mortality	Age-specific: 3. CIN 1 4. CIN 2/3	5. Age-specific HPV prevalence (any type) in normal cytology 6. Type-specific HPV (16/18) prevalence in cervical cancer overall
Dermarteau et al.	Italy	Yes	1	1. Age-specific incidence	–	–
La Torre et al.	Italy	Not reported	0	–		

Please cite this article in press as: 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 (2014), http://dx.doi.org/10.1016/j.vaccine.2014.08.048

Table 8 (*Continued*).

Author	Country	Country specific model validation?	Number of targets	Cervical cancer	Screening related[d]	HPV related
Mennini et al.	Italy	Yes	4	Age-specific: 1. incidence 2. mortality Age-standardised incidence	–	4. HPV prevalence (any type) in normal cytology overall
Coupe et al.	Netherlands	Yes	4	1. Age-specific incidence	2. Age-specific CIN 2/3	3. Age-specific prevalence of type-specific HPV (16, 18, 31,33, 45) in normal cytology Type-specific HPV (16,18,31,33,45) prevalence in overall cervical cancer
de Kok et al.	Netherlands	Yes	3	1. Age-specific incidence	2. Age-specific CIN	3. Age-specific HPV prevalence (any type) in normal cytology
Rogoza et al.	Netherlands	Yes	4	1. Age-specific incidence Age-specific mortality		3. Age-specific HPV prevalence (any type) in normal cytology Age-specific HPV prevalence (any oncogenic type) in normal cytology
Boot et al.[e]	Netherlands	Yes	4	1. Age-specific incidence	2. % lesion of CIN 1, 2, 3 (cytology det.) baseline and 3. % lesions of CIN 1, 2, 3 (cytology det.) FU	4. Age-specific HPV prevalence (HR types combined) in normal cytology
Author	Country	Country specific model validation?	Number of targets	Cervical cancer	Screening related[f]	HPV related
Dasbach et al.[g]	Norway	Yes	3	1. Overall incidence rate Overall mortality rate	3. Overall CIN 2/3 rate	
Suarez et al.	Poland	Not reported	0	No report on validation or calibration. However, quotes Debicki et al. as proof of model validation		
Obradovic et al.	Slovenia	Yes	3	1. Age-specific incidence 2. Age-standardise mortality Proportion cancer by FIGO stage	–	–
Gauthier et al.	Spain	Yes	7	Age-specific: 1. incidence 2. mortality	3. Age-specific CIN 2/3	4. Age-specific HPV prevalence (any type) in normal cytology 5. Type-specific HPV (unspecified) in normal cytology overall 6. Type-specific HPV (unspecified) in CIN 2/3 overall Type-specific HPV (unspecified) in cervical cancer overall
Kulasingam et al.	United Kingdom	Yes	3	1. Age-specific incidence Proportion cervical cancer by FIGO stage	–	3. Age-specific HPV prevalence (any type) in normal cytology
Jit t al.	United Kingdom	Yes	3	1. Age-specific number of cancer	–	2. Age-specific prevalence of type-specific HPV (16, 18, OHR) in unspecified cytology grade Age-specific prevalence of type specific HPV (16, 18, OHR) in absolute number of cancers

Please cite this article in press as: 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 (2014), http://dx.doi.org/10.1016/j.vaccine.2014.08.048

Table 8 (*Continued*).

Author	Country	Country specific model validation?	Number of targets	Cervical cancer	Screening related[f]	HPV related
Dasbach et al.	United Kingdom	Yes	1	–	–	1. T type-specific HPV (16/18) in overall cervical cancer incidence rate
Debicki et al.	United Kingdom	Yes	9	Age-specific: 1. incidence 2. mortality	3. Age-specific CIN 1 (cytology defined) 4. Age-specific CIN 2/3 (histologically confirmed) 5. Overall CIN	6. Age-specific HPV prevalence (any type) in normal cytology 7. Type-specific HPV (16,18, 31, 45, 52, OHR, LR) prevalence in LSIL overall 8. Type-specific HPV (16,18, 31, 45, 52, OHR, LR) prevalence in HSIL overall Type-specific HPV (16,18, 31, 45, 52, OHR, LR) prevalence in cervical cancer overall

CIN cervical intraepithelial lesions; HPV human papillomavirus; HR high risk HPV (oncogenic); LR low risk HPV (non-oncogenic).
CIN cervical intraepithelial lesions; HSIL high grade squamous intraepithelial lesion; LSIL low grade squamous intraepithelial lesion; HPV human papillomavirus; HR high risk HPV (oncogenic); LR low risk HPV (non-oncogenic).

[a] Unable to determine whether CIN terminology referred to cytology or histologically confirmed disease.
[b] Information obtained from Van de Velde et al. (55).
[c] Cannot determine whether HPV specific prevalence is overall or age-specific. Model was calibrated to reach steady state.
[d] Unable to determine whether CIN cytology or CIN histologically confirmed was used in most models unless specified.
[e] Information obtained from Berkhof et al. (56).
[f] Unable to Unable to determine whether CIN terminology referred to cytology or histologically confirmed disease.
[g] Model validated in the United States and the United Kingdom (60,62).

96% to detect any underlying cervical precancerous lesion, with a concomitant specificity of 48% [32,47,51,55,63]. Therefore, in such models approximately half of lesions are identified as being positive, even though there is no disease (false positives); this reflects actual data [71]. The implication is that a model that assumes perfect colposcopic accuracy will underestimate treatment rates, and thus will underestimate the quality of life decrements associated with treatment.

4.5. Modelling management of CIN

We assessed whether there were different approaches to the management of histologically confirmed CIN, and what type of follow-up was modelled. In general, histologically confirmed CIN 1 is managed with follow-up. However, some models did indicate treatment of CIN 1, with the proportion treated varying significantly within countries. For CIN 2/3 treatment is recommended. Following treatment, guidelines recommended intensive follow-up in the short term, as treatment is imperfect, and women with CIN 2/3 are at a higher risk of recurrence [75]. However, 7 of the 30 models assumed perfect treatment of histologically confirmed high grade disease, and thus may have overestimated treatment efficacy and underestimated total costs associated with re-treatment.

4.6. Validation of modelled 'background' cervical screening

In the studies considered in this review, model consistency with observed data (validation) was assessed using an array of targets. In the main, model validation was performed against observed age-specific cervical cancer incidence and mortality. However, only half the models also assessed consistency with observed cervical precancerous lesions. This was mainly in the form of CIN related outcomes (age-specific and overall rates). With reported measures using CIN it was difficult to determine whether the terminology referred to cytology predicting CIN or histologically confirmed CIN. It needs to be highlighted that the estimation of model parameters may be based on histological or cytological data but this should not be taken to represent the true underlying health state. In addition to these targets, approximately half validated to HPV prevalence related targets. The three main HPV related targets that were used were: (1) age-specific prevalence for any HPV type in normal cytology; (2) overall prevalence of type-specific HPV infection in cytology or histologically confirmed CIN (usually HPV 16 and/or 18)); and finally (3) overall prevalence of type-specific HPV infection (usually 16 and/or 18) in cervical cancer. It should be further noted that an issue that need to be considered when comparing model predicted versus observed age-specific cervical cancer incidence and mortality is the likely impact of cohort effects in cross-sectional data [5].

4.7. Organised screening model structure and implications for cost-effectiveness evaluation

The modelled effectiveness of organised screening in a particular setting can influence the calculated incremental costs and benefits from HPV vaccination. However, there are potentially two opposing influences arising from the modelling of screening that may impact the cost-effectiveness of HPV vaccination. Firstly, the modelled effectiveness of screening influences the incremental gains that can be realised after the introduction of vaccination. For example, for countries that are currently avoiding 50% of cervical cancers with screening, an HPV vaccination programme with a coverage level of 85% will increase the proportion of avoidable cervical cancers to 82% [76]. In countries with well organised screening, a high participation rate and a relatively compliant screening population, the addition of HPV vaccination should theoretically produce less favourable results in terms of life years saved compared with results in countries with poor quality cervical screening programmes (characterised by low coverage and/or opportunistic screening), if it is assumed that screening behaviour remains unchanged after vaccination.

Please cite this article in press as: 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 (2014), http://dx.doi.org/10.1016/j.vaccine.2014.08.048

However, where there is high screening coverage and a compliant screening population, a relatively higher number of cervical precancerous lesions are likely to be detected by cytology, and be subsequently histologically confirmed, compared to populations with low levels of screening. The detected cervical precancerous lesions that are intermediate health outcomes of cervical screening are potentially associated with quality of life decrements. Therefore, in countries with high screening coverage, the application of high disutility health state preference scores to screening health states in a cervical screening model has the potential to make HPV vaccination appear more cost-effective with respect to the ICER per QALY referenced to screening-only. This outcome occurs because HPV vaccination results in the averting detrimental quality of life effects associated with managing screen-detected abnormalities. The contribution of this effect may not be immediately apparent in the final outcome of the cost-effectiveness evaluation, and heavily depends on the utility weight sets chosen for the evaluation.

4.8. Limitations

There are several limitations with our review. For each of the reviewed models, restrictions on publication space may have prevented specific dimensions associated with screening to be reported. However, in an attempt to overcome these issues, an effort was made to examine other cited articles to determine how specific dimensions of screening were modelled. In many cases, other models were referenced and were adopted for the country of interest. However, in many cases, even after reviewing all related publications, it was still not entirely clear what parameters were used, and whether local data were assessed to examine model validity. Another particular limitation of this analysis is that models may not have had sufficient data to reflect each of the pre-defined parameters associated with organised screening. However, the aim in this systematic review was to assess screening in countries with organised programmes. The restriction was applied with the expectation that models could often be developed with the use of reports from local monitoring bodies to enable a more accurate approximation of real world behaviour.

In general terms, the optimal model design is dependent on the context of the screening programme. Where an organised screening programme has recently been introduced, or cancer incidence is expected to decrease further over time, then a model that can characterise parameter uncertainty should be developed to evaluate the benefit of HPV vaccination in this context. Otherwise, in situations where cancer incidence and precancer rates are expected to remain stable (such is the case in a few countries with long-established screening programmes and stable rates of the last several years), and there is sufficient data on the rates of events expected (for example event rates for treated CIN, number of colposcopies, etc.) then a proportionate model may be sufficient to evaluate the future impact of HPV vaccination on the screening programme. However, these types of models should be used cautiously since they do not account for actual screening behaviour or changes in outcomes over time (which generally continue to occur many years after the introduction of organised screening programmes). The modelled impact of screening on the screening-averted incidence and mortality of cervical cancer will influence the calculated incremental benefit of HPV vaccination.

In terms of the potential impact on evaluations of male vaccination, although cervical screening is expected to have only a small effect on the circulation of HPV in the population, in general terms any new intervention needs to be assessed against an accurate comparator which accounts for all the costs and QALYs of existing prevention programmes (including those related to cervical screening and its intermediate outcomes). Although simplification of the modelling of the costs and QALYs of cervical screening in assessments of either female or male HPV vaccination may be justified, our proposed framework allows such simplifications to be systematically documented and assessed.

5. Conclusion

The models of organised screening used for the evaluation of the cost-effectiveness of HPV vaccination varied in quality. One area of screening that was generally modelled relatively well was population participation. However, with respect to the other areas of screening (abnormal smear management, diagnostic follow-up and manage of CIN), models were inconsistent in structure and, in some cases, very simplified. Many models did not adequately reflect actual referral pathways for screening programmes. There was also considerable variability in the validation targets, especially with reference to CIN. Moreover, a number of models relied on using a small number of validation targets. These findings are most likely to be due to the lack of clear guidelines on what aspects of an organised screening model should be included and reported. In general terms, models of HPV vaccination in developed countries can be improved by further attention to the 'background' modelling of secondary protection via cervical screening. The quality framework developed for this review can be used to inform future HPV vaccination evaluations, including evaluations of the cost-effectiveness of male vaccination and next generation HPV vaccines. This quality framework will also be applicable to assessing models used to evaluate new cervical screening technologies and recommendations in the era of HPV vaccination.

Competing interests

KC declares that she is co-PI of a new trial of primary HPV screening in Australia ('Compass') which has received a funding contribution from Roche Molecular Systems, CA, USA.

Authors' contributions

LS developed the detailed quality framework, performed the systematic review and extracted data. Both authors were involved in the design of the study and drafting the manuscript.

Acknowledgements

The authors wish to thank Luke Testa for providing editorial assistance. This study was funded via a Cancer Council NSW PhD scholarship for LS and National Health and Medical Research Council Grants #440200 and #1007518. KC receives salary support from the National Health and Medical Research Council Australia (CDF 1007994).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.vaccine.2014.08.048.

References

[1] Dorleans F, et al. The current state of introduction of human papillomavirus vaccination into national immunisation schedules in Europe: first results of the VENICE2 2010 survey. Euro Surveill 2010;15(47).
[2] Gertig DM, Brotherton JM, Saville M. Measuring human papillomavirus (HPV) vaccination coverage and the role of the National HPV Vaccination Program Register: Australia. Sex Health 2011;8(2):171–8.
[3] Marra F, et al. Effectiveness and cost effectiveness of human papillomavirus vaccine: a systematic review. Pharmacoeconomics 2009;27(2):127–47.
[4] Jeurissen S, Makar A. Epidemiological and economic impact of human papillomavirus vaccines. Int J Gynecol Cancer 2009;19(4):761–71.

[5] Insinga RP, Dasbach EJ, Elbasha EH. Structural differences among cost-effectiveness models of human papillomavirus vaccines. Expert Rev Vaccines 2008;7(7):895–913.
[6] Brisson M, et al. Economic evaluation of human papillomavirus vaccination in developed countries. Public Health Genomics 2009;12(5–6):343–51.
[7] Newall AT, et al. Cost-effectiveness analyses of human papillomavirus vaccination. Lancet Infect Dis 2007;7(4):289–96.
[8] Ferko N, et al. Evolution of the health economics of cervical cancer vaccination. Vaccine 2008;26(Suppl. 5):F3–15.
[9] Canfell K. Cost-effectiveness and impact studies of prevention of HPV-associated disease in higher income countries. In: Public health workshop. 28th International human papillomavirus conference & clinical and public health workshops. 2012.
[10] Kim JJ, et al. Modeling cervical cancer prevention in developed countries. Vaccine 2008;26(Suppl. 10):K76–86.
[11] Beutels P, Jit M. A brief history of economic evaluation for human papillomavirus vaccination policy. Sex Health 2010;7(3):352–8.
[12] Puig-Junoy J, Lopez-Valcarcel BG. Economic evaluations of massive HPV vaccination: within-study and between study variations in incremental cost per QALY gained. Prev Med 2009;48(5):444–8.
[13] Canfell K, et al. Impact of HPV vaccination on cost-effectiveness of existing screening programs: example from New Zealand. In: 28th International human papillomavirus conference & clinical and public health workshops. 2012.
[14] Canfell K, et al. Modeling preventative strategies against human papillomavirus-related disease in developed countries. Vaccine 2012;30(Suppl. 5):F157–67.
[15] Brennan A, Akehurst R. Modelling in health economic evaluation. What is its place? What is its value? Pharmacoeconomics 2000;17(5):445–59.
[16] Philips Z, et al. Good practice guidelines for decision-analytic modelling in health technology assessment: a review and consolidation of quality assessment. Pharmacoeconomics 2006;24(4):355–71.
[17] 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.
[18] Craig BM, et al. Proceedings of the modeling evidence in HPV pre-conference workshop in Malmö: Sweden, May 9–10, 2009. Clin Ther 2010;32(8):1546–64.
[19] Decision analytic modelling in the economic evaluation of health, technologies. A consensus statement. Consensus Conference on Guidelines on Economic Modelling in Health Technology Assessment. Pharmacoeconomics 2000;17(5):443–4.
[20] Goldhaber-Fiebert JD, Stout NK, Goldie SJ. Empirically evaluating decision-analytic models. Value Health 2010;13(5):667–74.
[21] IARC Working Group on the Evaluation of Cancer. Effectiveness of screening in populations. In: IARC Working Group on the evaluation of cancer preventative strategies. IARC handbooks of cancer prevention. Cervix cancer screening, vol. 10. Lyon: IARC; 2005. p. 201–25.
[22] National Health and Medical Research Council (NHMRC). Screening to prevent cervical cancer: guidelines for the management of asymptomatic women with screen detected abnormalities. Canberra: Commonwealth Government; 2005.
[23] NHS Cancer Screening Programmes. Colposcopy and programme management: guidelines for the NHS cervical screening programme. Sheffield: NHS Cancer Screening Programmes; 2004.
[24] National Screening Unit. Guidelines for cervical screening in New Zealand: incorporating the management of women with abnormal cervical smears. Wellington: Ministry of Health; 2008.
[25] Sasieni P, Adams J, Cuzick J. Benefit of cervical screening at different ages: evidence from the UK audit of screening histories. Br J Cancer 2003;89(1):88–93.
[26] Lewis H, et al. Monitoring the performance of New Zealand's National Cervical Screening Programme through data linkage. N Z Med J 2009;122(1305):15–25.
[27] Melnikow J, et al. Cervical intraepithelial neoplasia outcomes after treatment: long-term follow-up from the British Columbia Cohort Study. J Natl Cancer Inst 2009;101(10):721–8.
[28] IARC, I.A.f.R.O.C. European Cancer Observatory, 2007 [cited 2011 March 1]. Available from: http://eu-cancer.iarc.fr/
[29] Kulasingam S, et al. A cost-effectiveness analysis of adding a human papillomavirus vaccine to the Australian National Cervical Cancer Screening Program. Sex Health 2007;4(3):165–75.
[30] Brisson M, et al. The potential cost-effectiveness of prophylactic human papillomavirus vaccines in Canada. Vaccine 2007;25(29):5399–408.
[31] Anonychuk AM, et al. A cost-utility analysis of cervical cancer vaccination in preadolescent Canadian females. BMC Public Health 2009;9:401.
[32] Debicki D, et al. Comparison of detailed and succinct cohort modelling approaches in a multi-regional evaluation of cervical cancer vaccination. Vaccine 2008;26(Suppl. 5):F16–28.
[33] Olsen J, Jepsen MR. Human papillomavirus transmission and cost-effectiveness of introducing quadrivalent HPV vaccination in Denmark. Int J Technol Assess Health Care 2010;26(2):183–91.
[34] Suarez E, et al. Cost-effectiveness of vaccination against cervical cancer: a multi-regional analysis assessing the impact of vaccine characteristics and alternative vaccination scenarios. Vaccine 2008;26(Suppl. 5):F29–45.
[35] Torvinen S, et al. Cost effectiveness of prophylactic HPV 16/18 vaccination in Finland: results from a modelling exercise. J Med Econ 2010;13(2):284–94.
[36] Bergeron C, et al. Cost-effectiveness analysis of the introduction of a quadrivalent human papillomavirus vaccine in France. Int J Technol Assess Health Care 2008;24(1):10–9.
[37] Demarteau N, et al. A generally applicable cost-effectiveness model for the evaluation of vaccines against cervical cancer. Int J Public Health 2010.
[38] Demarteau N, Standaert B. Modelling the economic value of cross- and sustained-protection in vaccines against cervical cancer. J Med Econ 2010;13(2):324–38.
[39] Dasbach EJ, et al. The cost effectiveness of a quadrivalent human papillomavirus vaccine (6/11/16/18) in Hungary. J Med Econ 2010;13(1):110–8.
[40] Oddsson K, et al. Cost-effectiveness of human papilloma virus vaccination in Iceland. Acta Obstet Gynecol Scand 2009;88(12):1411–6.
[41] Usher C, et al. Cost-effectiveness of human papillomavirus vaccine in reducing the risk of cervical cancer in Ireland due to HPV types 16 and 18 using a transmission dynamic model. Vaccine 2008. Q6
[42] Dee A, Howell F. A cost-utility analysis of adding a bivalent or quadrivalent HPV vaccine to the Irish cervical screening programme. Eur J Public Health 2010;20(2):213–9.
[43] La Torre G, et al. The health technology assessment of bivalent HPV vaccine Cervarix in Italy. Vaccine 2010;28(19):3379–84.
[44] Mennini FS, et al. Health and economic impact associated with a quadrivalent HPV vaccine in Italy. Gynecol Oncol 2009;112(2):370–6.
[45] de Kok IM, van Ballegooijen M, Habbema JD. Cost-effectiveness analysis of human papillomavirus vaccination in the Netherlands. J Natl Cancer Inst 2009;101(15):1083–92.
[46] Coupe VM, et al. HPV16/18 vaccination to prevent cervical cancer in The Netherlands: model-based cost-effectiveness. Int J Cancer 2009;124(4):970–8.
[47] Rogoza RM, et al. Cost-effectiveness of prophylactic vaccination against human papillomavirus 16/18 for the prevention of cervical cancer: adaptation of an existing cohort model to the situation in the Netherlands. Vaccine 2009;27(35):4776–83.
[48] Boot HJ, et al. Assessing the introduction of universal human papillomavirus vaccination for preadolescent girls in The Netherlands. Vaccine 2007;25(33):6245–56.
[49] Dasbach EJ, Largeron N, Elbasha EH. Assessment of the cost-effectiveness of a quadrivalent HPV vaccine in Norway using a dynamic transmission model. Expert Rev Pharmacoecon Outcomes Res 2008;8(5):491–500.
[50] Obradovic M, Mrhar A, Kos M. Cost-effectiveness analysis of HPV vaccination alongside cervical cancer screening programme in Slovenia. Eur J Public Health 2010;20(4):415–21.
[51] Gauthier A, et al. Long-term clinical impact of introducing a human papillomavirus 16/18 AS04 adjuvant cervical cancer vaccine in Spain. Eur J Public Health 2008;18(6):674–80.
[52] Kulasingam SL, et al. Adding a quadrivalent human papillomavirus vaccine to the UK cervical cancer screening programme: a cost-effectiveness analysis. Cost Eff Resour Alloc 2008;6:4.
[53] Choi YH, et al. Transmission dynamic modelling of the impact of human papillomavirus vaccination in the United Kingdom. Vaccine 2010;28(24):4091–102.
[54] Jit M, Choi YH, Edmunds WJ. Economic evaluation of human papillomavirus vaccination in the United Kingdom. BMJ 2008;337:a769.
[55] Kohli M, et al. Estimating the long-term impact of a prophylactic human papillomavirus 16/18 vaccine on the burden of cervical cancer in the UK. Br J Cancer 2007;96(1):143–50.
[56] Van de Velde N, Brisson M, Boily MC. Modeling human papillomavirus vaccine effectiveness: quantifying the impact of parameter uncertainty. Am J Epidemiol 2007;165(7):762–75.
[57] Berkhof J, et al. Natural history and screening model for high-risk human papillomavirus infection: neoplasia and cervical cancer in the Netherlands. Int J Cancer 2005;115(2):268–75.
[58] Myers ER, et al. Mathematical model for the natural history of human papillomavirus infection and cervical carcinogenesis. Am J Epidemiol 2000;151(12):1158–71.
[59] Goldie SJ, et al. Projected clinical benefits and cost-effectiveness of a human papillomavirus 16/18 vaccine. J Natl Cancer Inst 2004;96(8):604–15.
[60] Dasbach EJ, Elbasha EH, Insinga RP. Mathematical models for predicting the epidemiologic and economic impact of vaccination against human papillomavirus infection and disease. Epidemiol Rev 2006;28:88–100.
[61] Elbasha EH, Dasbach EJ, Insinga RP. Model for assessing human papillomavirus vaccination strategies. Emerg Infect Dis 2007;13(1):28–41.
[62] Elbasha EH, Dasbach EJ, Insinga RP. A multi-type HPV transmission model. Bull Math Biol 2008;70(8):2126–76.
[63] Dasbach EJ, Insinga RP, Elbasha EH. The epidemiological and economic impact of a quadrivalent human papillomavirus vaccine (6/11/16/18) in the UK. BJOG 2008;115(8):947–56.
[64] Garnett GP, Anderson RM. Balancing sexual partnerships in an age and activity stratified model of HIV transmission in heterosexual populations. IMA J Math Appl Med Biol 1994;11(3):161–92.
[65] Rogoza RM, et al. Optimization of primary and secondary cervical cancer prevention strategies in an era of cervical cancer vaccination: a multi-regional health economic analysis. Vaccine 2008;26(Suppl. 5):F46–58.
[66] Nanda K, et al. Accuracy of the Papanicolaou test in screening for and follow-up of cervical cytologic abnormalities: a systematic review. Ann Intern Med 2000;132(10):810–9.
[67] Fahey MT, Irwig L, Macaskill P. Meta-analysis of Pap test accuracy. Am J Epidemiol 1995;141(7):680–9.
[68] Cuzick J, et al. Human papillomavirus testing in primary cervical screening. Lancet 1995;345(8964):1533–6.

Please cite this article in press as: 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 (2014), http://dx.doi.org/10.1016/j.vaccine.2014.08.048

[69] Bulkmans NW, et al. Human papillomavirus DNA testing for the detection of cervical intraepithelial neoplasia grade 3 and cancer: 5-year follow-up of a randomised controlled implementation trial. Lancet 2007;370(9601):1764–72.
[70] Berkhof J, et al. Evaluation of cervical screening strategies with adjunct high-risk human papillomavirus testing for women with borderline or mild dyskaryosis. Int J Cancer 2006;118(7):1759–68.
[71] Mitchell MF, et al. Colposcopy for the diagnosis of squamous intraepithelial lesions: a meta-analysis. Obstet Gynecol 1998;91(4):626–31.
[72] Hopman EH, Kenemans P, Helmerhorst TJ. Positive predictive rate of colposcopic examination of the cervix uteri: an overview of literature. Obstet Gynecol Surv 1998;53(2):97–106.
[73] International Agency for Research on Cancer (IARC). Screening tests. In: IARC handbook of cancer prevention: cervical screening. Lyon: IARC Press; 2005. p. 59–115.
[74] MSAC. Automation-assisted and liquid based cytology for cervical cancer screening. MSAC application 1122, MSAC. Canberra: Department of Health and Ageing; 2009.
[75] Cecchini S, et al. Recurrence after treatment by loop electrosurgical excision procedure (LEEP) of high-grade cervical intraepithelial neoplasia. Tumori 2002;88(6):478–80.
[76] Franceschi S, et al. EUROGIN 2008 roadmap on cervical cancer prevention. Int J Cancer 2009;125(10):2246–55.

Please cite this article in press as: 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 (2014), http://dx.doi.org/10.1016/j.vaccine.2014.08.048