Published in final edited form as:

Med Decis Making. 2013 April ; 33(3): 407–414. doi:10.1177/0272989X12453503.

Multicohort Models in Cost-Effectiveness Analysis: Why Aggregating Estimates over Multiple Cohorts Can Hide Useful Information

James F. O'Mahony, MA, Joost van Rosmalen, PhD, Ann G. Zauber, PhD, and Marjolein van Ballegooijen, MD, PhD

Abstract

Background—Models used in cost-effectiveness analysis (CEA) of screening programs may include 1 or many birth cohorts of patients. As many screening programs involve multiple screens over many years for each birth cohort, the actual implementation of screening often involves multiple concurrent recipient cohorts. Consequently, some advocate modeling all recipient cohorts rather than 1 birth cohort, arguing it more accurately represents actual implementation. However, reporting the cost-effectiveness estimates for multiple cohorts on aggregate rather than per cohort will fail to account for any heterogeneity in cost-effectiveness between cohorts. Such heterogeneity may be policy relevant where there is considerable variation in cost-effectiveness between cohorts, as in the case of cancer screening programs with multiple concurrent recipient birth cohorts, each at different stages of screening at any one point in time.

Objective—The purpose of this study is to illustrate the potential disadvantages of aggregating cost-effectiveness estimates over multiple cohorts, without first considering the disaggregate estimates.

Analysis—We estimate the cost-effectiveness of 2 alternative cervical screening tests in a multicohort model and compare the aggregated and per-cohort estimates. We find instances in which the policy choices suggested by the aggregate and per-cohort results differ. We use this example to illustrate a series of potential disadvantages of aggregating CEA estimates over cohorts.

Conclusions—Recent recommendations that CEAs should consider the cost-effectiveness of more than just a single cohort appear justified, but the aggregation of estimates across multiple cohorts into a single estimate does not.

Keywords

cohort model; multicohort model; population model

The cost-effectiveness of health care interventions can differ between patient cohorts. Differences in cost-effectiveness between cohorts mean a common policy for all may not be optimal. Cost-effectiveness analyses (CEAs) should inform policy makers of heterogeneity in cost-effectiveness, enabling separate decisions for different cohorts if necessary.¹ Indeed, the need to reflect heterogeneity between groups is well recognized in CEA, as many analyses report separate cost-effectiveness estimates for separate groups.

Despite awareness of the relevance of cohort heterogeneity, the understanding of its implications for CEA model structure are evidently still incomplete in the particular case of CEAs of screening interventions. Screening CEAs conventionally model only a single birth cohort of screen recipients.² However, because screening programs typically last many years, introducing or changing a screening program affects multiple birth cohorts

simultaneously, and cost-effectiveness may vary between these cohorts. Consequently, some screening CEAs feature all recipient cohorts and report cost-effectiveness on aggregate over all cohorts^{3–7}; indeed, this has been explicitly advocated as more representative of actual implementation.² However, modeling all screened cohorts together and reporting cost-effectiveness on aggregate will fail to account for any cohort heterogeneity and does not facilitate separate decisions for specific cohorts. The purpose of this study is to illustrate the potential disadvantages of aggregating cost-effectiveness estimates over multiple cohorts, without first considering the disaggregate estimates. This study is intended to inform both guidance on appropriate model structures and the correct interpretation of results from multiple cohort models.

This study uses the specific context of screening to address the question of aggregation across cohorts for 2 reasons. First, the existing screening CEAs that model multiple cohorts and report cost-effectiveness on aggregate show that this approach is used in current practice. Second, the methodological research advocating this approach was motivated by the fact that screening typically involves multiple concurrent recipient birth cohorts.² Although our study most immediately applies to screening, the questions around aggregation apply more broadly to any example where cost-effectiveness varies between patient groups.

This article is structured as follows. The following section provides the background to this study, giving an overview of different model structures and terminology and briefly reviewing how cohort heterogeneity has been addressed in the literature. The Analysis section describes the methods and results of a CEA of cervical cancer screening and explains 4 potential problems of aggregating cost-effectiveness estimates, using the cervical screening CEA as an example. The discussion considers the implications of these issues for decision making and the potential tradeoffs between single- and multiple-cohort modeling.

BACKGROUND

Model Types and Terminology

Cohorts can be defined either as groups starting an intervention at a common point in time or of a common birth year. The birth year definition is appropriate in the case of screening, as screening eligibility is typically determined by age and because the cost-effectiveness of screening can vary with age. We categorize cohorts by adapting the terminology used by Hoyle and Anderson of prevalent and incident cohorts.⁸ The terms do not correspond to any disease state in this context but to screening eligibility. The single birth cohort starting screening at the program start age in the current period is the current incident cohort, cohorts already within the screening age range when a program is introduced or changed are the prevalent cohorts, and cohorts currently younger than the program start age are the future incident cohorts. Note that there can be many prevalent and future incident cohorts, each of different birth years. We collectively describe the prevalent cohorts and the current incident cohort as the current cohorts, whereas the future incident cohorts are simply described as the future cohorts.

Cost-effectiveness may vary between prevalent cohorts at any point in time because of differences in age and screening histories. Consequently, even if unit costs and treatment effectiveness are constant over time, screening interventions can lead to cohort heterogeneity in cost-effectiveness. The birth-year cohorts in this analysis are mutually exclusive; however, the cohort groups used in published CEAs are not always mutually exclusive.

Cohort Heterogeneity in the Literature

Single-cohort models are widely used in CEA.^{8–10} They model the costs and effects of an intervention for 1 cohort of patients. In the case of CEAs of screening, this cohort is typically a single birth-year cohort.^{11–14} Some CEAs models multiple cohorts. The most widely recognized rationale for multicohort modeling is in the context of infectious diseases, in which multiple cohorts are necessary to simulate herd immunity.^{15,16} However, the rationale considered in this study is the need to capture cohort heterogeneity. A single-cohort model is sufficient if that cohort's cost-effectiveness is representative of all recipient cohorts. Conversely, if cost-effectiveness is anticipated to vary significantly between cohorts, then a multicohort model will be more appropriate.^{8,17–18}

The literature addressing the cohort heterogeneity rationale for multicohort modeling is sparse. However, 2 previously published articles in this journal have directly addressed the topic and advocate multicohort models as more representative of the actual health care implementation.^{2,8} Dewilde and Anderson presented a CEA of cervical screening, showing its average cost-effectiveness ratio (ACER) to be considerably higher when prevalent cohorts are considered in addition to the incident cohort alone.² More recently, Hoyle and Anderson recognized that if cost-effectiveness varies between prevalent and incident cohorts, then the number of future incident cohorts modeled will influence the aggregate cost-effectiveness estimate.⁸ They also considered the possibility that cost-effectiveness may vary between incident cohorts. Consequently, they recommend CEAs include all current and all future cohorts and report a “combined cohorts ICER.” However, an accompanying editorial by Kuntz and others¹⁹ questioned whether it is appropriate to aggregate results over multiple cohorts. Similarly, Karnon and others¹⁸ noted that Dewilde and Anderson's multicohort approach can be used to account for the effect of cohort heterogeneity on aggregate cost-effectiveness but suggested separate per-cohort analyses be used where interventions can be applied differently to separate cohorts. Karnon and others' own analysis considered the influence of allowing incident cohorts to enter a CEA over time but primarily addressed the question as to how time horizons for the assessment of effects should be adapted to accommodate future incident cohorts, rather than questions of aggregation over cohorts.

Examples of multiple cohort models from the screening literature include models featuring the current cohorts only,^{3,4,7} whereas others also model future incident cohorts.^{5,6} Notably, cost-effectiveness is reported on aggregate over all cohorts in all these examples.

ANALYSIS

Methods

We present a CEA of 2 alternative cervical cancer screening tests as an illustrative example of the consequences of aggregating cost-effectiveness estimates over cohorts. We use the MISCAN-cervix microsimulation model developed by the Department of Public Health at Erasmus MC, the Netherlands. The model simulates the individual life histories of women from birth until death. The model generates age- and time-specific cancer incidence and mortality estimates. Alternative screening scenarios can be simulated in the model, whereby screening may detect disease in the preclinical phase, permitting early treatment to prevent further disease development and subsequent death from the cancer. Quality-of-life weights are applied to the different health states to estimate health effects of screening in terms of quality-adjusted life-years (QALYs). The model and its data sources can be found in previous publications.^{20–23}

The model also includes vaccination against human papillomavirus (HPV) types 16 and 18, precursors to cervical cancer. The characteristics of the vaccine are as modeled in a recent

CEA of the bivalent vaccine in the Netherlands.²¹ The vaccine is administered in 3 doses and is assumed equally effective in all vaccinated cohorts. The vaccine coverage rate simulated is 85%, and there is no selection effect assumed regarding vaccine coverage and the risk of infection. Vaccine efficacy is assumed to be 70% against cancer, 35% against preinvasive lesions, and 1.5% against HPV infections. The protection of the vaccine is assumed to be lifelong in this analysis.

We compare 2 screening scenarios. The baseline scenario is the current Dutch cervical screening program of 7 lifetime screens between the ages of 30 and 60 at 5-y intervals, using a primary cytology screen, with cytology triage for abnormal primary smears. The alternative scenario is a change in 2012 to screening with a primary HPV DNA test, followed by 3 cytology triage tests for abnormal primary screen results. The screening age range and frequency remain unchanged.

The model simulates 46 separate birth-year cohorts. The size of each cohort is matched to the number of female live births in each respective birth year in the Netherlands. The current incident cohort starting screening in 2012 is aged 30. There are 30 prevalent birth cohorts aged 31 to 60 y in 2012. Those in the prevalent cohorts will have already experienced at least 1 screening round prior to 2012 using the current test. The 15 future incident cohorts that have not yet started screening are aged between 29 and 15 y in 2012. Vaccination against HPV types 16 and 18 was introduced in the Netherlands in 2009 for girls aged 12 y with catch-up for those aged up to and including 16 y. Consequently, the youngest 4 future incident cohorts in the model have been offered vaccination.

All costs and health effects due to the screening program are included in the model from 2012 until death for each cohort. Costs and effects are discounted at 4% to the discount year of 2012. The costs used are from 2008 and are reported in euro (1 euro = 1.45 US dollars; 6 July 2011). We assume the unit costs of screening, follow-up, and avoided treatment remain constant over time. The parameters determining health gain from the early detection of cancer and precancerous lesions are also assumed constant over time. Age-specific disease incidence is also assumed to remain the same, except for those vaccinated cohorts, in which disease incidence falls in accordance with the vaccine's expected effectiveness.

The costs and effects are estimated both on a disaggregated per-cohort basis and an aggregated basis over multiple cohorts. These results are presented as the ACER of each intervention compared with no screening and the incremental cost-effectiveness ratio (ICER) of switching from the current primary cytology test to the HPV test. Both ACERs and ICERs are reported in terms of euro per QALY.

No direct funding was provided for this project. The first author is supported by the Health Research Board of Ireland and by the US National Cancer Institute under grant U01 CA115953. Neither funding body had a role in the study.

Results

Figure 1 shows the ACERs of both tests for each cohort and on aggregate. Figure 2 shows the ICERs for HPV screening relative to cytology screening. Table 1 reports the effects, cost, ACERs, and ICERs for selected cohorts and the multicohort aggregates. The table also reports the proportion of total net costs of screening for the selected cohorts, using the proportion of total costs under primary cytology as an indication of the relative weight of each cohort in the aggregate.

The per-cohort ACERs and ICERs fall into groups of cohorts with common screening histories and vaccination status. For example, in 2012, the 5 oldest prevalent cohorts, –30 to

–26, all have only 1 lifetime screening left at age 60, whereas cohorts –25 to –21 have 2 left, 1 at 55 and 1 at 60. The common cost-effectiveness within each group reflects the assumption of constant costs and effects: The small remaining variation between cohorts is a consequence of the stochastic nature of the simulation model.

The per-cohort ACERs show both screening tests to be less cost-effective in vaccinated cohorts relative to unvaccinated incident cohorts and increasingly less cost-effective with age in the prevalent cohorts. HPV testing is more costly and more effective than cytology in every cohort. Cytology is weakly dominated by a combination of HPV testing and no intervention in cohorts –30 to –16, the oldest 15 prevalent cohorts, so no ICERs of HPV testing are reported for them. The per-cohort ICER of HPV relative to primary cytology screening is essentially uniform among the unvaccinated incident cohorts (0 to 11), at approximately €47 700 per QALY. The HPV test is more cost-effective in prevalent cohorts –10 to –1 than in the incident cohorts, being the most cost-effective in the youngest 5 prevalent cohorts with an ICER of €31 800 per QALY. HPV testing is markedly less cost-effective in the vaccinated future cohorts 12 to 15, with ICERs in excess of €500 000 per QALY. The aggregate ICER of HPV testing over the current cohorts is approximately €38 300 per QALY, whereas the aggregate ICER of all cohorts is somewhat higher at €45 400.

The results above show how the cost-effectiveness of interventions can vary significantly between cohorts. The markedly high ACERs for both screening methods in the vaccinated cohorts are a consequence of the simulated reduction in disease incidence following vaccination. Similarly, the rising ACERs with age within the prevalent cohorts are attributed to the fact that both disease incidence and the potential life-years that can be gained fall with age. The changing relative cost-effectiveness of the 2 tests is attributed to the fact that although HPV testing is more effective because of higher sensitivity, it can lead to over-detection of transient HPV infections and noncancerous lesions, which are more common in younger women.

Four Problems of Aggregation

Both Dewilde and Anderson² and Hoyle and Anderson⁸ call attention to the fact that many CEA models do not accurately represent the policy choices they are to inform: Although CEAs typically assess cost-effectiveness for only a single incident cohort, the policy decisions they inform are almost never for that cohort alone but also for other cohorts, both in the present and the future. Consequently, Hoyle and Anderson⁸ suggested that CEAs include all present and future recipient cohorts. Given the need to capture cohort heterogeneity, this suggestion appears sensible. However, reporting cost-effectiveness as a single aggregate estimate for all cohorts seems inappropriate for the following reasons: 1) aggregate estimates over many cohorts may hide useful information from decision makers; 2) the choice of which cohorts to include appears logically problematic; 3) aggregate estimates demand significant assumptions about the future, which results in large uncertainty in aggregate estimates; and 4) aggregate modeling prompts broader questions about decision making over multiple periods for health care priority setting. This section explains each of these problems in detail.

Aggregate Results Hide Differences

Reporting an aggregate cost-effectiveness estimate across all recipient cohorts carries an implicit assumption that a common policy decision will be taken for all cohorts. However, different reimbursement decisions can clearly be made for different cohorts in many cases. Where selective reimbursement is possible, decision makers would be better served by cost-effectiveness estimates disaggregated per cohort, allowing them to approve the intervention for those cohorts identified as cost-effective to treat and withhold it from those who are not.

Our example shows how aggregate results can lead to inappropriate policy choices in terms of cost-effectiveness. If the cost-effectiveness threshold was €40 000 per QALY, the aggregate ICER over the current cohorts of €38 300 per QALY would indicate the HPV test is cost-effective. However, according to the per-cohort estimates, HPV screening would not be cost-effective in any incident cohort or the oldest 20 of the 30 prevalent cohorts. Conversely, if the aggregate ICER over all cohorts of €45 400 per QALY was used, then the HPV test would be deemed not cost-effective, but the per-cohort estimates indicate that it would be cost-effective for prevalent cohorts -10 to -1. Similarly, relying on the single incident cohort ICER of €47 700 per QALY would also lead to a rejection of the test for all cohorts.

There are certain cases in which disaggregated per-cohort estimates are not appropriate. These include interventions in which the selective allocation of the intervention is not possible, such as water fluoridation. Aggregate estimates are also appropriate in cases in which the effects of an intervention are shared across cohorts, such as herd immunity from vaccination. Another consideration is that it may be impractical to offer different interventions to different cohorts, as the costs of tailoring the intervention to each cohort may outweigh the benefits. For example, in the case of screening, offering different screening strategies to different birth cohorts may reduce adherence, possibly leading to a greater health loss than would be gained by optimizing the program for each cohort.

Logical Problems of Including and Excluding Cohorts

A related problem of aggregation is that results depend on which cohorts are modeled. There may be many potential recipient cohorts, but not all may be cost-effective to treat. We included vaccinated women as an example of cohorts in which the HPV test has poor incremental cost-effectiveness. Presumably, these cohorts should be excluded when estimating aggregate cost-effectiveness of HPV primary screening, as they are unlikely to receive that strategy because of its poor cost-effectiveness. However, using cost-effectiveness as an inclusion criterion for cohorts in the model leads to a certain circularity: If only cost-effective cohorts are included in the analysis (assuming at least 1 cost-effective cohort exists), then the intervention will necessarily be cost-effective on aggregate for the selected cohorts.

An analogous problem arises if cost-effectiveness is not an inclusion criterion: If all cohorts have ICERs that are either all above or all below the threshold, then including some or all cohorts will make no difference to the reimbursement decision. However, if some cohorts' ICERs are above the threshold and others' are below it, then an aggregate ICER, which must be either above or below the threshold, will necessarily imply an incorrect decision for some cohorts. Consequently, the aggregate multicohort approach could be interpreted as either not enhancing decision making or necessarily leading to errors.

Requirement of Additional Assumptions and the Impact on Uncertainty

Hoyle and Anderson's recommendation that multicohort models include all present and future recipient cohorts means a CEA would then encompass the entire expected implementation lifetime of an intervention. Modeling all recipient cohorts requires considerable assumptions, especially regarding future cohorts. A common assumption within CEA is that unit costs and health effects remain constant over time.²⁴ However, as noted above, even with the assumption of constant costs and effects, cost-effectiveness can vary between prevalent and incident cohorts; in which case, the number and size of the future incident cohorts modeled will influence aggregate cost-effectiveness estimates.

Although in our example the advent of vaccination means there will be a predictable date by which screening is likely to change, most interventions do not have accurately predictable lifetimes. In a recent study, Hoyle²⁴ used evidence of the lifetime and volume of pharmaceutical interventions in the past to inform assumptions regarding the implementation of new interventions and adjust the cost-effectiveness estimates accordingly. However, the validity and reliability of past evidence of previous interventions as a predictor for new interventions in the future are certainly dubious. Consequently, the cost-effectiveness estimates of aggregate models may seem arbitrary in part, given their reliance on assumptions regarding highly uncertain future use.

Modeling cost-effectiveness over an intervention's entire implementation lifetime significantly increases the scope for uncertainty of estimates. Such additional uncertainty is unwelcome. Uncertainty in input parameters will grow the further into the future they are projected. Although this uncertainty also applies to disaggregate multicohort models that report cost-effectiveness separately for each cohort, it is limited, to a degree, to the cost-effectiveness estimates of future incident cohorts. Given that it will be possible to review the cost-effectiveness evidence for most interventions at a later date and revise the reimbursement decision if necessary, the cost-effectiveness of the current and near-future cohorts is more relevant to the policy choices currently facing decision makers than that of cohorts far in the future. Consequently, modeling an intervention over its entire expected implementation lifetime unnecessarily adds uncertainty to current cost-effectiveness estimates.

Decision Making over Multiple Periods

Cohort heterogeneity arises in our example because of differences in age and screening history at the time of the policy change and vaccination status. Cohort heterogeneity could also occur if costs or effects change over time or if differential discounting of costs and health effects is applied, as is required in Belgium, the Netherlands, and Poland.^{25–27} Hoyle⁸ and Hoyle and Anderson²⁴ use the examples of falling drug prices and differential discounting, respectively, as rationales for modeling interventions' cost-effectiveness over their entire expected lifetimes.

Falling drug prices or the application of differential discounting can lead to inconsistencies between the per-cohort and lifetime analyses of an intervention's cost-effectiveness. For example, a drug may not be cost-effective at its current patent protected price but may be cost-effective if sufficient future cohorts enjoying a lower postpatent price are added to a CEA.²⁴ Conversely, postpatent price reductions of comparator drugs can lead to interventions becoming not cost-effective when analyses are extended to include future periods.²⁸ Similarly, as cost-effectiveness improves with the inclusion of more future cohorts under differential discounting,²⁹ the cost-effectiveness rank order of 2 interventions may switch when compared first on a per-cohort basis and then over their implementation lifetimes, where those lifetimes are unequal. These examples of inconsistencies between the per-cohort and the lifetime perspectives prompt the question as to which is more appropriate, especially as health priority setting is a repeated resource allocation problem reoccurring each year, not a once-off decision over a finite horizon.

DISCUSSION

Determining the most appropriate model structure requires an awareness of both the actual policy question faced by decision makers and any assumptions implicit in the model structure and the presentation of results. Relying on single-cohort models can be interpreted as embodying an assumption that all cohorts will exhibit the same cost-effectiveness as the current incident cohort. However, recommending aggregate reporting of cost-effectiveness

estimates over all current and future cohorts could equally be interpreted as implicitly assuming that interventions cannot be selectively reimbursed and that the lifetime of their use is the appropriate basis for comparisons to other interventions. Consequently, rather than estimating cost-effectiveness over all recipients as suggested by Dewilde and Anderson and Hoyle and Anderson, we support Karnon and others' suggestion that results be reported on a disaggregate basis,¹⁸ except where there is a reasonable rationale to report aggregate estimates, as described above. Reporting the results of multicohort models on a per-cohort basis is equivalent to using multiple single-cohort models. However, this is simply a semantic distinction; what matters is the useful reporting of results to decision makers.

If per-cohort and aggregate optimized strategies differ, then the aggregate strategy will be suboptimal, given the assumption that separately specified interventions can be implemented without any additional costs. However, where it is not costless to provide per-cohort optimized strategies, the tradeoff between costs and benefits of per-cohort optimization should be considered.

Providing disaggregated per-cohort estimates may demand more work, as estimates have to be made for each cohort separately. Given that the principal modeling effort is the gathering of parameter estimates and model specification, the additional effort of generating disaggregated estimates should be relatively small. However, a remaining drawback for decision makers is the potential difficulty of interpreting multiple estimates.

CONCLUSION

The recommendations of the current literature on multicohort modeling that CEAs should include not only the current incident cohort but also the current prevalent and future incident cohorts seem sound. However, the aggregation of cost-effectiveness estimates for all cohorts into a single estimate does not appear useful to the actual choices faced by decision makers in most cases. Therefore, we suggest consideration of the cost-effectiveness of more than just a single incident cohort in cases in which cost-effectiveness is likely to vary between cohorts, with estimates reported on a disaggregate per-cohort basis in addition to the overall estimate for all cohorts. This applies to CEAs in general but is particularly relevant for analyses of screening.

References

1. Torrance, GW.; Siegel, JE.; Luce, BR. Framing and designing the cost-effectiveness analysis. In: Gold, MR.; Siegel, JE.; Russell, LB.; Weinstein, MC., editors. *Cost-Effectiveness in Health and Medicine*. New York: Oxford University Press; 1996. p. 54-81.
2. Dewilde S, Anderson R. The cost-effectiveness of screening programs using single and multiple birth cohort simulations: a comparison using a model of cervical cancer. *Med Decis Making*. 2004; 24(5):486–92. [PubMed: 15358997]
3. Berchi C, Bouvier V, Réaud J-M, Launoy G. Cost-effectiveness analysis of two strategies for mass screening for colorectal cancer in France. *Health Econ*. 2004; 13(3):227–38. [PubMed: 14981648]
4. Maxwell GL, Carlson JW, Ochoa M, Krivak T, Rose GS, Myers ER. Costs and effectiveness of alternative strategies for cervical cancer screening in military beneficiaries. *Obstet Gynecol*. 2002; 100(4):740–8. [PubMed: 12383543]
5. Ramsey SD, Wilschut J, Boer R, van Ballegooijen M. A decision-analytic evaluation of the cost-effectiveness of family history-based colorectal cancer screening programs. *Am J Gastroenterol*. 2010; 105(8):1861–9. [PubMed: 20461066]
6. Stout NK, Rosenberg MA, Trentham-Dietz A, Smith MA, Robinson SM, Fryback DG. Retrospective cost-effectiveness analysis of screening mammography. *J Natl Cancer Inst*. 2006; 98(11):774–82. [PubMed: 16757702]

7. Tosteson ANA, Stout NK, Fryback DG, et al. Cost-effectiveness of digital mammography breast cancer screening. *Ann Intern Med.* 2008; 148(1):1–10. [PubMed: 18166758]
8. Hoyle M, Anderson R. Whose costs and benefits? Why economic evaluations should simulate both prevalent and all future incident patient cohorts. *Med Decis Making.* 2010; 30(4):426–37. [PubMed: 20228287]
9. Davies R, Roderick P, Raftery J. The evaluation of disease prevention and treatment using simulation models. *Eur J Oper Res.* 2003; 150(1):53–66.
10. Mauskopf J. Prevalence-based economic evaluation. *Value Health.* 1998; 1(4):251–9. [PubMed: 16674550]
11. Ahern CH, Shen Y. Cost-effectiveness analysis of mammography and clinical breast examination strategies: a comparison with current guidelines. *Cancer Epidemiol Biomarkers Prev.* 2009; 18(3): 718–25. [PubMed: 19258473]
12. Anderson R, Haas M, Shanahan M. The cost-effectiveness of cervical screening in Australia: what is the impact of screening at different intervals or over a different age range? *Aust N Z J Public Health.* 2008; 32(1):43–52. [PubMed: 18290913]
13. Goldhaber-Fiebert JD, Stout NK, Salomon JA, Kuntz KM, Goldie SJ. Cost-effectiveness of cervical cancer screening with human papillomavirus DNA testing and HPV-16, 18 vaccination. *J Natl Cancer Inst.* 2008; 100(5):308–20. [PubMed: 18314477]
14. Lansdorp-Vogelaar I, Kuntz KM, Knudsen AB, Wilschut JA, Zauber AG, van Ballegooijen M. Stool DNA testing to screen for colorectal cancer in the Medicare population. *Ann Intern Med.* 2010; 153(6):368–77. [PubMed: 20855801]
15. Brisson M, Edmunds WJ. Economic evaluation of vaccination programs: the impact of herd-immunity. *Health Econ.* 2003; 23(1):76–82.
16. Edmunds WJ, Medley GF, Nokes DJ. Evaluating the cost-effectiveness of vaccination programmes: a dynamic perspective. *Stat Med.* 1999; 18(23):3263–82. [PubMed: 10602150]
17. Feuer EJ. Chapter 1: modeling the impact of adjuvant therapy and screening mammography on U.S. breast cancer mortality between 1975 and 2000: introduction to the problem. *J Natl Cancer Inst Monogr.* 2006; (36):2–6. [PubMed: 17032887]
18. Karnon J, Brennan A, Akehurst R. A critique and impact analysis of decision modeling assumptions. *Med Decis Making.* 2007; 27(4):491–9. [PubMed: 17761961]
19. Kuntz KM, Fenwick E, Briggs A. Appropriate cohorts for cost-effectiveness analysis: to mix or not to mix? *Med Decis Making.* 2010; 30(4):424–5. [PubMed: 20634547]
20. Habbema JDF, van Oortmarssen GJ, Lubbe JTN, van der Maas PJ. The MISCAN simulation program for the evaluation of screening for disease. *Comput Methods Programs Biomed.* 1985; 20(1):79–93. [PubMed: 3849380]
21. de Kok IMCM, van Ballegooijen M, Habbema JDF. Cost-effectiveness analysis of human papillomavirus vaccination in the Netherlands. *J Natl Cancer Inst.* 2009; 101(15):1083–92. [PubMed: 19571256]
22. van den Akker-van Marle ME, van Ballegooijen M, van Oortmarssen GJ, Boer R, Habbema JDF. Cost-effectiveness of cervical cancer screening: comparison of screening policies. *J Natl Cancer Inst.* 2002; 94(3):193–204. [PubMed: 11830609]
23. de Kok IMCM, van Rosmalen J, Sasieni P, Dillner J, Iftner T, van Ballegooijen M. A comparison of primary HPV to cytology cervical cancer screening in different European settings: a cost-effectiveness analysis based on a Dutch microsimulation model. *BMJ.* In press.
24. Hoyle M. Accounting for the drug life cycle and future drug prices in cost-effectiveness analysis. *Pharmacoeconomics.* 2011; 29(1):1–15. [PubMed: 21142275]
25. Cleemput, I.; van Wilder, P.; Vrijens, F.; Huybrechts, M.; Ramaekers, D. Guidelines for Pharmacoeconomic Evaluations in Belgium. Brussels: Health Care Knowledge Centre (KCE); 2008.
26. Oostenbrink, JB.; Bouwmans, CAM.; Koopmanschap, MA.; Rutten, FFH. Guidelines for Pharmacoeconomic Research, Updated Version. Diemen, the Netherlands: College voor zorgverzekeringen; 2004.
27. Król, Z. Guidelines for conducting Health Technology Assessment (HTA). Warsaw, Poland: Agency for Health Technology Assessment; 2009.

28. Shih Y-CT, Han S, Cantor SB. Impact of generic drug entry on cost-effectiveness analysis. *Med Decis Making*. 2005; 25(1):71–80. [PubMed: 15673583]
29. O'Mahony JF, de Kok IMCM, van Rosmalen J, Habbema JDF, Brouwer W, van Ballegooijen M. Practical implications of differential discounting in cost-effectiveness analyses with varying numbers of cohorts. *Value Health*. 2011; 14(4):438–42. [PubMed: 21669368]

Figure 1.

Per-cohort and aggregate average cost-effectiveness ratios of cytology and HPV testing compared with no screening. The prevalent cohorts are numbered -30 to -1 starting with the oldest cohort, 0 for the current incident cohort, and from 1 to 15 for the future incident cohorts. The ACERs of cytology and the HPV test are in gray and black, respectively, with dashes for each individual cohort, solid lines for all cohorts on aggregate, and dotted lines for current cohorts on aggregate.

Figure 2.

Per-cohort and aggregate incremental cost-effectiveness ratios of HPV testing compared with cytology. The ICER of HPV testing relative to cytology for those cohorts in which cytology is not weakly dominated by a combination of HPV testing and no screening is shown with dashes. The aggregate ICERs over all cohorts and the current cohorts are shown with the solid and dotted lines, respectively.

Table 1

Discounted Costs, Effects, Average Cost-Effectiveness Ratios of Primary Cytology and Human Papillomavirus Primary (HPV)-Based Screening Compared with No Screening, and Incremental Cost-Effectiveness Ratios of Primary HPV Testing Compared with Primary Cytology for Selected Cohorts and Multicohort Aggregates

Cohort	Cytology Primary Test Screening				HPV Primary Test Screening			
	Effects, QALYs, 000s ^a	Costs, €Millions ^b	ACER, €QALY ^c	Proportion of Total Costs, %	Effects, QALYs, 000s	Costs, €Millions ^b	ACER, €QALY ^c	ICER, €QALY ^d
-30	38	2296	60 200	0.8	58	2800	48 700	N/A
-25	84	4347	51 600	1.6	112	5325	47 400	N/A
-20	144	6469	45 000	2.4	177	7954	44 900	N/A
-15	237	7806	33 000	2.9	274	9645	35 200	49 400
-10	433	8051	18 600	3.0	486	10 039	20 700	37 900
-5	611	7110	11 600	2.6	672	9036	13 500	31 800
0	795	7917	10 000	2.9	851	10 602	12 500	47 700
5	709	7049	9900	2.6	759	9439	12 400	47 500
10	614	6115	10 000	2.3	658	8188	12 500	47 700
15	205	5623	27 400	2.1	208	7326	35 200	592 800
-30 to 0	7797	171 858	22 000	63.4	8903	214 158	24 100	38 300
1 to 15	8421	99 085	11 800	36.6	8968	131 812	14 700	59 800
-30 to 15	16 218	270 943	16 700	100.0	17 871	345 970	19 400	45 400

^aQuality-adjusted life-years.

^bCosts in 2008 prices.

^cAverage cost-effectiveness ratio compared with no screening.

^dIncremental cost-effectiveness ratio of primary HPV testing compared with primary cytology.