Supplemental Table 2. Characteristics of included systematic reviews

	Smartt 2009 [14]	OHTAC 2010 [15]	Ward 2011 [16]	Hornberger 2012 [17]	Rouzier 2013 [18]
Dates of search and databases searched	2007 to September 2009 • EMBASE • MEDLINE • The Cochrane Library • CINAHL The searchers were limited to English (EMBASE, MEDLINE, and CINAHL). Non-human studies were excluded from EMBASE and MEDLINE searches.	1 January 2006 to19 March 2010 • OVID MEDLINE • MEDLINE In-process • EMBASE • CINAHL • INAHTA	January 2009 to May 2011 • MEDLINE: 1950 to May 2011 • MEDLINE In-process: 1950 to May 2011 • EMBASE: 1980 to May 2011 • Cochrane Central Register of Controlled Trials (CENTRAL): Issue 3, 2011 • Cochrane Database of Systematic Reviews – Issue 8, 2011 • NHS Database of Abstracts of Reviews of Effectiveness (DARE) – via Cochrane Library, Issue 3, 2011 • Health Technology Assessment Database (HTA) – via Cochrane Library, Issue 3, 2011 • BIOSIS previews: 1926 to May 2011 • Web of Knowledge: 1899 to May 2011 Additional sources were searched including contacting manufacturers, experts in the field, screening reference list of included studies, citation searching of key papers. Conference proceedings	2000 to 2011 • PubMed • CINAHL • Cochrane Database of Systematic Reviews Searches were supplemented by using Web of Science to identify additional studies, and searching publications from the American Society of Clinical Oncology, the San Antonio Breast Cancer Symposium, and the European Society of Medical Oncology.	1 January 2002 to 7 January 2012 • PubMed • Cochrane Library HTA websites of the UK, Canada, Australia, and USA were searched Conference proceedings were search from 2009 to 2012: • San Antonio Breast Cancer Symposium • American Society of Clinical Oncology • European Breast Cancer Conference • St. Gallen Oncology Conference • European Society of Medical Oncology • European Cancer Organization • International Society for Pharmacoeconomics and Outcomes Research

			from the St. Gallen International Breast Cancer were screened as well as relevant reviews and guidelines.		
Population	Women with early-stage breast cancer	Women with early stage (I-IIIa) invasive breast cancer that is: • Estrogen receptor (ER) positive and/or progesterone receptor (PR) *positive* • Lymph node (LN) *negative* • Human epidermal growth factor receptor 2 (HER2) *negative*	Women with early-stage invasive breast cancer (stage I, II, or III) • Lymph node (LN) *negative* or *positive (up to 3)* • Estrogen receptor (ER) *positive* or *negative* • HER2 *positive* or *negative*	Women with early-stage breast cancer	"early-stage, nonmetastic breast cancer patients who underwent curative-intent surgery"
GEP test(s)	Oncotype DX MammaPrint H/I ratio test	Oncotype DX	Randox Breast Cancer Array MammaPrint BluePrint PAM50 Oncotype DX Breast Cancer Index	Mammostrat MammaPrint Oncotype DX Molecular Grade Index BreastOncPx	Commercially available multi-gene assasys (MGAs) : OncotypeDx MammaPrint BluePrint PAM50 Breast Cancer Index Mammostrat NPI+
Aims and objectives/Key Questions focused on clinical utility	To assess the evidence of GEP tests for improving prognostic accuracy, treatment choice, and health outcomes.	• What is the predictive value of Oncotype-DX? • How does Oncotype-DX impact patient quality of life and clinical/patient decision-making?	"The overall aim of the assessment is to assess the clinical effectiveness, effect on patient outcomes, and cost-effectiveness of the new GEP and expanded IHC tests".	"The primary aim of ourstudy was to systematically grade the Level-of-Evidence (LOE) instudies that assessed the clinical validity/utility of risk stratifiersforESBC. A secondary aim was to document studies that provided evidenceon changes in practice patterns and health economic implicationsof the stratifiers".	"This systematic review summarizes the available evidence from health economic analyses on MGAs and molecular markers in breast cancer"

Inclusion/ exclusion criteria	"Articles were considered to be ineligible if the study: • only applied to breast cancer biology (ie were not clinical studies) • did not involve OncotypeDX™ or MammaPrint® or H/I gene expression profiling tests • did not involve original data or original data analysis • did not involve breast cancer patients • was not reported in English • did not apply to any of the key questions of the review • other (give reason) • eligibility was unclear. " letters, editorial, comments and news articles were excluded conference abstracts were included if they presented relevant data	***"Inclusion Criteria*** ▪ Any observational trial, controlled clinical trial, randomized controlled trial (RCT), meta-analysis or systematic review that reported on the laboratory performance, prognostic value and/or predictive value of Oncotype-DX testing, or other outcome relevant to the Key Questions, specific to the target population was included. ***Exclusion Criteria*** ▪ Studies that did not report original data or original data analysis, ▪ Studies published in a language other than English, ▪ Studies reported only in abstract or as poster presentations (such publications were not sought nor included in this review since the Medical Advisory Secretariat (MAS)does not generally consider evidence that is not subject to peer review nor does the MAS consider evidence that lacks detailed description of methodology)."	**Exclusion criteria:** ▪ Animal models ▪ Preclinical and biological studies ▪ Editorials and opinion pieces ▪ Non-English publications ▪ Conference abstracts ▪ Studies related to breast cancer biology ▪ Studies conducted in the neo-adjuvant treatment setting "…studies will be excluded if they…appear to be methodologically unsound, or do not report methods and/or results in the necessary detail."	**Inclusion criteria:** ▪ Original data on an assay's ability to predict risk of progression or response to chemotherapy ▪ Studies reporting assay's impact on clinical decisions, practice patterns, or economics **Exclusion criteria:** ▪ Studies reporting individual components of the assay ▪ Pathophysiological or in vitro studies ▪ No original data ▪ Non early-stage breast cancer ▪ Non-English language publication ▪ Did not report on the clinical validity, clinical decisions, or economics	**Inclusion criteria:** "publications assessing the cost-effectiveness or budget impact of prognostic MGAs"

Included studies addressing key questions related to clinical utility	OncotypeDx: • KQ 1: 3 studies (Akashi-Tanaka 2009 [20]; Kok 2009 [21]; Li 2009 [22]) • KQ 2: 3 studies (Asad 2008 [27]; Henry 2009 [29]; Rayhanabad 2008 [28]) MammaPrint: • KQ 2: 1 study (Bueno-de-Mesquita 2007 [39])	Oncotype Dx: • KQ 1: 2 studies(Paik 2006 [23]; Albain 2010 [24]) • KQ 2: 3 studies (Asad 2008 [27]; Geffen 2009 [30]; Lo 2010 [31])	OncotypeDx: • KQ 1: 3 studies (Albain 2010 [24]; Tang 2011 [25]; Tang 2010 [26]) • KQ 2: 4 studies (Ademuyiwa 2011 [32];Geffen 2009 [30]; Holt 2011 [33]; Lo 2010 [31]) MammaPrint: • KQ 2: 1 study (Gevensleben 2010 [40])	Oncotype DX: • KQ 1: 3 studies (Paik 2006 [23]; Tang 2011 [25]; Albain 2010 [24]) • KQ 2: 10 studies (Ademuyiwa 2011 [32]; Asad 2008 [27]; Henry 2009 [29]; Hornberger 2011 [36]; Joh 2011 [37]; Klang 2010 [35]; Lo 2010 [31]; Oratz 2007 [34]; Partin 2011 [38]; Rayhanabad 2008 [28]) • KQ3: 8 included studies were not considered in this overview as their methodological quality was not evaluated. MammaPrint: • KQ 2: 1 study (Bueno-de-Mesquita 2007 [39]) • KQ 3: 3 included studies were not considered in this overview as the methodological quality was not evaluated.	OncotpyeDX: • KQ 3: 22 studies (Hornberger 2005 [41]; Lyman 2007 [42]; Kondo 2008 [43]; Cosler 2009 [44]; de Lima Lopes 2010 [45]; Klang 2010 [35]; O'Leary 2010 [46]; Tsoi 2010 [47]; de Lima Lopes 2011 [48]; Holt 2011b [49]; Hornberger 2011 [36]; Kondo 2011 [50]; Hassan 2011 [51]; Lacey 2011a [52]; Lacey 2011b [53]; Paulden 2011 [54]; Ragaz 2011 [55]; Vanderlaan 2011 [56]; Hall 2012 [57]; Lamond 2012 [58]; Madaras 2012 [59]; Wilson 2012[60]) MammaPrint: • KQ 3: 5 studies (Oestreicher 2005 [64]; Zarca 2009 [65]; Chen 2010 [63]; Retel 2010 [61]; Kondo 2012 [62])
Quality assessment	Quality of individual studies was assessed with a 44-item checklist based on REMARK criteria, and STARD criteria were used to compare quality between studies.	The quality of the body of evidence was assessed according to the GRADE criteria for: 1) quality; 2) consistency; and 3) directness.	Quality of individual studies was assessed using six dimensions related to internal validity purposed by Altman 2001: 1) sample of participants; 2) follow-up of participants; 3) outcomes; 4) prognostic variables; 5) analysis; and 6) treatment subsequent to inclusion in cohort.	Rated included studies according to the level of evidence outlined by Simon 2009[19].	"The Quality Health Economic Studies (QHES) instrument was used to evaluate the quality of economic evaluations"