Appendix 2. Methodological quality of observational studies that evaluated screening programs (CASP appraisal checklist)

Item #	Agardh 2011[25]	Kohner 2001 [27]	Kristinsson 1995 [28] Olafsdottir 2007[31]	Looker 2013 [40]	Maguire 2005 [29]	Misra 2009 [30] Jones 2012 [26]	Soto-Pedre 2009 [32]	Stratton 2013 [39]	Thomas 2012 [1]	Younis 2003a [34] Younis 2003b [33]
Study design	Prospective cohort study	Retrospective cohort study	Retrospective cohort study	Retrospective cohort study	Retrospective cohort study	Dynamic cohort study	Retrospective, two arm cohort	Retrospective cohort study with risk stratification	Retrospective cohort study	Retrospective cohort study
1. Did the study address a clearly focused issue?	To determine the incidence of STDR in T2DM patients without DR at baseline.	To establish the relationship between severity of DR and the progress to photocoagulation for T2DM patients.	Whether biennial examinations suffice for T1DM and T2DM patients without DR.	To establish whether there are any subgroups of the population at low risk of transition to a referable state of retinopathy in a 2 year interval.	To identify the optimum screening frequency for children and adolescents with T1DM.	To estimate incidence, prevalence, and progression rates of DR in relation to baseline DR severity and other prognostic factors in T2DM patients.	To estimate safe screening intervals for sight-threatening diabetic retinopathy.	To develop a simple model to estimate risk of STDR.	To determine the incidence of referable DR in people with T2DM over a 4 year period.	To investigate incidence of DR in relation to severity of DR in patients with T1DM and T2DM and calculate optimum screening intervals.
	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
2. Was the cohort recruited in an acceptable way?	Study recruited cohort of T2DM patients at one hospital that fit criteria.	Cohort of 3,709 T2DM patients with good quality retinal photographs selected from a wider cohort.	Two Icelandic cohorts followed for comparison (81 children aged<15 and 185 adult patients with T1DM and T2DM with no retinopathy; selection process not clear, was this the total population of eligible participants?	Cohort of 155,114 people with T1DM and T2DM diagnosed before May 2008 who were screened by the Scottish Screening Service and had either no retinopathy, mild background retinopathy, observable background retinopathy or observable background maculopathy in the worst eye at baseline and at least one subsequent screen.	Cohort of 668 children and adolescents recruited from a hospital. All patients who had a baseline and at least one follow up screen before age 20 were included.	Cohort of all people with T2DM (n=20,788) identified through GP diabetes register were screened; patients with STDR were excluded. It was a dynamic cohort where individuals enter and leave the cohort at different times.	Two cohorts of patients referred to one hospital between 1998 and 2004 with at least one baseline screening and one subsequent screening. Cohort 1: (n-=286), no DR at baseline Cohort 2: Mild nonproliferative DR at baseline (n=144).	Population based cohort of 14,554 people aged 12 or over with T1DM and T2DM who were screened by the Gloucestershire Diabetic Eye Screening Service between 2005 and 2010 and had no DR or mild nonproliferative DR at two consecutive annual screenings.	Cohort of 57,199 patients with T2DM aged > 12 registered with a GP with no DR at baseline.	Cohort of all patients with T1DM (n=501) and T2DM (n=4,770) registered with GPs and screened with DR data available at baseline and at least one further screening.
	Yes	Yes	Can't tell	Yes	Yes	Yes	Yes	Yes	Yes	Yes

3. Was the exposure accurately measured to minimise bias?	All patients who were available for follow up invited for assessment after 3 years. Reminders were sent to non-responders, so some may have been followed up after a longer time period. This is not reported.	Participants of underwent a full medical examination, including retinal examination yearly.	Screening intervals for group 2 described as 'at least annually'; screening intervals for group 1 (children) not reported.	Patients screened yearly under a standard protocol.	Patients were followed longitudinally and different screening frequencies were observed. Interval lengths were recorded and patients were grouped by interval length for analysis.	All patients invited for screening annually, although people with clinical indications or those with questionable images or technical problems were rescreened at 6 months. It is not clear if non attendees were followed up or whether actually attendance was at sooner or later than the 12 month mark.	Study reviews all patients with a baseline screening and subsequent screening. It is not clear what the standard interval is for this cohort.	Patients screened yearly under a standard protocol.	Patients screened yearly under a standard procedure.	Patients with no DR or background DR were screened yearly. Patients with DR without sight threatening maculopathy were followed up every 6 months.
	Can't tell	Yes	Can't tell	Yes	Yes	Can't tell	Can't tell	Yes	Yes	Yes
4. Was the outcome accurately measured to minimise bias?	Eyes were assessed using the same method and graded by specially trained ophthalmic nurses using a validated international scale. HbA1c was measured using the same method.	The same protocol for screening and grading was used for the cohort. All images were graded by physicians at the Retinopathy Grading Centre.	Standardised examination and reporting method used across both patient groups. Retinopathy reported for each patient was based on the worst eye.	Methodology for screening and quality assurance is not described.	Outcomes were verified on a proportion of photographs graded independently by a second grader for quality control and 'good agreement' was found between the graders. Outcome classification was standardised.	Some measurement bias is possible over the period of the study. Two different methods of imaging and grading systems (scales) were used over the period of the study.	Standard procedure for screening described for all patients.	Standard protocol for screening and grading used for all patients.	Standard protocol for screening and grading used for all patients.	Standardised protocol for screening, grading (Wisconsin algorithm), and reporting DR. Provision for rescreening or validating results in place.
	Yes	Yes	Yes	Can't tell	Yes	No	Yes	Yes	Yes	Yes
5. Have the authors identified all important confounding factors?	The authors measured HbA1c at baseline showing no significant difference between attendees and non-attendees.	The study did not consider confounders such as HbA1c or therapy allocation.	For children (group 1), onset of puberty was considered. Age at onset and duration of disease were reported for adults in group 2.	At baseline, BMI, HbA1c, blood pressure were recorded nearest to the initial screening. Age, sex, and diabetes duration were recorded for T1DM and T2DM.	At each eye examination, height, weight, pubertal staging, blood pressure and HbA1c and DM duration were recorded. Other confounders	Age, duration of DM, DM treatment and hypertension treatment were measured at baseline. Smoking history, blood glucose, blood pressure,	At baseline, Age, sex, glycated hemoglobin level, type of DM, diabetes duration and treatment of DM were recorded. Number of	The study did not consider confounders, but are allocated to a group in accordance to identified risk factors.	Age at DM diagnosis, duration of DM, DM treatment and sex were recorded at baseline. HbA1c percentage was not used in the study.	The authors considered age, duration of DM, age at diagnosis, follow up duration, number of screening visits, sex and treatment at baseline.

	Age at diagnosis, duration of DM and DM treatment method were recorded.		HbA1c was not reported or analysed in either group.		such as DM treatment were not recorded.	sex and ethnicity were not recorded as part of the screening programme.	screening visits was not recorded.			
	Yes	No	No	Yes	Yes	No	Yes	No	No	Yes
6. Was the follow up of the subjects complete enough?	Of the 1,691 recruited subjects 1,322 were followed up. The authors state that HbA1c levels did not differ in those with and without follow up data.	The study only included patients who had all the required follow up images and data.	The study only included patients who were still alive at the end of the 10 year follow up period. There was no information on patients who may have died during the follow-up whose outcomes may have been different to the outcomes in the overall group.	The study only included patients who had all the required follow up images and data	Only patients with a baseline and follow up screen were included. It appears that screens were conducted as a diagnostic rather than as a screening programme, meaning that these findings could perhaps represent either more unwell children or children or families that were more likely to seek medical support.	Patients with evidence of retinopathy were referred to the hospital eye service and the authors report that the quality of data referring to risk factors and outcomes was poor meaning that they were unable to provide analysis for these patients.	The study only included patients who had all the required follow up images and data.	The study only included patients that met the criteria for recruitment	Of the 57,199 individuals recruited at baseline, 7,436 (13%) did not attend a further screening. Of the 7,436 subjects, 449 were not eligible for a second screen (recruited less than 12 months from the end of the study). It is unclear why the remaining 6,987 patients did not attend the second screen The authors stated that the non-attendees were more likely to be older and have a longer duration of DM.	A large proportion of patients (31%, n=2388) had not undergone a repeat screening by the end of the study period and were not included in the cohort analysis of baseline data plus one other screen. Non-attenders to a second appointment may have differed from attenders factors which could affect onset of DR. Non-participation rate in T1DM patients was high. Of the 1050 eligible patients, only 79% (n=831) accepted invitation for a baseline screen, of whom only 501 participated in a follow up screen.
	Yes	No	Can't tell	No	No	No	No	No	No	No
7. What are the results of this study?	See Table 1	See Table 1	See Table 1	See Table 1	See Table 1	See Table 1	See Table 1	See Table 1	See Table 1	See Table 1
8. How precise are	95% CIs not reported	The proportion of patients requiring	Only proportions of outcomes are	P-values were used when comparing	The authors used General Estimating	Annual incidence for was provided	95% CIs were reported for all	Hazard ratios that compare the risk	The reported incidence or	95% CIs were reported for all

the results?		photocoagulation 95% CIs (reported only graphically) were wider for patients with severe DR, probably due to smaller numbers and as time progresses	reported for both groups. No further statistical analysis is provided. 95% CIs not reported	incidence data between patients with no visible retinopathy and mild background retinopathy and between no visible retinopathy and observable retinopathy and maculopathy.	Equations (GEEs) to compare risk of retinopathy at yearly intervals to the baseline based on the available data for the whole group and the two age divisions. P-values were used when comparing incidence data between patient groups for whom the corresponding estimates were less precise.	with 95% CIs, which are wider for non-proliferative DR as numbers are smaller, particularly as time progresses. 95% CIs were narrow for up to 4 years for patients with no retinopathy at baseline.	findings; although the sample size is much smaller than other studies.	of DR progression between groups are reported with 95% confidence intervals. P Values are used when comparing the incidence of more serious DR between groups.	progression of DR from 1 to 4 years had narrow 95% CIs.	findings; given the large study group, the reported estimates were precise.
	Can't tell	Can't tell	Can't tell	Can't tell	Can't tell	Yes	Can't tell	Yes	Yes	Yes
9. Do you believe the results?	Measurement and grading methods are robust and characteristics of followed up and non-followed up patients are reported as not differing making the results believable.	The cohort is large and the design of the study is robust with each patient having the same data evaluated with the same protocol making the results believable.	For both groups, the study population is small. The lack of precision and short follow up periods would make the findings difficult to rely, however considered in the wider context of the review the findings are in line with other studies	Large cohort. As the screening and grading methodology is not clearly outlined, it is not possible to be assured of the results reported for individual patients.	The findings are in line with findings from other groups and significance is tested which makes the findings believable. However the numbers are much smaller than other studies which would promote caution if relying on these findings only.	The lack of description of the characteristics of those not attending for screening or taking part in the programme is concerning as they may have characteristics such greater non-compliance to diabetes treatment that may have affected progression to retinopathy.	The findings are in line with findings from other groups and significance is tested which makes the findings believable. It is useful to see the two cohorts analysed separately although the study population for each cohort is small. The authors report the limitations of the specificity of the screening methodology.	Large cohort with robust recruitment and measurement methods.	The cohort is large and the authors are careful to report any potential limitations of the study. The analysis of the results is robust, making the findings believable.	The study had a robust methodology and the authors acknowledge the limitations of the smaller sample size and the impact of the non-participants on the findings. The results are more believable for the larger group of patients with no DR than for those with mild pre-proliferative DR at baseline.
	Yes	Yes	No	Can't tell	Yes	Can't tell	Yes	Yes	Yes	Yes

10. Can the results be applied to the population of England?	Compliance to screening has been reported to be higher in Sweden than in England meaning that non-compliance may lead to a longer than 3 year interval in a larger subgroup of a screening cohort.	Study uses a sample of UK patients making the findings applicable to the rest of England acceptable.	The study uses a small Icelandic population which makes the findings less applicable to the English population as characteristics such as screening compliance and DM control can vary	Study uses a Welsh population and is broadly applicable to a UK population.	The study is based in Australia, which in terms of access to care and treatment and overall patient characteristics is similar to an English population.	Study is of a large population in the UK, making application to the rest of England acceptable.	The study uses a Spanish population in one small area. No information is given about the screening policy or compliance in a Spanish population.	The study uses a population form Gloucestershire in England which is broadly comparable to the UK population	Study uses a Welsh population and is applicable to a UK population.	Study is of a large population in Liverpool, UK making application to the rest of England applicable.
	Can't tell	Yes	No	Yes	Yes	Yes	Can't tell	Yes	Yes	Yes
11. Do the results fit with other evidence?	The authors conclude that longer screening interval is safe for low risk T2DM patients with no retinopathy; however, the recommendation for a 3 year interval is longer than other studies recommend.	As with other studies, the authors found that retinopathy incidence was low in patients without retinopathy over a 3 - 6 year period.	The authors conclude that biennial screening for both T1DM and T2DM without retinopathy is reasonable. Other studies have reported that people with T1DM should remain on yearly intervals.	As with other studies, the authors report lower risk of retinopathy in patients with T2DM and no background retinopathy and with no visible retinopathy at two consecutive screens. Although the study does not go as far as recommending biennial screening they state that the risk of low risk patients developing retinopathy is small.	This study in children, finds that STDR is unlikely to occur within an interval of 2 years in patients with no baseline DR. As with other studies, the authors recommend that upon detection of retinopathy, frequency should change to annual.	As with other studies, the authors find incidence of retinopathy over 5-10 years in patients with no retinopathy at first screen as low, recommending that intervals longer than one year may be appropriate for this group of people.	As with other studies, the authors report low incidence of retinopathy in patients with no retinopathy at baseline over a 6 year period. As with other studies, poor metabolic control is identified as a risk factor. The recommendation for a 3-4 year interval for patients with no retinopathy baseline is longer than other studies recommend.	As with other studies, the authors report that those in higher risk groups (defined by presence of DR at baseline or previous screen) are most likely to progress to more serious DR. The study does not make explicit recommendations on the frequency of screening.	Similarly to other studies, the authors recommend longer intervals for patients with no retinopathy at baseline. Identified risk factors were similar to those identified in other studies (age, insulin use and duration of diabetes).	Results of this generally fit with the other studies that include a small cohort of T1DM patients. Optimum screening interval of 5.7 years for a 95% likelihood of remaining free of STDR for those without DR at baseline is much longer than the intervals recommended by other studies Optimum screening interval for those with background retinopathy was 1.3 years, which is more in line with findings from other studies.
	No	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
	STDR=sight-threatening diabetic retinopathy; T1DM= type 1 diabetes mellitus; T2DM=type 2 diabetes mellitus; DR=diabetic retinopathy; DM=diabetes mellitus ; 95% CI=95 percent confidence interval									