Appendix 1 Data extraction tables

Study 1 of 8 – Cachulo and colleagues

Reference and design	Diagnostic tests	Participants	Outcome measures
Condition being diagnosed / detected: Choroidal neovascular-isation (CNV) in exudative AMD **First author:** Cachulo[99] **Publication year:** 2011 **Country:** Portugal **Study design:** Prospective observational longitudinal 2 year study **Number of centres:** One **Funding:** Not reported **Competing interests:** Not reported 1 author appears to be employed by Pfizer Inc.	**Index test:** Fundus autofluorescence (FAF): acquired with confocal scanning laser ophthalmoscopy (cSLO) HRA II (Heidelberg Retina Angiograph) Excitation 488nm; barrier filter beginning at 500nm. Each FAF image was compiled from at least 17 single scans in movie mode and automatically aligned and averaged. **Reference standard:** Fluorescein angiography (FA): acquired using the HRA II (Heidelberg Retina Angiograph) scanning laser ophthalmoscope **Comparator:** 1) Colour fundus photography 2) Fluorescein angiography 3) Indocyanine green angiography 4) Optical coherence tomography 5) Retinal angiography (retinal leakage analysis – RLA – measuring retinal fluorescein leakage from the blood stream into the vitreous using cSLO)	**Number of participants:** 62 (52 included in analysis) **Number of eyes:** 52 **Sample attrition/dropout:** 52 participants completed the 2 year follow-up, dropout was due to death (4 patients), withdrawal of informed consent (4 patients), hospitalisation (1 patient), loss to follow-up (1 patient treated in another country) **Selection of participants:** Patients with neovascular AMD in one eye and early AMD in the fellow eye (study eye) at risk for development of CNV. Not reported whether patients selected consecutively **Inclusion criteria for study entry:** 1) Older than 50 years 2) Any race and either sex 3) Clinical diagnosis of wet AMD in one eye (non-study eye) 4) Presence of the following characteristics in the	**Primary outcome of study:** Presence of conversion from early AMD to wet AMD: sensitivity and specificity (repeated imaging assessments at 6-monthly intervals for 2 years or until CNV presence was confirmed in the study eye) **Other relevant outcomes:** None **Diagnostic threshold:** FAF (observed from results, but not stated in methods): patchy pattern; reticular pattern; speckled pattern; focal increased pattern; lacelike pattern FA: not reported **Recruitment dates:** Not reported

		study eye: a) 5 or more intermediate soft drusen >63µm or 1 large soft druse >125µm, and/or confluent drusen within 3,000 µm of the foveal centre b) With or without pigmentary changes **Exclusion criteria for study entry:** 1) Current or past medical condition that would preclude scheduled visits or completion of the study 2) Current or past history of ophthalmic disease in the study eye (other than AMD), that would likely compromise the visual acuity of the study eye 3) Clinical signs of myopic retinopathy or refractive power of >8 diopters or funduscopic evidence of degenerative myopia 4) Past history of intraocular surgery within 60 days prior to enrolling in the study 5) Evidence of past or present CNV in the study eye	

Participant characteristics	
Sex, m:f (%male)	26:26 (50)
Age, years, mean (SD)	76 (6), range 56-92

Results – FAF versus FA

Calculations are based on number of eyes (single eyes of 52 subjects)	**Population with disease on FA reference standard**		**Population without disease on FA reference standard**		**Total**
FAF imaging positive	15	a	23	c	38
FAF imaging negative	2	b	12	d	14
Total	17		35		52

Diagnosis		**95% CI**
Clinical sensitivity a / (a + b)	88.24 %	63.52 to 98.20
Clinical specificity d / (c + d)	34.29 %	19.15 to 52.21
PPV a / (a + c)	39.47 %	24.05 to 56.61
NPV d / (b + d)	85.71 %	57.16 to 97.80
Positive likelihood ratio [sensitivity/(100-specificity)]	1.34	1.00 to 1.80
Negative likelihood ratio [(100-sensitivity)/specificity]	0.34	0.09 to 1.36
Diagnostic odds ratio (a x d)/(b x c)	3.91	0.77 to 20.02

Comments: Calculations do not agree with values reported in paper. Reported values for FAF are: sensitivity 93%, specificity 37%, positive predictive value 57% and negative predictive value 93%. This may be because of different ways that the reviewer and authors categorised the 2 eyes in FAF in which the pattern of autofluorecence could not be determined because of poor quality images.

Interpretability and acceptability of test	
Numbers excluded from analysis due to poor image quality	2/52 (3.85%)
Inter-observer agreement	Not reported
Intra-observer agreement	Not reported
Test acceptability (patients / clinicians)	Not reported
Adverse events	Not reported

AMD: age-related macular degeneration; CNV: choroidal neovascularisation; cSLO: confocal scanning laser ophthalmoscopy; FA: fluorescein angiography; FAF: fundus autofluorescence; NPV: negative predictive value; PPV: positive predictive value.

Cachulo and colleagues[99] critical appraisal criteria (based on Reitsma et al.[81] adaptation of the QUADAS Tool[93])

	Item	Description	Judgement
1	Was the spectrum of patients representative of the patients who will receive the test in practice?	Study is prospective but unclear if it involved consecutive patients. Participants had confirmed CNV in one eye, so may be an atypical case-mix	No
2	Is the reference standard likely to classify the target condition correctly?	FA is described as the gold standard for assuming conversion from early AMD to wet AMD	Yes
3	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	It is reported that each patient underwent study assessments at baseline and every six months for two years. However, no detail is given about the time between tests	Unclear
4	Did the whole sample or a random selection of the sample receive verification using the intended reference standard?	Of 62 patients enrolled in the study, 10 dropped out, due to death (n=4), withdrawal of consent (n=4), hospitalisation (n=1) and loss to follow up (n=1). It is confirmed in the results section that 52/52 of the remaining patients underwent the fluorescein angiography test	Yes
5	Did patients receive the same reference standard irrespective of the index test result?	The results confirm that 52/52 patients underwent the fluorescein angiography test	Yes
6	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes	Yes
7	Were the reference standard results interpreted without knowledge of the results of the index test?	Not reported	Unclear
8	Were the index test results interpreted without knowledge of the results of the reference standard?	Not reported	Unclear
9	Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	Not reported	Unclear
10	Were uninterpretable or intermediate test results reported?	It is stated that in 2 eyes the pattern of autofluorescence could not be determined because of poor quality images	Yes
11	Were withdrawals from the study explained?	Yes	Yes

Study 2 of 8 – Dinc and colleagues

Reference and design	Diagnostic tests	Participants	Outcome measures
Condition being diagnosed / detected: Cystoid macular oedema (CMO) (secondary to diabetic retinopathy, retinal vein occlusions, uveitis, cataract surgery, epiretinal membrane or age-related macular degeneration) **First author:** Dinc[83] **Publication year:** 2010 **Country:** Turkey (not stated explicitly) **Study design:** Patients were selected from a FAF database (no further details given); informed consent was obtained from all patients, suggesting the study was prospective **Number of centres:** Not explicitly reported but appears to be single centre **Funding:** No information provided **Competing interests:** No information provided	**Index test:** Fundus autofluorescence (FAF) acquired with confocal scanning laser ophthalmoscopy (cSLO) (Heidelberg Retinal Angiograph 2, Heidelberg Engineering, Germany). View mode 30°; pupil dilated to a diameter ≥ 6 mm. Excitation 488nm; barrier filter 500nm. Stated that a mean of 9 frames was obtained. **Reference standard:** Fluorescein angiography (FA). Method not reported except that in the late phase of FA, path-ognomonic leakage of fluorescein at the fovea in a petaloid configur-ation with feathery margins was considered as CMO. **Comparator:** Optical coherence tomography (OCT) (type not reported)	**Number of participants:** 55 **Number of eyes:** 67 **Sample attrition/dropout:** None (results reported for all eyes) **Selection of participants:** Stated only that the patients diagnosed with CMO were selected from a FAF database (no criteria specified) **Inclusion criteria for study entry:** Patients with CMO secondary to diabetic retinopathy, retinal vein occlu-sions, uveitis, cataract surgery, epiretinal membrane or age-related mac-ular degeneration **Exclusion criteria for study entry:** Eyes with significant media opacity, cataract, poor FAF images, or having subfoveal serous retinal detachment on OCT	**Primary outcome of study:** Detection of CMO by FAF and FA **Other relevant outcomes:** Central macular thickness assessed by OCT (data not extracted here) **Diagnostic threshold:** Not explicitly stated but implied to be increased autofluorescence in a round or oval fashion at the fovea (example image given for reference) **Recruitment dates:** Unclear. Stated that patients were selected from the FAF database between January 2008 and June 2009

Participant characteristics	
Sex, m:f (%male)	28:27 (51)
Age, years, mean (SD)	62.1 (14.4)
Origin of CMO (n= no. of eyes)	Diabetic retinopathy, n=36 Branch retinal vein occlusion, n=13 Macular epiretinal membrane, n=5 Age-related macular degeneration, n=5 Uveitis, n=4 Cataract extraction, n=3 Central retinal vein occlusion, n=1

Results – FAF compared against FA

Calculations are based on the numbers of eyes (both eyes of 12 subjects and single eyes of 43 subjects)	**Population with CMO on FA**		**Population without CMO on FA**		**Total**
FAF imaging positive	64	a	2	c	66
FAF imaging negative	1	b	0	d	1
Total	65		2		67

Diagnosis		**95% CI**
Clinical sensitivity a / (a + b)	98.46%	91.69 to 99.74
Clinical specificity d / (c + d)	0.00%	0.00 to 80.71
PPV a / (a + c)	96.97%	89.46 to 99.54
NPV d / (b + d)	0.00%	0.00 to 83.45
Positive likelihood ratio [sensitivity/(100-specificity)]	0.98	0.96 to 1.01
Negative likelihood ratio [(100-sensitivity)/specificity]	Not calculable	
Diagnostic odds ratio (a x d)/(b x c)	8.60	0.28 to 268.48

Comments: Diagnostic outcomes are not reported in paper – calculated by reviewer

Interpretability and acceptability of test	
Numbers excluded from analysis due to poor image quality	None – results are reported for all 67 study eyes
Inter-observer agreement	Not reported
Intra-observer agreement	Not reported
Test acceptability (patients / clinicians)	Not reported
Adverse events	Not reported

CMO: cystoid macular oedema; cSLO: confocal scanning laser ophthalmoscopy; FA: fluorescein angiography; FAF: fundus autofluorescence; NPV: negative predictive value; OCT: optical coherence tomography; PPV: positive predictive value

Dinc and colleagues[83] critical appraisal criteria (based on Reitsma et al.[81] adaptation of the QUADAS Tool[93])

	Item	Description	Judgement
1	Was the spectrum of patients representative of the patients who will receive the test in practice?	Unclear if study is prospective, but it involved consecutive patients. CMO was secondary to a range of conditions, and patients with CMO and serous retinal detachment were excluded, so is an atypical case-mix	No
2	Is the reference standard likely to classify the target condition correctly?	FA is described as the gold standard for detecting CMO	Yes
3	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Order, but not timing, of tests specified	Unclear
4	Did the whole sample or a random selection of the sample receive verification using the intended reference standard?	Whole sample	Yes
5	Did patients receive the same reference standard irrespective of the index test result?	Yes	Yes
6	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Separate tests applied at different times	Yes
7	Were the reference standard results interpreted without knowledge of the results of the index test?	Stated that the data on FA and FAF images were evaluated by as single clinician; masking not reported	Unclear
8	Were the index test results interpreted without knowledge of the results of the reference standard?	FAF images were obtained prior to FA images	Yes
9	Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	Not reported	Unclear
10	Were uninterpretable or intermediate test results reported?	All eyes included in analysis but not stated whether image quality was an inclusion or exclusion criterion	No
11	Were withdrawals from the study explained?	Results data reported for all eyes – implies no withdrawals	Not applicable

Study 3 of 8 – Hogg and colleagues

Reference and design	Diagnostic tests	Participants	Outcome measures
Condition being diagnosed / detected: Reticular pseudodrusen (RPD) in age-related macular degeneration (AMD) **First author:** Hogg[96] **Publication Year:** 2014 **Country:** Italy, Portugal, UK (Northern Ireland) **Study design:** Prospective cohort study **Number of centres:** 3 **Funding:** Educational grant from Pfizer Inc. **Competing interests:** Authors declared financial support or consultancies from Pfizer, Heidelberg Engineering, Zeiss Meditec, Novartis, Allergan, Zeiss, Alcon, Bayer, and THEA	**Index test:** fundus autofluorescence (FAF) acquired using scanning laser ophthalmoscopy (SLO): Spectralis HRA+OCT (Heidelberg Engineering, Heidelberg, Germany). Excitation not stated; barrier filter not stated. Settings: Field of view 30° centred on the macula; automatic image brightness (also called gain); high-speed mode; movie duration 30 seconds; average of 15 frames (Spectralis mean function); and tomography settings 7mm for Z-scan images **Reference standard:** (1) Reference standard relevant to the current review: Colour fundus photography (CFP): Stereopair colour images acquired using a Topcon 50X fundus camera. No further details given. (2) Reference standard according to the primary study: Presence or absence of RPD on >1 of 5 modalities: CFP, red-free photography (RF), Infrared photography (IR), fundus auto-fluorescence (FAF), and optical coherence tomography (OCT) CFP: details as above IR: acquired using same equipment as index test and same settings	**Number of participants:** 105 **Number of eyes:** 105 **Sample attrition/dropout:** Not reported, but appears to have excluded 12 eyes with poor image quality (n=93 after exclusion) **Selection of participants:** Patients attending retina clinics at each study site who had a diagnosis of neovascular AMD in 1 eye were approached and invited to take part. Neovascular AMD not defined in the publication **Inclusion criteria for study entry:** Men and women older than 50 years with a confirmed diagnosis of neovascular AMD in 1 eye; study eye (fellow eye) free of any features of late AMD (i.e., no neovascularization or geographic atrophy) with a visual acuity of 20/40 or better; sufficiently clear ocular media and adequate pupillary dilatation to permit good-quality fundus imaging of the study eye; and	**Primary outcome of study:** Presence of RPD **Other relevant outcomes:** Between-grader repeatability (κ statistics) for each imaging method **Diagnostic threshold:** Definitions of RPD: FAF: "clusters of ill-defined hypo-autofluorescent lesions interspersed against a background of mildly increased AF occurring in a regular and well-defined array." CFP: yellow interlacing networks ranging from 125 to 250 μm in width or lesions that occurred in regular, well-defined domains. IR and RF: "clusters of ill-defined hypo-reflective lesions interspersed against a background of mild hyper-reflectance." OCT: discrete accumulations of material anterior to the RPE often occurring as sharp

	RF: acquired using same equipment as index test and same settings OCT acquired using same equipment as index test. Centred on the macula, using evenly spaced lines in the scan area: 30° (horizontal) x 15° (vertical) area; number of sections set to 37; mean function used with 5 scans per line; high-speed acquisition mode Note: SD-OCT implied but not stated	willing and able to comply with scheduled visits, laboratory tests, and other trial procedures. **Exclusion criteria for study entry:** Evidence of a neovascular lesion on FA in the study eye; any other feature of neovascular AMD (eg. subretinal or intraretinal fibrosis within the macular region, RPE tear); significant media opacities, cataracts, lens opacification requiring cataract surgery within 2 year follow-up; other retinal disease eg. pathologic myopia (spherical equivalent of -8 diopters or more or axial length of 25 mm or more), ocular istoplasmosis syndrome, angioid streaks, choroidal rupture, multifocal choroiditis; ocular progressive disease, eg. glaucoma or diabetic retinopathy in the study eye; medical condition that would interfere with the patient's ability to complete the trial; concurrent enrolment in any other observational or interventional clinical study; treatment with an ocular or systemic investigational agent in the past 60 days for medical	peaks visible within the layers corresponding to the outer regions of the photoreceptors **Recruitment dates:** Not reported

		condition; or known serious allergies to the dye used in FA or ICGA.	

Participant characteristics	
Sex, m:f (%male)	53:52 (50)
Age, years, mean (SD)	75.6 (7.5), range 52-93
Visual acuity in patients with vs. without drusen 1) Distance visual acuity (letters), mean (SD) 2) Near visual acuity (logarithm of the minimum angle of resolution), mean (SD) 3) Low luminescence visual acuity (SKILL score), mean (SD)	 1) 83 (6) vs. 81 (6) 2) 0.3 (0.1) vs. 0.2 (0.1) 3) 38 (12) vs. 33 (9)

Results – (1) FAF versus Spectralis OCT

Calculations are based on numbers of eyes (single eyes of 93 subjects)	**Population with disease on Spectral OCT**		**Population without disease on Spectral OCT**		**Total**
FAF imaging positive	29	a	9	c	38
FAF imaging negative	4	b	48	d	52
Total	33		57		90

Diagnosis		**95% CI**
Clinical sensitivity a / (a + b)	87.88 %	71.78 to 96.52
Clinical specificity d / (c + d)	84.21 %	72.13 to 92.30
PPV a / (a + c)	76.32 %	59.75 to 88.53
NPV d / (b + d)	92.31 %	81.44 to 97.82
Positive likelihood ratio [sensitivity/(100-specificity)]	5.57	3.02 to 10.27
Negative likelihood ratio [(100-sensitivity)/specificity]	0.14	0.06 to 0.36
Diagnostic odds ratio (a x d)/(b x c)* *0.5 added to each number to avoid division by zero	38.67	10.92 to 136.97
Interpretability and acceptability of test – see table below		

Results – (2) FAF versus CFP

Calculations are based on numbers of eyes (single eyes of 93 subjects)	**Population with disease on CFP**		**Population without disease on CFP**		**Total**
FAF imaging positive	15	a	26	c	41
FAF imaging negative	0	b	52	d	52
Total	15		78		93

Diagnosis		**95% CI**
Clinical sensitivity a / (a + b)	100.00 %	78.03 to 100.00
Clinical specificity d / (c + d)	66.67 %	55.08 to 76.94
PPV a / (a + c)	36.59 %	22.13 to 53.06
NPV d / (b + d)	100.00 %	93.08 to 100.00
Positive likelihood ratio [sensitivity/(100-specificity)]	3.00	2.19 to 4.11
Negative likelihood ratio [(100-sensitivity)/specificity]	0.00	Not calculable
Diagnostic odds ratio (a x d)/(b x c)* *0.5 added to each number to avoid division by zero	61.42	3.54 to 1066.71

Comments: CFP is the usual method for diagnosing RPD but was not the reference standard in the primary study. Diagnostic outcomes for this comparison were not reported in the paper but have been calculated by reviewers from data in Table 4 in the paper.

Interpretability and acceptability of test – see table below

Results – (3) FAF versus >1 imaging modality

Calculations are based on numbers of eyes (single eyes of 93 subjects)	**Population with disease on >1 imaging modality**		**Population without disease on >1 imaging modality**		**Total**
FAF imaging positive	41	a	0	c	41
FAF imaging negative	2	b	50	d	52
Total	43		50		93

Diagnosis		**95% CI**
Clinical sensitivity a / (a + b)	95.35 %	84.16 to 99.30
Clinical specificity d / (c + d)	100.00 %	92.82 to 100.00
PPV a / (a + c)	100.00 %	91.31 to 100.00
NPV d / (b + d)	96.15 %	86.76 to 99.42
Positive likelihood ratio [sensitivity/(100-specificity)]	Not calculable	
Negative likelihood ratio [(100-sensitivity)/specificity]	0.05	0.01 to 0.18
Diagnostic odds ratio (a x d)/(b x c)	1676.60	78.30 to 35903.35

Comments: The diagnostic odds ratio was not reported in the paper. The calculation of specificity differs as the paper reported specificity to be 98%.

Interpretability and acceptability of test	
Numbers excluded from analysis due to poor image quality	Not reported. Appears to have excluded 12 eyes that were ungradable for RPD: Instead of 105 eyes, results are presented for 93 eyes comparing FAF with >1 imaging modality; 93 eyes comparing FAF with fundus photography; and 90 eyes comparing FAF with OCT. However, the numbers that were ungradable on each imaging modality are not specified.
Inter-observer agreement (only for the UK [Belfast] site, n=35), kappa statistics	Colour photography, 0.72 (P<0.001); IR, 0.87 (P<0.001); RF, 0.53 (P = 0.002); FAF, 0.94 (P<0.001); OCT, 0.86 (P<0.001); ICGA, 0.93 (P<0.001); RPD on 1 or more imaging method, 1.0.
Intra-observer agreement	Not reported
Test acceptability (patients / clinicians)	Not reported
Adverse events	Not reported

AF: autofluorescence; AMD: age-related macular degeneration; CFP: colour fundus photography; cSLO: confocal scanning laser ophthalmoscopy; FA: fluorescein angiography; FAF: fundus autofluorescence; ICGA: indocyanine green angiography; IR: infrared photography; NPV: negative predictive value; OCT: optical coherence tomography; PPV: positive predictive value; RF: red-free photography; RPD: reticular pseudodrusen; RPE: retinal pigment epithelium; SD-OCT: spectral-domain optical coherence tomography

Hogg and colleagues[96] critical appraisal criteria (based on Reitsma et al.[81] adaptation of the QUADAS Tool[93])

	Item	Description	Judgement
1	Was the spectrum of patients representative of the patients who will receive the test in practice?	Study is prospective, and involved consecutive patients. Patients had neovascular AMD in only one eye, and no signs of AMD or other eye conditions in the other eye, so this is an atypical case-mix	No
2	Is the reference standard likely to classify the target condition correctly?	Study reference standard is positive result on ≥1 of 5 modalities (CFP, RFP, IRP, FAF, OCT), but CFP is the standard approach in clinical practice, with SD-OCT and FA also useful for detecting RPD	Unclear
3	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Not reported	Unclear
4	Did the whole sample or a random selection of the sample receive verification using the intended reference standard?	Multiple imaging methods were used, but it is unclear whether all participants received all tests but data was excluded, or whether some participants did not receive all tests. Attrition is not reported	Unclear
5	Did patients receive the same reference standard irrespective of the index test result?	No, the combination of the ≥1 test modalities making up the reference standard varied between patients.	No
6	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	The index test was one of the tests contributing to a diagnosis	No
7	Were the reference standard results interpreted without knowledge of the results of the index test?	Not reported	Unclear
8	Were the index test results interpreted without knowledge of the results of the reference standard?	Not reported	Unclear
9	Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	Not reported	Unclear
10	Were uninterpretable or intermediate test results reported?	12 subjects were ungradable for RPD and these appear to have been excluded from analysis – but the number of ungradable images on the index test is not reported	No
11	Were withdrawals from the study explained?	No	No

Study 4 of 8 – McBain and colleagues

Reference and design	Diagnostic tests	Participants	Outcome measures
Condition being diagnosed/detected: Cystoid macular oedema (CMO) **First author:** McBain[100] **Publication Year:** 2008 **Country:** UK **Study design:** Retrospective, consecutive, observational case series **Number of centres:** 1 **Funding:** Not stated **Competing interests:** Stated none	**Index test:** FAF imaging using cSLO. This was obtained using Heidelberg retina angiograph which consisted of a solid-state argon blue excitation laser (488nm) and barrier filter (500nm). 30 degree field-of-view mode was used for the images. Sequential images were obtained and 20 frames were selected and averaged to assess the distribution of FAF. **Reference standard:** Fluorescein angiography (FA) Digital stereo images obtained using Topcon-Imagenet system **Comparator:** None Time period between tests: within 2 weeks of each other; there was a minimum gap of 4 days washout if FAF was obtained following FA	**Number of participants:** 34 **Number of eyes**: 34 **Sample attrition/dropout:** 106 consecutive patients with clinically suspected CMO had FAF imaging, of which 34 patients were eligible for inclusion and 62 were excluded.* **Selection of participants:** Consecutive patients with clinically suspected CMO were selected from FAF imaging database of the Ophthalmology Department. **Inclusion criteria for study entry:** CMO secondary to cataract extraction, inherited retinopathies, inflammatory eye disease or idiopathic cases, where both FAF and FA were obtained to confirm diagnosis. One eye per person included, left eye chosen in bilateral cases. Patients were eligible if FAF was performed within 2 weeks of FA **Exclusion criteria for study entry:** No additional criteria cited.	**Primary outcome of study:** Diagnostic accuracy (sensitivity and specificity) **Other relevant outcomes:** Interpretability and acceptability of test; adverse events **Diagnostic threshold:** FAF: CMO considered present whenever there were round or oval areas of fundus autofluorescence at the fovea with a fundus auto-fluorescence signal similar to back-ground levels. FAF signal is usually reduced at the fovea compared with background, due to blockage of the signal by the luteal pigment. FA: CMO was considered present whenever leakage of fluorescein dye was observed in a petaloid pattern around the fovea in the late phase of the angiogram (recirculation phase or later) **Recruitment dates:** February 2004 - May 2007*

*The numbers do not add up to 106 but 96. There is a discrepancy in reporting the total numbers in the abstract (which reports 96) vs the text (reporting 106). There is also a discrepancy in recruitment dates in the abstract (reported as between Aug 2004 and June 2006) vs the text (Feb 2004 and May 2007). It appears that the main text has been updated but the abstract has not, and that 10 exclusions have not been accounted for.

Participant characteristics	
Sex, m:f (%male:female)	20:14 (59)
Age, years, mean (SD)	59 (range 17-89)
CMO secondary to inflammatory disease, n (%)	17/34 (50)
CMO following cataract surgery, n (%)	11/34 (32)
CMO associated with inherited retinal dystrophies, n (%)	3/34 (9)
CMO idiopathic, n (%)	4/34 (12)

Results – FAF versus FA

Calculations are based on numbers of eyes (= number of patients as only one eye per patient was included)	**Population with disease on FA**		**Population without disease on FA**		**Total**
FAF imaging positive	17	a	4	c	21
FAF imaging negative	4	b	9	d	13
Total	21		13		34

Diagnosis		**95% CI**
Clinical sensitivity a / (a + b)	80.95	58.08 to 94.44
Clinical specificity d / (c + d)	69.23	38.61 to 90.72
PPV a / (a + c)	80.95	58.08 to 94.44
NPV d / (b + d)	69.23	38.61 to 90.72
Positive likelihood ratio [sensitivity/(100-specicifcity)]	2.63	1.13 to 6.10
Negative likelihood ratio [(100-sensitivity)/specificity]	0.28	0.11 to 0.71
Diagnostic odds ratio (a x d)/(b x c)	9.56	1.92 to 47.57

Comments: Calculations agree with values reported in paper except for values for PPV and NPV, which are switched in the paper.

Interpretability and acceptability of test	
Poor FAF images related to media opacities (cataract), n (%)	9/96 (9.4%)
Inter-observer agreement	Not reported
Intra-observer agreement	Not reported
Test acceptability (patients / clinicians)	Not reported
Adverse events	No side effects were observed related to FAF or AF images during the study period.
The percentage has been calculated by reviewers using the denominator 96 rather than 106, as the reasons for 10 exclusions appear to have been omitted from the paper (see above).	

CMO: cystoid macular oedema; FA: fluorescein angiography; FAF: fundus autofluorescence; NPV: negative predictive value; PPV: positive predictive value

McBain and colleagues[100] critical appraisal criteria (based on Reitsma et al.[81] adaptation of the QUADAS Tool[93])

	Item	**Description**	**Judgement**
1	Was the spectrum of patients representative of the patients who will receive the test in practice?	Study is retrospective, but involved consecutive patients. Patients had CMO secondary to specific conditions, so this may be an atypical case-mix	No
2	Is the reference standard likely to classify the target condition correctly?	FA is described as used routinely for diagnosis of CMO	Yes
3	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Within two weeks of each other, with a minimum gap of 4 days if FAF followed AF.	Yes
4	Did the whole sample or a random selection of the sample receive verification using the intended reference standard?	All included patients received both tests.	Yes
5	Did patients receive the same reference standard irrespective of the index test result?	FA was used in all analysed patients	Yes
6	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	FAF was not part of reference standard	Yes
7	Were the reference standard results interpreted without knowledge of the results of the index test?	Evaluation of results was done in a masked fashion by a single observer; images evaluated independently from one another.	Unclear
8	Were the index test results interpreted without knowledge of the results of the reference standard?	As above	Unclear
9	Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	States no information was provided to the observer with regards to the patient.	Unclear
10	Were uninterpretable or intermediate test results reported?	Reports number of patients excluded due to poor FAF images (9/96)	Yes
11	Were withdrawals from the study explained?	62/96 patients excluded: 23 no AF, 14 had more than two weeks	Yes

		between tests, 2 had FAF less than four days after FA, 9 had poor AF images related to media opacities (cataract), 14 had CMO related to other diseases, e.g. precious branch vein occlusion, diabetic retinopathy or AMD. However, there is a discrepancy between the recruitment dates and numbers recruited between the abstract and main text, which suggests that there are 10/106 exclusions which are not accounted for.	

Study 5 of 8 – Smith and colleagues

Reference and design	Diagnostic tests	Participants	Outcome measures
Condition being diagnosed / detected: Reticular pseudodrusen (RPD) in age-related macular degeneration (AMD) **First author:** Smith[101] **Publication year:** 2006 **Country:** UK and USA **Study design:** Retrospective case series (2 distinct case series combined) **Number of centres:** Not reported **Funding:** New York Community Trust (lead author) and unrestricted funds from Research to Prevent Blindness **Competing interests:** Stated none	**Index test:** Fundus autofluorescence (FAF) imaging. No details of method reported; introduction suggests probably confocal scanning laser ophthalmoscopy (cSLO) **Reference standard:** Colour fundus photography (CFP). Colour photographs were studied both in their original state and as highly contrast-enhanced versions, to facilitate RPD identification. No further details of method reported. **Comparator:** None reported	**Number of participants:** 138 **Number of eyes:** 221 (166 eyes of 83 patients with early AMD or GA, without evidence of choroidal neovascularisation (CNV)) and 55 unaffected eyes of 55 patients with unilateral CNV) **Sample attrition/dropout:** None (retrospective database selection) **Selection of participants:** From two databases: an AMD study database at the UK Institute of Ophthalmology; and a database of patients imaged at Columbia Eye University, USA. Not reported whether patients were selected consecutively. **Inclusion criteria for study entry:** Not explicitly reported. Stated only that the eyes had either: bilateral soft drusen ± pigment abnormalities, but no evidence of CNV; or they had unilateral CNV. **Exclusion criteria for study entry:** Eyes that did not receive both FAF imaging and colour fundus photography.	**Primary outcome of study:** The fraction and relative probability of focally increased autofluorescence corresponding spatially with drusen and pigment as identified by fundus colour photography; and the presence or absence of reticular FAF and RPD **Other relevant outcomes:** None reported **Diagnostic threshold:** FAF: Reticular pattern of autofluorescence (hypofluorescent lesions) CFP: Image segmentation method reported, but morphological criteria for diagnosing RPD on CFP not reported **Recruitment dates:** Not reported

Participant characteristics	
Sex, m:f (%male)	Not reported
Age, years, mean (SD)	Not reported
Other key characteristics	None reported

Results – FAF imaging

Calculations are based on numbers of eyes (both eyes of 83 subjects and single eyes of 55 subjects)	**Population with RPD on colour fundus photography**	**Population without RPD on colour fundus photography**	**Total**
FAF imaging positive	28 a	4 c	32
FAF imaging negative	2 b	187 d	189
Total	30	191	221

Diagnosis		**95% CI**
Clinical sensitivity a / (a + b)	93.33%	77.89 to 98.99
Clinical specificity d / (c + d)	97.91%	94.72 to 99.41
PPV a / (a + c)	87.50%	70.99 to 96.41
NPV d / (b + d)	98.94%	96.22 to 99.84
Positive likelihood ratio [sensitivity/(100-specificity)]	44.57	16.82 to 118.08
Negative likelihood ratio [(100-sensitivity)/specificity]	0.07	0.02 to 0.26
Diagnostic odds ratio (a x d)/(b x c)	654.50	114.50 to 3741.07

Comments: Sensitivity calculation agrees with statement in the paper that "AF imaging was over 90% sensitive" (no other diagnostic results were reported in the paper).
Sensitivity and specificity are also calculable for a sub-group of patients based on unaffected fellow eyes of those with unilateral CNV ("CNV-R" group). However, subgroup is defined by auto-fluorescence pattern (reticular AF and / or RPD) and does not include all patients with unilateral CNV. Therefore data have not been extracted here.

Interpretability and acceptability of test	
Numbers excluded from analysis due to poor image quality	None. However, reported that for one patient with RPD only this was "perhaps due to marginal scan quality" and another patient had "bilateral RPD and an AF image in the left eye that could not be graded for reticular AF". Unclear whether these were the only poor-quality images present.
Inter-observer agreement	Not reported
Intra-observer agreement	Not reported
Test acceptability (patients / clinicians)	Not reported
Adverse events	Not reported

AF: autofluorescence; AMD: age-related macular degeneration; CFP: colour fundus photography; CNV: choroidal neovascularisation; cSLO: scanning laser ophthalmoscopy; FAF: fundus autofluorescence; GA: geographic atrophy; NPV: negative predictive value; PPV: positive predictive value; RPD: reticular pseudodrusen

Smith and colleagues[101] critical appraisal criteria (based on Reitsma et al.[81] adaptation of the QUADAS Tool[93])

	Item	Description	Judgement
1	Was the spectrum of patients representative of the patients who will receive the test in practice?	Study is retrospective but unclear if it involved consecutive patients. Two separate cohorts of eyes were combined: patients from the UK without evidence of CNV and unaffected eyes of patients from the USA who had unilateral CNV, so is an atypical case mix	No
2	Is the reference standard likely to classify the target condition correctly?	Although not stated explicitly, CFP is a standard approach for detecting RPD	Yes
3	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Not reported	Unclear
4	Did the whole sample or a random selection of the sample receive verification using the intended reference standard?	Whole sample – appears to have been purposively selected to ensure patients had received both tests	Yes
5	Did patients receive the same reference standard irrespective of the index test result?	Yes	Yes
6	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes	Yes
7	Were the reference standard results interpreted without knowledge of the results of the index test?	Not reported	Unclear
8	Were the index test results interpreted without knowledge of the results of the reference standard?	Not reported	Unclear
9	Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	Not reported	Unclear
10	Were uninterpretable or intermediate test results reported?	No. Unclear whether patient selection purposively excluded those with uninterpretable or indeterminate test results. Some poor-quality images may have influenced results classification (see 'Interpretability and acceptability of test' above)	No
11	Were withdrawals from the study explained?	No withdrawals	Not applicable

Study 6 of 8 – Ueda-Arakawa and colleagues

Reference and design	Diagnostic tests	Participants	Outcome measures
Condition being diagnosed / detected: Reticular pseudodrusen (RPD) in age-related macular degeneration (AMD) **First author:** Ueda-Arakawa[97] **Publication year:** 2013 **Country:** Japan **Study design:** Retrospective case series **Number of centres:** One **Funding:** Not stated **Competing interests:** Stated none	**Index test:** (1) Fundus auto-fluorescence (FAF): acquired using confocal scanning laser ophthal-moscope (cSLO) (Spectralis HRA+OCT; Heidelberg Engineering, Heidelberg, Germany). Excitation 488nm; barrier filter beginning at 500nm. (2) Near-infrared fundus autofluorescence (NIR-FAF): acquired with cSLO (same equipment as FAF), in the indocyanine green angiography mode (excitation: 790nm; detection 800nm). Each FAF or NIR-FAF image was compiled from an average of 15 to 20 scans by the cSLO software. **Reference standard:** At least 2 of 7 imaging modalities (in any combination) positive for RPD: (1) Contrast-enhanced colour fundus photo-graphy (CFP): 30°–40° field acquired digitally using Topcon TRC NW6S non-mydriatic retinal camera (Topcon, Tokyo, Japan). Blue channel examined using Image J software (Nat-ional Institutes of Health, Bethesda, MD, USA). (NB: paper notes that this has been the traditional method for detecting RPD).	**Number of participants:** 114 **Number of eyes:** 220 **Sample attrition/dropout:** 8/228 eyes excluded, due to phthisis bulbi (n=2) or poor image quality in ≥3 imaging modalities (n=6). Further excluded due to poor image quality: FAF imaging: 3/220; NIR-FAF imaging: 84/220. **Selection of participants:** Consecutive patients with AMD who first visited ophthal-mology department during recruitment dates **Inclusion criteria for study entry:** Early AMD, neovascular AMD or geographic atrophy in at least one eye. Early AMD defined as presence of soft drusen (≥ 63 μm) or areas of hyper- or hypopigmentation in the RPE. Geographic atropy defined on colour fundus photography as a sharply delineated area (≥ 175 μm) ie hypopigmentation, depigmentation or apparent absence of RPE in which choroidal vessels were clearly visible.	**Primary outcome of study:** Not stated. Paper focuses on sensitivity and specificity of each imaging modality at detecting RPD. **Other relevant outcomes:** Inter-grader agreement rates for detecting RPD in each imaging modality. **Diagnostic threshold:** RPD diagnosed if reticular patterns showed on at least two of the follow-ing: blue-channel CFP, IR, FAF, NIR-FAF, CBR, IA, or SD-OCT. Characterisation of reticular lesions: FAF and NIR-FAF: A group of ill-defined, hypo-fluorescent lesions against a back-ground of mildly elevated AF. Blue-channel CFP and CBR: light interlacing networks 125-250μm wide. IR: groups of hypo-reflectant lesions against a back-ground of mild hyper-reflectance with anomalous characteristics.

	(2) Infrared reflectance (IR): acquired using cSLO (same equipment as the index test). Light stimulus 820nm. (3) FAF imaging (i.e., an index test – see above). (4) NIR-FAF imaging (i.e., an index test – see above). (5) Confocal blue reflectance (CBR): acquired with cSLO (same equipment as the index test). Light stimulus 488nm; field of view 30° x 30°, centred on the macula. (6) Late-phase indocyanine green angiography (ICGA): acquired with cSLO (same equipment as the index test). Excitation: 790nm; detection 800nm. (7) Spectral domain optical coherence tomography (SD-OCT): conducted using Spectralis HRA+OCT (Heidelberg Engineering). Horizontal and vertical line scans through the fovea centre obtained at a 30° angle, followed by serial horizontal scans with an examination field size ranging from 30° x 10° to 30° x 25. At each location of interest on the retina, 50 images were averaged to reduce speckle noise.	Neovascular AMD defined as neovascularisation detected using FA or indocyanine angiography. Only images of eligible quality were analysed and only eyes with eligible image quality in at least five imaging modalities were included. **Exclusion criteria for study entry:** People aged < 50 years, eyes with high myopia, eyes with other macular abnormalities.	ICGA: A distinctive grouping of hypo-fluorescent dots present in the late angiogram phases. SD-OCT: ≥5 hyper-reflective mounds or triangular lesions above the RPE in ≥1 B-scan. **Recruitment dates:** January 2010 – November 2010

Participant characteristics	
Sex, m:f (%male)	79:35 (69)
Age, years, mean (SD)	73.8 (9.4), range 52-92
Visual acuity (logarithm of the minimum angle of resolution), mean (SD)	0.396 (0.512), range 0.01-1.5

Results

FAF versus ≥2 (of 7) imaging modalities

Calculations are based on numbers of eyes, including both eyes of each subject	**Population with disease on ≥2 imaging modalities**	**Population without disease on ≥2 imaging modalities**	**Total**
FAF imaging positive	32 a	9 c	41
FAF imaging negative	5 b	171 d	176
Total	37	180	217

Diagnosis		**95% CI**
Clinical sensitivity a / (a + b)	86.49%	71.21 to 95.41
Clinical specificity d / (c + d)	95.00%	90.72 to 97.68
PPV a / (a + c)	78.05%	62.38 to 89.42
NPV d / (b + d)	97.16%	93.49 to 99.06
Positive likelihood ratio [sensitivity/(100-specificity)]	17.30	9.04 to 33.11
Negative likelihood ratio [(100-sensitivity)/specificity]	0.14	0.06 to 0.32
Diagnostic odds ratio (a x d)/(b x c)	121.60	38.25 to 386.57

Comments: Calculations agree with values reported in paper (except diagnostic odds ratio not reported). Paper also reports (in Supplementary Table 2) that the sensitivity of FAF imaging is 86.5% when the field size is limited to the same imaging area as SD-OCT, i.e. 30^{o} x 10^{o} – but sample sizes (n/N) for this calculation (32/37) are not explained.

Note that CFP is the test usually considered as a reference standard for diagnosing RPD. Although diagnostic outcomes for a comparison of FAF versus CFP are given in supplementary Table 1 of the paper, these relate only to a subset of 37 eyes that had a reticular pattern on ≥2 imaging modalities, therefore these data have not been extracted.

Interpretability and acceptability of test	
Number of eyes excluded from analysis due to poor image quality	3/220 (1.4%)
Inter-observer agreement (grading reticular pattern)	89.3%; kappa = 0.700
Intra-observer agreement	Not reported
Test acceptability (patients / clinicians)	Not reported
Adverse events	Not reported

NIR-FAF versus ≥2 (of 7) imaging modalities

Calculations are based on numbers of eyes, including both eyes of each subject	**Population with disease on ≥2 imaging modalities**	**Population without disease on ≥2 imaging modalities**	**Total**
NIR-FAF imaging positive	9 a	5 c	14
NIR-FAF imaging negative	19 b	103 d	122
Total	28	108	136

Diagnosis		**95% CI**
Clinical sensitivity a / (a + b)	32.14%	15.91% to 52.35%
Clinical specificity d / (c + d)	95.37%	89.52% to 98.46%
PPV a / (a + c)	64.29%	35.18% to 87.11%
NPV d / (b + d)	84.43%	76.75% to 90.35%
Positive likelihood ratio [sensitivity/(100-specificity)]	6.94	2.53 to 19.08
Negative likelihood ratio [(100-sensitivity)/specificity]	0.71	0.55 to 0.92
Diagnostic odds ratio (a x d)/(b x c)	9.76	2.95 to 32.33
Comments: Calculations agree with values reported in paper (except diagnostic odds ratio not reported). Paper also reports (in Supplementary Table 2) that the sensitivity of NIR-FAF imaging is 28.6% when the field size is limited to the same imaging area as SD-OCT, i.e. 30^o x 10^o – but sample sizes (n/N) for this calculation (8/28) are not explained.		
Interpretability and acceptability of test		
Number of eyes excluded from analysis due to poor image quality	64/220 (29%)	
Inter-observer agreement (grading reticular pattern)	84.2%; kappa = 0.563	
Intra-observer agreement	Not reported	
Test acceptability (patients / clinicians)	Not reported	
Adverse events	Not reported	

Number of eyes with good image quality – results for all imaging tests	
FAF	217/220 (99%)
NIR-FAF	136/220 (62%)
Blue-channel CFP	220/220 (100%)
IRR	220/220 (100%)
ICGA	220/220 (100%)
SD-OCT	220/220 (100%)
CBR	204/220 (93%)

AMD: age-related macular degeneration; CBR: confocal blue reflectance; CFP: colour fundus photography; CNV: choroidal neovascularisation; cSLO: confocal scanning laser ophthalmoscopy; FAF: fundus autofluorescence; GA: geographic atrophy; ICGA: indocyanine green angiography; IRR: infrared reflectance; NIR-FAF: near-infrared fundus autofluorescence; NPV: negative predictive value; PPV: positive predictive value; RPD: reticular pseudodrusen; RPE: retinal pigment epithelium; SD-OCT: spectral-domain optical coherence tomography

Ueda-Arakawa and colleagues[97] critical appraisal criteria (based on Reitsma et al.[81] adaptation of the QUADAS Tool[93])

	Item	Description	Judgement
1	Was the spectrum of patients representative of the patients who will receive the test in practice?	Study is retrospective but involved consecutive patients. Patients are Japanese, with newly diagnosed AMD, and patients with comorbidities excluded (including, among others, certain types of CNV and central serous chorioretinopathy), so is an atypical case-mix	No
2	Is the reference standard likely to classify the target condition correctly?	Paper describes CFP as traditional test for detecting RPD. Study reference standard is positive result on ≥2 of 7 modalities, but individual eyes were diagnosed using different combinations of modalities	Unclear
3	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Not reported	Unclear
4	Did the whole sample or a random selection of the sample receive verification using the intended reference standard?	Multiple imaging methods used in all patients but the diagnostic tests differed between patients	Yes
5	Did patients receive the same reference standard irrespective of the index test result?	No, the combination of the ≥2 test modalities making up the reference standard varied between patients	No
6	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	The index test(s) could have been one (or both) of the two tests contributing to a diagnosis	No
7	Were the reference standard results interpreted without knowledge of the results of the index test?	Not reported	Unclear
8	Were the index test results interpreted without knowledge of the results of the reference standard?	Not reported	Unclear
9	Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	Not reported	Unclear
10	Were uninterpretable or intermediate test results reported?	Stated that 3/220 eyes for FAF (1%), 84/220 eyes for NIR-FAF (38%) and 16 eyes for CBR (7%) did not have good quality images and were excluded	Yes
11	Were withdrawals from the study explained?	Yes – phthisis bulbi or poor image quality	Yes

Study 7 of 8 – Vujosevic and colleagues

Reference and design	Diagnostic tests	Participants	Outcome measures
Condition being diagnosed / detected: diabetic macular oedema (DMO) **First author:** Vujosevic[98] **Publication year:** 2012 **Country:** Italy **Study design:** Cross-sectional study. Probably prospective (not explicitly stated but all patients provided consent) **Number of centres:** Not reported (>1 clinic implied) **Funding:** None received **Competing interests:** No information provided	**Index test:** Fundus autofluorescence (FAF) acquired with confocal scanning laser ophthalmoscopy (cSLO) (Heidelberg Retinal Angiograph, HRA 2; Heidelberg Engineering, Heidelberg, Germany). Solid-pumped laser; excitation 488nm; emission detected above 500nm using barrier filter. FAF signal amplified by calculating a mean of 15 aligned images after correction of eye movements using image analysis software. **Reference standard:** Retromode scanning laser ophthalmoscopy (RM-SLO): images taken with F-10 SLO (Nidek Co, Gamagori, Japan), which uses 4 wavelengths: blue, green, red and infrared. Infrared laser was set at 790nm. F-10 contains 8 apertures (five confocal and 3 with a central stop) and five stops. To obtain a RM-SLO image, a central stop and a laterally oriented oval-shaped opening was used, from both right and left sides. **Comparators:** Time domain OCT (TD-OCT) Fluorescein angiography (FA)	**Number of participants:** 137 **Number of eyes:** 263 **Sample attrition/dropout:** Not reported explicitly but none evident **Selection of participants:** Recruited consecutively from unspecified diabetic retinopathy (DR) clinics **Inclusion criteria for study entry:** Type 1 or 2 diabetes mellitus; any stage of untreated or treated DR; and having TD-OCT, FAF, FA and RM-SLO performed on the same day **Exclusion criteria for study entry:** Significant media opacities	**Primary outcome of study:** Inter-method agreement in identifying different patterns of DMO **Other relevant outcomes:** Sensitivity and specificity for detecting DMO **Diagnostic threshold:** Identification of DMO: FAF: Not reported (stated only that images were graded for different foveal patterns [normal, single spot increased and multiple spots increased] and presence⁄absence of decreased⁄ increased auto-fluorescence in the macula) TD-OCT: central retinal thickness > 230 μm (measured in the central foveal zone) FA: Not reported (stated only that late-phase FA images of the macula were graded for the presence of fluorescein leakage and pattern of leakage [cystoid and non-cystoid]). RM-SLO: Not reported (stated only that images were graded for presence/absence of

			DMO) For all methods, 2 masked retinal specialists trained in imaging grading independently graded all images on a 17-inch monitor dedicated to DR screening. In case of disagreement, a 3rd specialist adjudicated. **Recruitment dates:** March to August 2009

Participant characteristics	
Sex, m:f (%male)	87:50 (64)
Age, years, mean (SD)	Type I diabetes: 48.8 (11.5), range 28-64 Type II diabetes: 66.6 (8.1), range 41-85 Overall : Not reported (numbers with diabetes do not account for all patients – see below)
With Type I or II diabetes, N(%)	Type I: 12 (8.8) [reported as 10.1% in the paper] Type II: 107 (78.1) [reported as 89.9% in the paper] Not reported: 18 (13.1)
Duration of diabetes, years, mean (SD)	Type I: 28.8 (11.9), range 5-51 Type II: 15.4 (8.8), range 1-39
Central macular thickness, mean (SD), µm	323.4 (125.2), range 154.0-884.0

Results - FAF versus RM-SLO

Calculations based on numbers of eyes (both eyes of 126 subjects and single eyes of 11 subjects)	**Population with DMO on RM-SLO**		**Population without DMO on RM-SLO**		**Total**
FAF imaging positive	195	a	8	c	203
FAF imaging negative	16	b	44	d	60
Total	211		52		263

Diagnosis		**95% CI**
Clinical sensitivity a / (a + b)	92.42%	87.98 to 95.60
Clinical specificity d / (c + d)	84.62%	71.91 to 93.10
PPV a / (a + c)	96.06%	92.38 to 98.28
NPV d / (b + d)	73.33%	60.34 to 83.92
Positive likelihood ratio [sensitivity/(100-specificity)]	6.01	3.17 to 11.38
Negative likelihood ratio [(100-sensitivity)/specificity]	0.09	0.06 to 0.15
Diagnostic odds ratio (a x d)/(b x c)	67.03	26.99 to 166.45

Comments: Data reported in the paper are for RM-SLO compared against a FAF reference; recalculated by reviewers to give sensitivity and specificity of FAF compared against a RM-SLO reference.

Paper states (in the Discussion) that OCT is the 'new gold standard' for diagnosing DMO. However, a diagnostic accuracy comparison of FAF versus TD-OCT is not possible from the reported data (only the diagnostic accuracy of RM-SLO versus TD-OCT, FA and FAF are reported and calculable – not extracted here).

Interpretability and acceptability of test	
Numbers excluded from analysis due to poor image quality	Not reported
Inter-observer agreement	Not reported
Intra-observer agreement	Not reported
Test acceptability (patients / clinicians)	Not reported
Adverse events	Not reported

cSLO: confocal scanning ophthalmoscopy; DMO: diabetic macular oedema; DR: diabetic retinopathy; FA: fluorescein angiography; FAF: fundus autofluorescence; NPV: negative predictive value; OCT: optical coherence tomography; PPV: positive predictive value; RM-SLO: retromode scanning laser ophthalmoscopy; TD-OCT: time domain optical coherence tomography

Vujosevic and colleagues[98] critical appraisal criteria (based on Reitsma et al.[81] adaptation of the QUADAS Tool[93])

	Item	Description	Judgement
1	Was the spectrum of patients representative of the patients who will receive the test in practice?	Unclear if study is prospective, but it involved consecutive patients. Patients had any stage of untreated or treated diabetic retinopathy so this may be a typical case-mix	Unclear
2	Is the reference standard likely to classify the target condition correctly?	Paper describes OCT (i.e. not RM-SLO) as the 'new gold standard' for classifying DMO	Unclear
3	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Same day	Yes
4	Did the whole sample or a random selection of the sample receive verification using the intended reference standard?	Whole sample	Yes
5	Did patients receive the same reference standard irrespective of the index test result?	Yes	Yes
6	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes	Yes
7	Were the reference standard results interpreted without knowledge of the results of the index test?	Order of tests not reported but stated that images were independently graded in a masked fashion	Yes
8	Were the index test results interpreted without knowledge of the results of the reference standard?	Order of tests not reported but stated that images were independently graded in a masked fashion	Yes
9	Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	Not reported	Unclear
10	Were uninterpretable or intermediate test results reported?	Not reported	No
11	Were withdrawals from the study explained?	No withdrawals evident	Not applicable

Study 8 of 8 – Waldstein and colleagues

Reference and design	Diagnostic tests	Participants	Outcome measures
Condition being diagnosed / detected: Diabetic macular oedema (DMO) **First author:** Waldstein[84] **Publication year:** 2012 **Country:** Not stated, appears to be UK **Study design:** Retrospective cross-sectional. **Number of centres:** One **Funding:** 2 authors received Marie Curie Intra European Fellowship; Worshipful Company of Barbers-Waitangi Foundation Fellowship; and funding from the University of Auckland. **Competing interests:** Stated none	**Index tests:** FAF imaging using cSLO (modified HRA classic, Heidelberg Engineering, Heidelberg, Germany) with an external source of a solid-state laser emitting at 488nm at a 30˚ field of view; and FAF imaging using cSLO with an argon-ion laser emitting at 514nm at a 30˚ field of view; mean of 16 images. **Reference standard:** SD-OCT (Heidelberg Engineering, software version 1.6.4.0). Each B-scan consisted of 512 A-scans and was averaged nine times using the ART mode. A 20˚ x 20˚ scan pattern using 25 sections with an inter-scan distance of 240µm was recorded. **Comparator:** Macular Pigment Optical Density (MPOD) imaging (sequential use of both 488nm and 514nm FAF allowed calculation of macular pigment optical density (MPOD) maps that topographically illustrate the relative distribution of macular pigment)	**Number of participants:**71 **Number of eyes:** 125 **Sample attrition/dropout:** Not reported explicitly; but all included eyes were analysed. **Selection of participants:** Patients who underwent OCT and two-wavelength FAF imaging in the diabetic retinopathy clinic of a university hospital were selected consecutively. **Inclusion criteria for study entry:** The presence of diabetic retinopathy with or without DMO; clear ocular media that allow recording of high-quality FAF images; availability of both two-wavelength FAF and OCT imaging within a 2-week period. **Exclusion criteria for study entry:** Presence of any ocular comorbidity affecting the macula, such as retinal vein occlusion or age-related macular degeneration.	**Primary outcome of study:** Comparison of sensitivity and specificity of FAF and MPOD for detection of DMO **Other relevant outcomes:** Inter-grader variability. (Cohen's kappa was used to estimate inter-grader agreement) **Diagnostic threshold:** Diagnosis of DMO was based on DMO visibility which was compared across the technologies using the following scoring system: -no DMO visible -DMO suspected -DMO clearly visible FAF: Relatively bright, single or multiple, round or oval areas that are mostly bordered by darker rims. OCT: Intraretinal cysts (no details given) **Recruitment dates:** Between May 2009 and November 2010

Participant characteristics	
Sex, m:f (%male)	46:25 (65%)
Age, years, mean (SD)	63 (15)

Results: FAF (488nm) versus OCT

Calculations are based on no. of eyes (single eyes from 17 subjects and both eyes from 54 subjects)	**Eyes with signs of DMO on SD-OCT**		**Eyes without signs of DMO on SD-OCT**		**Total**
FAF imaging positive	54	a	6	c	60
FAF imaging negative	13	b	52	d	65
Total	67		58		125

Diagnosis		**95% CI**
Clinical sensitivity a / (a + b)	80.60	69.11 to 89.24
Clinical specificity d / (c + d)	89.66	78.82 to 96.08
PPV a / (a + c)	90.00	79.48 to 96.22
NPV d / (b + d)	80.00	68.23 to 88.89
Positive likelihood ratio [sensitivity/(100-specicifcity)]	7.79	3.62 to 16.77
Negative likelihood ratio [(100-sensitivity)/specificity]	0.22	0.13 to 0.36
Diagnostic odds ratio (a x d)/(b x c)	36.00	12.73 to 101.81

Comments: Diagnostic values are calculated by the reviewer from the reported sensitivity and specificity. The calculations agree with the results reported in the paper.

Sensitivity and specificity are reported also for MPOD based on combining FAF488 nm and 514nm images. MPOD sensitivity and specificity were very similar to those of FAF 488nm alone (data not extracted here).

Results: FAF (514nm) versus OCT

Calculations are based on numbers of eyes (single eyes from 17 subjects and both eyes from 54 subjects)	**Eyes with signs of DMO on SD-OCT**		**Eyes without signs of DMO on SD-OCT**		**Total**
FAF imaging positive	37	a	3	c	40
FAF imaging negative	30	b	55	d	85
Total	67		58		125

Diagnosis		**95% CI**
Clinical sensitivity a / (a + b)	55.22	42.58 to 67.39
Clinical specificity d / (c + d)	94.83	85.60 to 98.86
PPV a / (a + c)	92.50	79.59 to 98.34
NPV d / (b + d)	64.71	53.59 to 74.77
Positive likelihood ratio [sensitivity/(100-specificity)]	10.68	3.47 to 32.82
Negative likelihood ratio [(100-sensitivity)/specificity]	0.47	0.36 to 0.62
Diagnostic odds ratio (a x d)/(b x c)	22.61	6.43 to 79.54

Comments: Diagnostic values are calculated by the reviewer from the reported sensitivity and specificity. The calculations agree with the results reported in the paper.

Distinct patterns of DMO on OCT (no. of eyes, %):	
Predominantly foveal intraretinal cysts	51 (76)
Predominantly extrafoveal intraretinal cysts	5 (7)
Diffuse, small intraretinal cysts	11 (16)
Sensitivity for detecting Foveal cysts compared to OCT imaging	
FAF (488nm)	90.0%
FAF (514nm)	20.0%
MPOD	96.0%
Sensitivity for detecting Extrafoveal or diffuse cysts compared to OCT imaging	
FAF (488nm)	60.8%
FAF (514nm)	70.0%
MPOD	45.5%
MPOD vs OCT	
Clinical sensitivity	80.6%
Clinical specificity	91.4%

FAF: Fundus Autofluorescence; NPV: negative predictive value; PPV: positive predictive value; MPOD: Macular Pigment Optical Density; SD-OCT: Spectral Domain Optical Coherence Tomography; cSLO: Confocal Scanning Laser Ophthalmoscope; DMO: Diabetic Macular Oedema

Interpretability and acceptability of test	
Numbers excluded from analysis due to poor image quality	Not reported
Intra-observer agreement	Not reported
Test acceptability (patients/clinicians)	Not reported
Adverse events	Not reported
Inter-observer agreement (Cohen's kappa)	
FAF (488nm)	0.84
FAF (514nm)	0.63
MPOD	0.79

Waldstein and colleagues[84] critical appraisal criteria (based on Reitsma et al.[81] adaptation of the QUADAS Tool[93])

	Item	Description	Judgement
1	Was the spectrum of patients representative of the patients who will receive the test in practice?	Study is retrospective, but involved consecutive patients. Patients had diabetic retinopathy with or without DMO, with no macular comorbidities, so this may be an atypical case-mix	No
2	Is the reference standard likely to classify the target condition correctly?	Paper describes OCT as clinical standard for the non-invasive diagnosis of DMO	Yes
3	Is the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests?	Both two-wavelength FAF and OCT imaging were available within a 2 week period.	Yes
4	Did the whole sample or a random selection of the sample receive verification using the intended reference standard?	Patients were required to have both tests for inclusion	Yes
5	Did patients receive the same reference standard irrespective of the index test result?	Yes	Yes
6	Was the reference standard independent of the index test (i.e. the index test did not form part of the reference standard)?	Yes	Yes
7	Were the reference standard results interpreted without knowledge of the results of the index test?	Same grader who assessed OCT scans was one of the FAF graders. No masking reported.	Unclear
8	Were the index test results interpreted without knowledge of the results of the reference standard?	All FAF and MPOD were evaluated by two independent graders who were blinded to the patient but not to the mode of imaging. No further details provided.	Unclear
9	Were the same clinical data available when test results were interpreted as would be available when the test is used in practice?	Not reported	Unclear
10	Were uninterpretable or intermediate test results reported?	No	No
11	Were withdrawals from the study explained?	All included eyes were analysed	Not applicable