H.1 Classification

H.1.1 Classification systems for age-related macular degeneration (AMD)

RQ6: What effective classification tool should be used to inform people with AMD?

Validation outcomes for existing classification systems of AMD

Agreement outcomes: Interobserver agreement

Studies	Classification System	Risk of bias	Inconsistency	Indirectness	Imprecision	Clinical population (n)	Effect	Quality
AREDS 17 (2006) Case-control study	AREDS 9-step severity scale	Serious[1]	Not applicable (N/A)	Not serious	Not serious	1225 eyes from the Age Related Eye Disease Study (AREDS)	Complete agreement: 63.4% of eyes, Agreement within 1 step: 86.6%, Agreement within 2 steps in 93.6%. Unweighted κ statistic (SE): 0.58 (0.015), κ weighted to give 75% credit for 1-step disagreement: 0.73(0.013).	MODERATE
Danis et al (2013) Retrospective cohort	AREDS 9-step severity scale	Serious[1]	N/A	Not serious	Not serious	1335 eyes from the AREDS2 study	Contemporaneous regrades, (interobserver agreement) (n=1335) Agreement: 96% Weighted Kappa (SE): 0.76	MODERATE

Studies	Classification System	Risk of bias	Inconsistency	Indirectness	Imprecision	Clinical population (n)	Effect	Quality
							(0.01) Historical AREDS Temporal Drift (AREDS Report 6 and 17), (n=119) Agreement: 94% Weighted Kappa (SE): 0.73 (0.01)	
AREDS 6, (2001) Retrospective cohort	AREDS 4-step severity scale	Serious[1]	N/A	Not serious	Not serious	1230 eyes from the AREDS study	Interobserver contemporaneous reproducability AMD severity level Agreement- 82.8% Agreement within 1 step: 98.7% Kappa, unweighted (SE)- 0.77 (0.01) Kappa, weighted (SE)- 0.88 (0.01)	MODERATE
Seddon 2006 Retrospective cohort	CARMS	Serious[1]	N/A	Not serious	Not serious	492 eyes recruited for the Progression of Age-Related Macular Degeneration Study	Agreement between Clinical observations and Reading Centre. Agreement: 75% Agreement within 1 step: 89% Kappa, unweighted (95% CI): 0.63 (0.53-0.74) Kappa, weighted (95% CI): 0.78 (0.62-0.93)	MODERATE

Studies	Classification System	Risk of bias	Inconsistency	Indirectness	Imprecision	Clinical population (n)	Effect	Quality
							Agreement between 2 observers assessments of Age-Related Maculopathy. Agreement: 84% Agreement within 1 step: 90% Kappa, unweighted (95% CI): 0.79 (0.47-1.1) Kappa, weighted (95% CI): 0.86 (0.41-1.3)	
Hamada (2006) Retrospective cohort	The Modified International Classification of ARM	Serious[1]	N/A	Not serious	Not serious	164 images of 106 patients taken from consecutive patients referred to the Retinal Research Unit at King's College Hospital.	Interobserver consistency between the two graders: Kappa value of 0.82 (SE 0.34).	MODERATE
Leeuwen (2003) Retrospective cohort	The Modified International Classification of ARM	Serious[1]	N/A	Not serious	Not serious	91 subjects in the EUREYE study. 131 images of eyes taken to represent the full range of AMD.	On all 8 stages: digital images Agreement: 59.0 Weighted kappa: 0.72 On all 8 stages: 35-mm film Agreement: 65.7% Weighted kappa: 0.78 On the 5 main stages: digital images Agreement: 64.9% Weighted kappa: 0.74	MODERATE

Studies	Classification System	Risk of bias	Inconsistency	Indirectness	Imprecision	Clinical population (n)	Effect	Quality
							On the 5 main stages: 35-mm film Agreement: 72.3% Weighted kappa: 0.79	
Klein (2014) Retrospective cohort	Harmonized Three Continent AMD Consortium Severity Scale	Serious[1]	N/A	Not serious	Not serious	60 images from participants of the Beaver Dam Eye Study	Interobserver agreement Exact grading agreement of the 60 eyes between centers: 61.0 - 81.4%, Within-one-step agreement was 84.7- 98.3% between centers. Weighted kappa scores varied from 0.66 to 0.86	MODERATE
1. Downgraded one level for risk of bias due to lack of clarity regarding baseline characteristics of included participants								

Agreement outcomes: Intraobserver Agreement

Studies	Classification System	Risk of bias	Inconsistency	Indirectness	Imprecision	Clinical population (n)	Effect	Quality
Danis et al (2013) Retrospective cohort	AREDS 9-step severity scale	Serious[1]	N/A	Not serious	Not serious	1335 eyes from the AREDS2 study	AREDS2 Temporal Drift Regrade Year 4 Compared to BL, (intraobserver agreement) (n=88) Agreement: 92% Weighted Kappa (SE): 0.73 (0.02)	MODERATE
AREDS 6, (2001) Retrospec	AREDS 4-step severity scale	Serious[1]	N/A	Not serious	Not serious	1230 eyes from the AREDS study	Intraobserver temporal reproducibility AMD severity level	MODERATE

Studies	Classification System	Risk of bias	Inconsistency	Indirectness	Imprecision	Clinical population (n)	Effect	Quality
tive cohort							Agreement- 88.2% Agreement within 1 step: 98.3% Kappa, unweighted (SE)- 0.83 (0.04) Kappa, weighted (SE)- 0.88 (0.04)	
Seddon 2006 Retrospective cohort	Clinical Age-Related Maculopathy Staging (CARMS) system	Serious[1]	N/A	Not serious	Not serious	492 eyes recruited for the Progression of Age-Related Macular Degeneration Study	Intraobserver agreement Agreement: 94% Agreement within 1 step: 100% Kappa, unweighted (95% CI): 0.92 (0.58-1.3) Kappa, weighted (95% CI): 0.97 (0.49-1.4)	MODERATE

1. Downgraded one level for risk of bias due to lack of clarity regarding baseline characteristics of included participants

Validation outcomes for existing sub-classification systems of late wet AMD

Studies	Classification System	Risk of bias	Inconsistency	Indirectness	Imprecision	Clinical population (n)	Effect	Quality
Interobserver agreement								
Classification: 1) Classic only, 2) predominantly classic, 3) minimally classic, 4) occult without PED (with or without RAP) and 5) vascularised PED (with or without RAP).								
Cohen (2007) Prospective cohort	CAMRS	Very serious[1, 3, 4]	N/A	Not serious	Serious[2]	207 patients with newly diagnosed exudative AMD	Lesion classification: Kappa: 0.59 Location of lesion: Kappa: 0.52	VERY LOW
(1) AMD with type 1 CNV; (2) AMD with type 1 + 2 CNV; (3) AMD with type 2 CNV only; (4) Chorioretinal anastomosis (RAP) (5) PCV, (using fundus phot, FA,								

Studies	Classification System	Risk of bias	Inconsistency	Indirectness	Imprecision	Clinical population (n)	Effect	Quality
ICG and OCT)								
Coscas (2014) Prospective cohort	CAMRS	Very Serious[1, 3,]	N/A	Not serious	Serious[7]	99 consecutive Japanese eyes and 94 consecutive French eyes with exudative AMD	Crude agreement with final diagnosis: Range, Kyoto patients (n= 99) AMD with type 1 CNV: 79.4 - 91.1% AMD with type 1+2 CNV: 33.3- 66.6% AMD with type 2 CNV: 60.0- 100% Chorioretinal anastomosis (RAP): 83.3% PCV with type 1 or 2 CNV: 66.6% PCV without type 1 or 2 CNV: 95.6% Other: 100% Range, French patients (n= 94) AMD with type 1 CNV: 95.8 - 97.9% AMD with type 1+2 CNV: 68.4 - 89.5% AMD with type 2 CNV: 60.0 - 100% Chorioretinal anastomosis: 80.0- 100% PCV without type 1 or 2	VERY LOW

Studies	Classification System	Risk of bias	Inconsistency	Indirectness	Imprecision	Clinical population (n)	Effect	Quality
							CNV: 66.6-87.5% Other: 75-100%	
(1) AMD with type 1 CNV; (2) AMD with type 1 + 2 CNV; (3) AMD with type 2 CNV only; (4) Chorioretinal anastomosis (RAP) (5) PCV, (using fundus phot, FA)								
Coscas (2014) Prospective cohort	CAMRS	Very Serious[1, 3,]	N/A	Not serious	Serious[7]	99 consecutive Japanese eyes and 94 consecutive French eyes with exudative AMD	Crude agreement with final diagnosis: Range, Kyoto patients (n= 99) AMD with type 1 CNV: 79.4 – 82.3% AMD with type 1+2 CNV: 16.6- 66.6% AMD with type 2 CNV: 40-80% Chorioretinal anastomosis: 66.6- 83.3% PCV with type 1 or 2 CNV: 33.3% PCV without type 1 or 2 CNV: 56.5-91.3% Other: 66.6-88.8% Range, French patients (n= 94) AMD with type 1 CNV: 89.5% AMD with type 1+2 CNV: 36.8- 78.9% AMD with type 2 CNV: 60.0- 100%	VERY LOW

Studies	Classification System	Risk of bias	Inconsistency	Indirectness	Imprecision	Clinical population (n)	Effect	Quality
							Chorioretinal anastomosis (RAP): 60-80% PCV without type 1 or 2 CNV: 33.3-75% Other: 50-100%	
Anatomic classification (OCT, photo and FA): 1) type 1 (sub–retinal pigment epithelium [RPE], incl PCV), 2) type 2 (subretinal), 3) type 3 (intraretinal, RAP), or 4) mixed NV. **MPS criteria and the Digital Angiographic Reading Center (DARC): occult or classic CNV**								
Jung (2014) Prospective cohort	CARMS	Serious[1, 6]	N/A	Serious[5]	Not serious	374 treatment naïve patients with neovascular AMD in at least 1 eye	Agreement between FA and anatomic classification: Kappa 0.65	LOW
1) Classic only, 2) occult only, 3) mixed, or 4) unable to determine								
Friedman (2000) Retrospective cohort	CARMS	Very serious[1, 3, 4, 6]	N/A	Serious[2]	Not serious	6 fluorescein angiograms read by 21 ophthalmologists	Membrane type Mean agreement, % (SD): 72.5 (23.0) Mean kappa (SD): 0.64 (0.30)	VERY LOW
1) classic, 2) occult, or 3) mixed with classic component less or equal/greater than 50%								
Holz (2003) Prospective cohort	CARMS	Very serious[1, 3, 4]	N/A	Serious[2]	Not serious	40 patients with neovascular ARMD, graded by 16 retinal specialists.	Mean kappa agreement (SD): Randomised series A: 0.40 (0.05) Randomised series B: 0.37 (0.05)	VERY LOW
Predominantly classic, minimally classic, or occult								
Olsen (2004)	CAMRS	Very serious[1, 4, 6]	N/A	Serious[2]	Not serious	200 cases of nAMD from 2 centres	kappa agreement: 0.63	VERY LOW

Studies	Classification System	Risk of bias	Inconsistency	Indirectness	Imprecision	Clinical population (n)	Effect	Quality
Retrospective cohort								
1) Classic only 2) Occult only 3) Classic and Occult (mixed <50%/>50% classic) 4) Disciform scar 5) cannot determine 6) Serous PED (present/absent)								
Maguire (2008) Retrospective cohort	CAMRS	Serious[1]	N/A	Serious[2]	Not serious	282 eyes developed CNV or serous PED in CAPT trial	Agreement: 80-100% Weighted kappa: 0.75-100	LOW
Intraobserver agreement								
classic, occult, or mixed with classic component less or equal/greater than 50%								
Holz (2003) Prospective cohort	CAMRS	Very serious[1, 3, 4]	N/A	Serious[2]	Not serious	40 patients with neovascular ARMD, graded by 16 retinal specialists.	Mean kappa agreement (SD): 0.64 (SD 0.11)	VERY LOW

1. Downgraded one level for risk of bias due to lack of clarity regarding baseline characteristics of included participants
2. Downgraded one level for people with PCV excluded or unclear inclusion
3. Downgraded one level for lack of clear pre-specified criteria for diagnosis or unclear
4. Downgraded one level for some participants received an extra investigation (e.g. ICG angiography) without a clear criteria RE who should receive the extra investigation, possibly inconsistent between graders. Or unclear consistency of investigation.
5. Downgraded one level for agreement between classifications systems with multiple graders, unclear if relevant.
6. Downgraded one level for unclear grading was done without knowledge of other graders decision
7. Downgraded one level for only crude agreement, no adjustment possible

Validation outcomes for existing sub-classification systems of late dry AMD

Studies	Classification System	Risk of bias	Inconsistency	Indirectness	Imprecision	Clinical population (n)	Effect	Quality
CAPT classification of late dry AMD								
Brader (2011)	CAMRS	Serious[1]	N/A	Serious[2]	Not serious	Sample of 15 photographic sets, some	Interobserver variability kappa: 0.536	LOW

Studies	Classification System	Risk of bias	Inconsistency	Indirectness	Imprecision	Clinical population (n)	Effect	Quality
Retrospective cohort						of which included lesions that met the new criteria but not the previously used criteria. Regraded 6m.		
Intraobserver agreement								
classic, occult, or mixed with classic component less or equal/greater than 50%								
Brader (2011) Retrospective cohort	CAMRS	Serious[1]	N/A	Serious[2]	Not serious	Sample of 15 photographic sets, some of which included lesions that met the new criteria but not the previously used criteria. Regraded 6m.	Intraobserver agreement kappa: 0.845	LOW
1. Downgraded one level for risk of bias due to lack of clarity regarding baseline characteristics of included participants 2. Downgraded one level for people with PCV excluded or unclear inclusion								

Clinical risk assessment models: risk outcomes

Studies	Classification system	Risk of bias	Inconsistency	Indirectness	Imprecision	Clinical population (n)	Units	Effect	Quality
Risk of developing neovascular AMD									
Simple Severity Score									
Perlee et al (2013) Prospective cohort study	Simple severity score	Very serious [1, 2, 5]	N/A	Not serious	Not serious	Participants in the Age-Related Eye Disease Study (n=2415)	HR (95% CI)	Hazard Ratios for Progression to neovascular AMD 0) referent 1) 4.76 (2.43-9.34)	LOW

Studies	Classification system	Risk of bias	Inconsistency	Indirectness	Imprecision	Clinical population (n)	Units	Effect	Quality
								2) 12.66 (6.87-23.36) 3) 26.56 (14.53-48.58) 4) 35.89 (19.75-65.21)	
Sandberg 4-point scale									
Sandberg (1998) Prospective cohort study	Sandberg 4-point scale	Very Serious [1, 2, 3]	N/A	Not serious	Very serious7	patients with unilateral neovascular AMD (127)	HR (95% CI)	Hazards ratio for development of choroidal neovascular membrane (95% confidence intervals) 1.76 (1.18-2.73)	VERY LOW
Risk of developing geographic atrophy									
Simple Severity Score									
Perlee et al (2013) Prospective cohort study	Simple severity score	Very serious [1, 2, 5]	N/A	Not serious	Nots serious	Participants in the Age-Related Eye Disease Study (n=2415)	HR (95% CI)	Hazard Ratios for Progression to geographic atrophy 0) referent 1) 6.97 (3.01-16.14) 2) 9.33 (4.13-21.05) 3) 23.29 (10.59-51.22)	LOW

Studies	Classification system	Risk of bias	Inconsistency	Indirectness	Imprecision	Clinical population (n)	Units	Effect	Quality
								4) 34.81 (16.02-75.65)	
Risk of developing advanced AMD									
Simple Severity Score									
Klein et al (2011) Prospective cohort study	Simple severity score	Very serious [1, 2, 3]	N/A	Not serious	Not serious	Participants in the Age-Related Eye Disease Study (n=2846)	HR (95% CI)	Hazard Ratios for Progression to Advanced Age-Related Macular Degeneration at 2, 5, and 10 Years (95% Confidence Interval) Simple scale score 0- referent 1- 6.38 (3.48-11.69) 2- 14.12 (8.06-24.75) 3- 34.53 (19.79-60.26) 4- 50.65 (28.86-88.89)	LOW

1. Downgraded one level for risk of bias due to the study sample (for example, the paper is not clear about how many people were eligible for the study and were not included, there was no meaningful comparison between those included in the study and the population of interest for important differences)
2. Downgraded one level for risk of bias due to the study attrition (for example, the paper is not clear about how many people were lost to follow up in the study and/or had missing data, there was no meaningful comparison between those lost to follow up or with missing data in the study and the rest of the included sample)
3. Downgraded one level for risk of bias due to the confounding factor measurement (for example, the paper is not clear about how the confounding factors were

Studies	Classification system	Risk of bias	Inconsistency	Indirectness	Imprecision	Clinical population (n)	Units	Effect	Quality
measured, it is not clear which confounders were adjusted for in analysis, not all the important confounders were adjusted for) 4. Downgraded one level for imprecision was defined by crossing the minimum important difference defined by NICE for showing an effect (0.80 or 1.25), if the confidence intervals crossed two lines of minimum important difference this was defined as very serious imprecision. 5. Downgraded one level for risk of bias due to adjustment for confounders (confounding measurement and account).									