

Prospective validation of a virtual clinic pathway in the management of choroidal naevi
The NAEVUS study: Report No 1 – Safety Assessment

Lamis Al Harby¹, Zaria Ali², Azita Rajai², Stephen A. Roberts², Tunde Peto³, Irene Leung¹, Jane Gray², Gordon Hay¹, Amit K. Arora¹, Pearse Keane¹, Victoria M. L. Cohen¹, *Mandeep S. Sagoo¹, *Konstantinos Balaskas^{1,2}

1. NIHR Biomedical Research Centre for Ophthalmology at Moorfields Eye Hospital NHS Foundation Trust and University College London Institute of Ophthalmology, London, UK
2. University of Manchester, School of Biological Sciences, Manchester, UK
3. Department of Ophthalmology, Queen's University, Belfast, N. Ireland, UK

*Joint Senior authors

Correspondence:

Mr K. Balaskas, Moorfields Eye Hospital, City Road, London, EC1V 2PD, United Kingdom.

Email: k.balaskas@nhs.net

Tel: 020 72533411

Running title: Prospective Validation of Virtual Naevus Clinic

Keyword: Naevus, Melanoma, Choroid, Virtual Clinic

Number of Tables: 5

Number of Supplementary Tables: 2

Number of Figures: 1

Word count: 3422

SYNOPSIS

Choroidal naevi are a common benign incidental finding, often prompting referral to ocular oncologists. This study validates a virtual imaging-based clinic model for the safe management of choroidal naevi.

ABSTRACT

Background

Choroidal naevi are a common incidental finding prompting specialist referrals to ocular oncology. Rarely, such lesions have sufficient suspicious features to diagnose a small melanoma. The aim of the study is to show that ‘virtual’ imaging-based pathways are a safe and efficient option to manage such referrals.

Methods

A prospective cohort study at the Manchester Royal Eye Hospital and Moorfields Eye Hospital between June 2016 - July 2017 of the management decision of 400 patients reviewed by an ophthalmologist in a face-to-face consultation (gold standard) supported by fundus photography, optical coherence tomography (OCT), autofluorescence (AF) and B-mode ultrasound.. The images were also read independently by masked graders (non-medical) and masked ophthalmologists and a management decision was made based on image review alone (virtual pathway). The two pathways were compared for safety.

Results

The agreement for management decisions between face-to-face and virtual pathways was 83.1% (non-medical) and 82.6% (medical). There were more over-referrals in the virtual pathway (non-medical 24.3%, medical 23.3% of gold standard discharge) and only 2 under-referrals (10.5% of gold standard referrals), both borderline cases with minimal clinical risk. Agreement for risk factors of growth (orange pigment, subretinal fluid, hyper-autofluorescence) ranged between 82.3% and 97.3%.

Conclusions

We prospectively validated a virtual clinic model for the safe management of choroidal naevi. Such a model of care is feasible with low rate of under-referral. An over-referral rate of almost 24% from the virtual pathway needs to be factored into designing such pathways in conjunction with evidence on their cost-effectiveness.

INTRODUCTION

Choroidal naevus describes an ocular fundus lesion akin to a freckle or mole. They are common, present in up to 7% of the general population and are increasingly identified during routine eye checks in community optometry services as an incidental finding.^{1,2} The majority are benign but the potential for missing an early melanoma often prompts urgent over-referrals to ophthalmologists and even ocular oncologists³. Overall, they carry a small potential for transformation into melanoma.^{1,4-8} There are exceptions, such as giant choroidal naevi (10 mm or more in diameter), which were estimated to transform into melanoma in 18% over 10 years.^{9,10}

In several reports, risk factors of growth have been outlined, including thickness (T) more than 2 mm, subretinal fluid (F), symptoms (S) flashes/floaters/blurred vision, orange (O) lipofuscin pigment and margin (M) less than 3 mm from optic disk.¹¹⁻¹³ This schema has been modified recently to include imaging and autofluorescence to judge orange pigment and optical coherence tomography to ascertain the presence of fluid.^{14,15}

In our previous study,¹⁶ we described a model of managing such lesions in a virtual pathway, comparing retrospectively the outcomes from imaging graded by a masked non-medical grader or a masked ophthalmologist versus clinical assessment of patients by an ophthalmologist, the gold standard. In 101 cases, agreement was impressive at 96.1 % between the masked non-medical grader and 100 % between the masked ophthalmologist with the gold standard clinical encounter. However, we cautioned against the widespread adoption of this pathway until a larger cohort had been prospectively assessed. In the digital era, evolutions in imaging and communication technologies offer unprecedented opportunities for service transformations that are particularly relevant to Ophthalmology. Although such innovative service redesigns are often piloted at small scale and demonstrate early potential, the decisions by policy makers to transition to such novel service models needs to be informed by robust evidence from well-designed implementation science research. The NAEVUS study was designed to address all aspects of implementation science for the proposed virtual model of care, including safety, cost effectiveness and acceptability by patients and healthcare practitioners.

In this report we present the results of the NAEVUS study on a large prospective cohort to validate a virtual model for managing choroidal naevi referrals in terms of its safety.

METHODS

Study design

This is a prospective analysis of the clinical details and imaging of patients attending the naevus clinics at the Manchester Royal Eye Hospital (MREH) and Moorfields Eye Hospital (MEH). Our aims were to evaluate the outcomes of a one-stop, dedicated virtual clinic to image 400 eyes referred with choroidal naevi, comparing masked decisions by non-medical graders and ophthalmologists based on review of images alone against the gold standard of full clinical evaluation by an ophthalmologist in a face-to-face consultation, including indirect fundoscopy.

The study had ethical approval (16/NW/0288) and adhered to the tenets of the Declaration of Helsinki. The study was supported by a National Institute of Health Research/Research for Patient Benefit (PB-PG-0215-36081) - 'the NAEVUS study' research grant and sponsored by Manchester University NHS Foundation Trust. Consecutive patients attending the naevus clinics of MREH and MEH gave informed consent and were included. Patients with pathologies other than choroidal naevi (erroneous referrals, including congenital hypertrophy of the retinal pigment epithelium) were excluded from agreement analysis. Accuracy in recognition of alternate pathologies between the virtual and face-to-face pathways was an additional outcome measure.

Patient pathway, clinical and imaging assessments, management decisions

Study subjects were assessed by an ophthalmologist with ophthalmic examination, including fundoscopy. Imaging tests were performed by a technician and stored in the relevant databases in MREH or MEH, including wide-field colour imaging Optos (Optos, Dumfermline, UK), fundus autofluorescence imaging AF (Optos, Dumfermline, UK), optical coherence tomography OCT (Spectralis, Heidelberg Engineering GmbH, Heidelberg, Germany) and B-scan ultrasound. Both centres used the same imaging modalities.

Management decisions (Table 1) were made by the ophthalmologist (the gold standard) and allowed one of three options:

1. Discharge to the care of referring community optometrist (following a single confirmatory follow-up visit at 6 months)
2. Follow-up in hospital-based naevus clinic at a set interval; or
3. Referral for specialist opinion to ocular oncology, on the basis of a specific decision-making algorithm.

All patients were offered a second appointment at 6 months (including in cases deemed of negligible risk suitable for discharge to the care of community optometrists) as per existing

protocols. This fail-safe process was meant to verify the appropriateness of original decision-making and also in order to fulfil the guideline of the Royal College of Ophthalmologists for lesion growth to be examined amongst the criteria for patient referral to specialist ocular oncology services.

Masked grading and definition of gold standard

Imaging data, including wide-field fundus images, AF and OCT was transferred to the Reading Centre, UCL Institute of Ophthalmology in a pseudo-anonymised coded form. These were assessed by both experienced non-medical graders (trained optometrists) and by medical graders (ophthalmologists) within the Reading Centre on the basis of strict guidelines, with same possible management outcomes as the gold standard (Table 1). The images were reviewed by two individual optometrists (non-medical graders in Reading Centre) and two ophthalmologists. Each case was assessed by a single grader after project-specific training.

The gold standard management decision for validating the virtual service model was defined as the decision reached following combined clinical examination and assessment by imaging tests, confirmed at a follow-up visit at 6 months. If a different management decision was reached at the clinical follow-up visit than baseline, this was the one considered the gold standard.

Outcomes analysis

Three pathways for reaching management decisions for patients with choroidal naevi were evaluated: decisions made in clinic on the basis of imaging tests and direct clinical assessment as confirmed at 6 months (face-to-face pathway, gold standard); decisions made by an expert ophthalmologist in a blinded fashion on the basis of imaging tests alone (KB, MSS) and decisions made by Reading Centre non-medical graders (specialist optometrists) on the basis of imaging tests alone (virtual pathways).

The degree of agreement between the gold standard management decisions and those reached by the two alternative 'virtual' pathways were calculated. In cases of disagreement in management decisions that questioned the gold standard, adjudication was made by an experienced ocular oncologist (MSS). These criteria were applied to cases where patient safety could be affected by a failure in the gold standard ('discharge' or 'follow-up' recommended by gold standard as opposed to 'referral' by grader).

The primary outcome was the agreement in clinical management decisions between face-to-face clinical pathway (gold standard) and each of the two virtual clinical pathways (medical and non-medical).

Secondary outcomes included:

- Rate of incorrect and unsafe decisions in the virtual pathways (refer to oncology by gold standard and otherwise by masked non-medical grader/ masked ophthalmologist in virtual pathway),
- Rate of incorrect yet safe (discharge by gold standard and otherwise by masked non-medical grader/ masked ophthalmologist),
- Agreement in correct diagnosis (choroidal naevi or other) between gold standard and virtual pathways. Other lesions included Congenital Hypertrophy of the Retinal Pigment Epithelium, scar tissue, RPE changes and any other lesions referred as suspicious choroidal naevi that was deemed to be an alternative pathology,
- Agreement on feature-based measurements for known risk factors for growth (orange pigment, subretinal fluid, increased autofluorescence and its origin) between gold standard and virtual pathways.

For dichotomous ratings, agreement was measured by kappa score and for ordinal ratings agreement was measured by weighted kappa score using quadratic weights in which the weight assigned to each cell is $1 - ((i-j)/(k-1))^2$ where i and j index the rows and columns of the ratings by the two graders and k is the maximum number of possible ratings (herein 3). Therefore, the further apart the two ratings are, the smaller weight (less participation in agreement) is assigned to the relevant cell. A bootstrapping method was used to estimate a bias corrected 95% confidence interval (CI) for weighted kappa score (K). All analyses were performed with Stata software version IC 13.1 (StataCorp, College Station, TX, USA).

An economic evaluation and patient and healthcare practitioner acceptability analysis were also undertaken (to be reported separately) to inform implementation plans for virtual pathways in terms of their safety, cost-effectiveness and acceptability.

RESULTS

During the study period (June 2016 - July 2017), 400 cases were collected and analysed in three pathways of gold standard assessment in naevus face-to-face clinic, virtual image analysis by masked ophthalmologist and virtual image analysis by a masked trained non-medical grader. 200 patients from MREH and 200 patients from MEH were recruited, of which 4 (1%) did not

have a complete set of images and were excluded from analysis.

Image grading at the Moorfields Reading Centre was based on the presence or absence of known clinical imaging features in line with established risk stratification algorithms for choroidal naevi (e.g. Royal College of Ophthalmologists Guidelines).¹⁷ Features include orange pigment (n=256), subretinal fluid on OCT scan (n=218), increased autofluorescence (n=220) and drusen presence (n=198).

Agreement of clinical management decision between gold standard versus masked ophthalmologist and masked non-medical grader

As primary outcome of this study the absolute agreement for clinical management decision between gold standard and masked non-medical grader was 83.1% (329/396, 95% CI 79.0% to 86.6%, Kappa 0.77), while absolute agreement between gold standard and masked ophthalmologist was 82.6% (327/396, 95% CI 78.5% to 86.2%, Kappa 0.76) (Table 2). Images A-C in Figure 1 show an example of a case where all graders and the gold standard agreed to discharge back to the care of community optometrist.

More patients were sent for follow-up or refer by masked non-medical grader or masked ophthalmologist while more patients were discharged by gold standard.

The agreement for management decision between masked ophthalmologist and masked non-medical grader was high at 98.5% (390/396, 95% CI 96.7% to 99.4%, Kappa 0.98) (Table 3).

‘Under-referral’ or ‘incorrect and potentially unsafe’ were cases which were referred to ocular oncologist by gold standard but decided otherwise (follow-up or discharge) by masked non-medical grader or masked ophthalmologist. The sensitivity of the masked non-medical grader or masked ophthalmologist in making decision to refer was calculated at 89.5% (95% CI 66.9% to 98.7%). The rate of under-referral was 10.5% (2/19, 95% CI 1.3% to 33.1%) by both masked grader and ophthalmologist. No cases of lesion growth were identified at the 6-month safety-net appointment for the group with a decision to discharge to community optometry.

There were 2 cases that were referred to ocular oncology by gold standard, but both masked non-medical grader and masked ophthalmologist decided not to refer. Review of these cases showed that both were borderline, and the masked non-medical grader and masked ophthalmologist recommended close monitoring within two months. In one case, both masked non-medical grader and masked ophthalmologist decided to follow-up because the lesion was within 1 DD of the optic disc while the gold standard advised referral to oncology (Fig 1, D-F). In that case, both masked assessors did not grade for presence of subtle peripapillary SRF.

Adjudication by the expert ocular oncologist confirmed that management of these cases was safe. Therefore, the overall rate of potentially unsafe decisions was 2/186 (1.1%, 95% CI 0.1% to 3.8%) in both virtual pathways and the sensitivity in decision to refer or follow-up was 98.9% (95% CI 96.2% to 99.9%). (Supplementary Table 1) (Supplementary Table 2).

Similarly, ‘over-referral’ or ‘incorrect yet safe’ were cases that were discharged by gold standard but managed otherwise (follow-up or refer) by masked non-medical grader or masked ophthalmologist. The rate of over-referral was 24.3% (51/210, 95% CI 18.6% to 31.0%) by masked non-medical grader and 23.3% (49/210, 95% CI 17.8% to 29.7%) by masked ophthalmologist. Images G-I in Figure 1 show an over-referred case where masked non-medical grader decided to refer to oncology because visual acuity (VA) was reduced in the affected eye associated with thickness >1.5 mm on ultrasound. The gold standard instead advised follow-up because reduced VA was not considered a related symptom in this case.

Agreement of diagnostic of choroidal naevi or other lesions between gold standard versus masked non-medical grader

9 cases were excluded from this part of the analysis (because of not recorded diagnosis). In distinguishing between choroidal naevi and ‘other’ lesions the masked non-medical grader demonstrated absolute agreement in 85.5% cases (331/387, 95% CI 81.6% to 88.9%). The Kappa agreement was 0.67 (95% CI 0.59 to 0.75) (Table 4).

Agreement between gold standard and masked non-medical grader for risk factors of growth

The agreement rate for the presence of orange pigment detected on clinical photography by a masked non-medical grader in comparison with gold standard was 97.3% (249/256, 95% CI 94.2% to 98.8%, Kappa 0.71), for subretinal fluid on OCT 96.3% (210/218, 95% CI 92.6% to 98.3%, Kappa 0.82) and for presence of increased autofluorescence in wide-field AF imaging 82.3% (181/220, 95% CI 76.4% to 86.9%, Kappa 0.27). Increased AF attributable to drusen only had agreement of 84.8% (168/198, 95% CI 79.9% to 89.4%, Kappa 0.48). (Table 5)

DISCUSSION

Choroidal naevi are a frequent incidental finding¹ but distinguishing between a choroidal naevus and a small choroidal melanoma is not universally straightforward.¹⁸ Choroidal naevi are generally stable, and growth of a choroidal naevus is usually a sign of malignant transformation,^{19, 20} especially in a relatively short time period. With an increasing number of such lesions referred from the community, there is a burden on ophthalmologists and ocular

oncologists for their timely management and the reassurance to the referrer and the patient that the lesion has no suspicious features. This prospective study aimed to answer the question of whether those low-risk, incidental findings might be safely managed using clinical imaging and an algorithm to make an appropriate management decision in a virtual clinical pathway by masked non-medical graders and/or masked ophthalmologist graders, using a prospective protocol rather than a retrospective chart review.¹⁶

We studied 400 participants over a 13-month period in 2 major centres being assessed in a dedicated naevus clinic, comparing masked decisions by a trained non-medical grader and an ophthalmologist against the gold standard of full clinical evaluation by an ophthalmologist confirmed at a 6-month follow-up visit (the gold standard).

The primary outcome of decision agreement between the gold standard ophthalmologist and masker grader or masked ophthalmologist were 83.1% and 82.6% respectively. Within these figures, more patients would be followed up in the naevus clinic by the masked grader and ophthalmologist alike, while more patients were discharged back to the community from the face-to-face clinic. This indicates that the discrepancies in performance between the face-to-face and the virtual pathways were driven by a tendency of virtual clinic graders to be conservative.

The suggestion that the virtual pathway is robust enough to be used by a masked ophthalmologist or trained non-medical grader is further supported by their agreement in decision making which was 98.5%, carrying a Kappa value of 0.98. Hence a virtual naevus clinic model could be staffed by trained non-medical graders rather than being reliant on ophthalmologists to carry out standardised decision making.

A potential limitation of this study is that non-medical graders, although trained optometrists, were also expert Reading Centre graders experienced in review of imaging data for research purposes. This may not be fully representative of optometrists in purely clinical roles or in community practice. However, the non-medical graders achieved slightly better agreement with gold standard than medical graders in a number of features. A potential explanation is that non-medical graders had limited previous clinical exposure to choroidal naevi, received task-specific training for assessing images of choroidal naevi and are generally more likely to adhere closely to the image grading protocol. This is a pattern observed in other use-cases where optometrists are upskilled to deliver shared care for diabetic retinopathy and AMD clinics where consistent adherence to grading and decision-making algorithms has led to high levels of performance.²¹

In the 6-month follow-up period, there were no cases of documented growth or malignant transformation. This is a short interval, and melanomas are slow growing tumours. Although this is a limitation, the design of the study was to test the virtual versus face-to-face decision at first presentation, with the follow-up being a safety net. There is scope for further studies to document follow-up protocols for malignant transformation, either virtual or with a clinician. In the Covid-19 pandemic, where progress in teleophthalmology has accelerated, it is possible that many naevi can be followed on virtual pathways to avoid patient travel and exposure. As Roelofs et al have recently reported, ultrasonography may be unnecessary for detecting progression in choroidal naevi.²² In that study, the sensitivity for detecting progression with sequential colour photographs alone was 97% (95% CI 93–100%) and increased to 99% (95% CI 97–100%) by including autofluorescence and OCT. In other words, increase in basal diameter occurs more frequently than an increase in only the ultrasound height, paving the way for virtual monitoring of naevi using readily available modalities of colour photographs, autofluorescence and OCT.

In planning a virtual naevus clinic service, the reported 23-24% rate of over-referrals must be factored in. In this scenario patient safety is not compromised but the service model of a combination of virtual and face-to-face encounters becomes less efficient. A cost-effectiveness analysis is being undertaken as part of this study combined with this primary analysis to help inform policy decision makers on commissioning such innovative virtual pathways.

The non-medical grader was tested in 2 other ways against the gold standard: distinction between choroidal naevi and 'other lesion' and for the risk factors for growth. In 85.5% of cases, the masked non-medical grader made the same diagnosis as the gold standard. The agreement for the individual risk factors of naevus growth ranges from 82.3% to 97.3%. These agreement rates are already high but in order to improve and optimise the decisions performed by non-medical professionals, a constant audit of grading outcomes should be undertaken with continuing training and education among graders to improve their skills.

The major aim of this study was to test the safety of the virtual naevus clinic pathway rather than diagnostic or clinical factors' agreement. As long as a decision by the masked ophthalmologist or masked non-medical grader resulted in a safe outcome, then the pathway and associated algorithm can be deemed unlikely to cause harm. An incorrect decision can be safe or unsafe, and it is the latter, an under-referral to an ocular oncologist that is potentially most serious. The overall sensitivity of all graders was 89.5% but 98.9% of management decisions were considered safe. The decision defined as potentially unsafe for the purposes of this analysis was not a complete discharge from review, but follow-up in the naevus clinic or

discharge back to a community optometrist for follow-up. It is likely that even with such an outcome a lesion suspicious for a small melanoma would have been picked up during that follow-up. Agreement on growth at follow-up is possible to include in a separate study.

The idea of a virtual naevus clinic is attractive for patients and healthcare systems. A fast, one-stop, streamlined and comprehensive service seeks to optimise patient experience. We anticipate reduced delays and inconvenience from repeated visits. A safe, standardised model ensures homogeneous management of cases, appropriate and prompt return of care closer to home to community-based optometrists and minimises unnecessary referrals to ophthalmologists and ocular oncology centres. For healthcare systems there is an efficient use of resources in terms of staff and clinical investigations. A virtual naevus clinic with graders maximises expanded skills for allied health professionals. This alternative model of healthcare tested in this study would be a potential alternative to established tertiary and quaternary care services and will be subject to a health economics analysis. This research work, as well as strategies such as the MOLES scoring system²²⁻²⁴ for triage, could empower community-based providers to deliver management of benign choroidal naevi without referral to specialist units.

In conclusion, our results show that images of choroidal naevi evaluated remotely using a decision-making algorithm by masked non-medical graders or masked ophthalmologists is safe. This work prospectively validates a virtual naevus clinic model focusing on patient safety as the primary consideration while also highlighting a potential for inefficiency associated with a high rate of over-referrals. This needs to be reviewed in conjunction with the economic evaluation of this model and any further iteration of the virtual pathways that could improve efficiency such as enhanced adjudication and quality review of performance by ocular oncologist specialists. Further development of the proposed virtual pathway could embed artificial intelligence Decision Support Systems for image grading and classification to make the process even more effective.

TABLES

Table 1 - Management algorithm		
Management decisions		
Refer to Ocular Oncology subspecialist clinic	<p>A: Any one of the following:</p> <ul style="list-style-type: none"> ▶ Thickness >2.0 mm ▶ Collar-stud configuration ▶ Documented growth 	<p>OR B: Any two of the following:</p> <ul style="list-style-type: none"> ▶ Thickness >1.5 mm ▶ Orange pigment ▶ Serous retinal detachment ▶ Symptoms
Follow-up in naevus clinic at set interval if anyone of the following	<ul style="list-style-type: none"> ▶ Naevus detectable on ultrasound ▶ Orange pigment ▶ Subretinal fluid ▶ Location in extreme periphery ▶ Base diameter >3 DD ▶ More than one naevus in one eye ▶ Within 1 DD of the optic disc 	
Discharge back to the care of community optometrist	If none of the criteria set out for follow-up or refer are present	
Reading Centre masked non-medical graders and masked ophthalmologists assessing imaging test results in isolation had the option of the above management decisions.		

		Table 2 - Summary of comparisons of Clinical Management Decisions between face-to-face (Gold Standard) and virtual pathways (Masked non-medical Graders)					
		Masked non-medical Grader			Masked Ophthalmologist		
		Clinical Decision					
		Discharge	Follow-up	Refer	Discharge	Follow-up	Refer
Gold Standard Clinical Decision	Discharge	159	51	0	161	49	0
	Follow-up	2	153	12	6	149	12
	Refer	0	2	17	0	2	17
	Absolute agreement, n (%) (95% CI)	329 (83.1%) (79.0% - 86.6%)			327 (82.6%) (78.5% - 86.2%)		
	Kappa agreement, (95% CI)	0.77 (0.72 - 0.82)			0.76 (0.71 - 0.82)		
	Sensitivity in decision to refer, (95% CI)*	89.5% (66.9% - 98.7%)			89.5% (66.9% - 98.7%)		
	Rate of under-referral, n (%) (95% CI)	2/19 (10.5%) (1.3% - 33.1%)			2/19 (10.5%) (1.3% - 33.1%)		
	Specificity in decision to discharge, (95% CI)^	75.7% (69.3% - 81.4%)			76.7% (70.4% - 82.2%)		
Rate of over-referral, n (%) (95% CI)	51/210 (24.3%) (18.6% - 31.0%)			49/210 (23.3%) (17.8% - 29.7%)			
*Sensitivity or true positive rate where positive is "Refer to oncology".							
^Specificity or true negative rate where negative is "Discharge".							

		Table 3 - Comparison of Clinical Management Decisions between two versions of virtual pathway (Masked Ophthalmologist versus Masked non-medical Grader)		
		Masked non-medical Grader		
		Discharge	Follow-up	Refer
Masked Ophthalmologist	Discharge	161	6	0
	Follow-up	0	200	0
	Refer	0	0	29
	Absolute agreement, n (%) (95% CI)	390 (98.5%) (96.7% - 99.4%)		
	Kappa agreement, (95% CI)	0.98 (0.96 - 0.99)		

		Table 4 - Comparison of Choroidal Naevi or 'Other' diagnosis between face-to-face (Gold Standard) and virtual pathway (Masked non-medical Grader)	
		Diagnosis by Masked non-medical Grader	
Diagnosis by Gold Standard		Naevus	Other
	Naevus	234	21
	Other	35	97
	Absolute agreement, n (%) (95% CI)	331/387 (85.5%) (81.6% - 88.9%)	
	Kappa agreement, (95% CI)	0.67 (0.59 - 0.75)	
Sensitivity in detecting erroneous referrals, (95% CI)	73.5% (65.1%, 80.8%)		

			Table 5 - Agreement between Gold Standard and Masked non-medical Grader for risk factors of growth		
			Masked non-medical Grader		
			No	Yes	Kappa (95% CI)
Gold Standard	Orange Pigment (n= 256)	No	240	2	0.71
		Yes	5	9	(0.45 - 0.89)
	Subretinal fluid on OCT scan (n= 218)	No	189	3	0.82
		Yes	5	21	(0.69 - 0.92)
	Increased autofluorescence (n=220)	No	170	31	0.27
		Yes	8	11	(0.11 - 0.43)
	Drusen present (n=198)	No	149	25	0.48
		Yes	5	19	(0.32 - 0.63)

STATEMENTS:

Acknowledgements:

The research supported by the National Institute for Health Research (NIHR) Biomedical Research Centre based at Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology. Furthermore, LAH was generously supported by SEHA Abu Dhabi Scholarships Division. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or SEHA Abu Dhabi.

Statement of Ethics:

The study had ethical approval (16/NW/0288) and adhered to the tenets of the Declaration of Helsinki.

Conflict of Interest Statement:

The authors have no conflicts of interest to report. This manuscript has not previously been submitted for publication. None of the authors have any financial disclosures or conflicts of interest to declare.

Funding Sources:

The study was supported by a National Institute of Health Research/Research for Patient Benefit (PB-PG-0215-36081) - 'the NAEVUS study' research grant and sponsored by Manchester University NHS Foundation Trust.

AUTHOR CONTRIBUTIONS:

Lamis Al Harby

Substantial contribution to acquisition, analysis, interpretation of data for the work

Drafting the work, revising it for critically important intellectual content

Final approval of version to be published

Agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved

Zaria Ali

Drafting the work, revising it for critically important intellectual content

Final approval of version to be published

Agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved

Azita Rajai

Drafting the work, revising it for critically important intellectual content

Final approval of version to be published

Agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved

Stephen A. Roberts

Drafting the work, revising it for critically important intellectual content

Final approval of version to be published

Agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved

Tunde Peto

Drafting the work, revising it for critically important intellectual content

Final approval of version to be published

Agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved

Irene Leung

Drafting the work, revising it for critically important intellectual content

Final approval of version to be published

Agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved

Jane Gray

Drafting the work, revising it for critically important intellectual content

Final approval of version to be published

Agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved

Gordon Hay

Drafting the work, revising it for critically important intellectual content

Final approval of version to be published

Agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved

Amit K. Arora

Drafting the work, revising it for critically important intellectual content

Final approval of version to be published

Agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved

Pearse A. Keane

Drafting the work, revising it for critically important intellectual content

Final approval of version to be published

Agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved

Victoria M.L. Cohen

Drafting the work, revising it for critically important intellectual content

Final approval of version to be published

Agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved

Mandeep S. Sagoo

Substantial contribution to acquisition, analysis, interpretation of data for the work

Drafting the work, revising it for critically important intellectual content

Final approval of version to be published

Agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved

Konstantinos Balaskas

Substantial contribution to acquisition, analysis, interpretation of data for the work

Drafting the work, revising it for critically important intellectual content

Final approval of version to be published

Agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved

REFERENCES

1. Sumich P, Mitchell P, Wang JJ. Choroidal nevi in a white population: the Blue Mountains Eye Study. *Arch Ophthalmol* 1998;116(5):645-50.
2. Brihaye-Van G. [CONTRIBUTION TO THE STUDY OF MELANOTIC TUMORS OF THE UVEA AND THEIR ORIGIN]. *Doc Ophthalmol* 1963;17:163-248.
3. Aleksidze N, Medina CA, Singh AD. De novo Evolution of a Small Choroidal Melanoma. *Ocul Oncol Pathol* 2015;1(2):83-7.
4. Singh AD, Kalyani P, Topham A. Estimating the risk of malignant transformation of a choroidal nevus. *Ophthalmology* 2005;112(10):1784-9.
5. Gonder JR, Augsburger JJ, McCarthy EF, Shields JA. Visual loss associated with choroidal nevi. *Ophthalmology* 1982;89(8):961-5.
6. Shields CL, Furuta M, Mashayekhi A, et al. Visual acuity in 3422 consecutive eyes with choroidal nevus. *Arch Ophthalmol* 2007;125(11):1501-7.
7. Mims JL, 3rd, Shields JA. Follow-up studies of suspicious choroidal nevi. *Ophthalmology* 1978;85(9):929-43.
8. Butler P, Char DH, Zarbin M, Kroll S. Natural history of indeterminate pigmented choroidal tumors. *Ophthalmology* 1994;101(4):710-6; discussion 7.
9. Jovanovic P, Mihajlovic M, Djordjevic-Jocic J, et al. Ocular melanoma: an overview of the current status. *Int J Clin Exp Pathol* 2013;6(7):1230-44.
10. Kaiserman I, Kaiserman N, Pe'er J. Long term ultrasonic follow up of choroidal naevi and their transformation to melanomas. *Br J Ophthalmol* 2006;90(8):994-8.
11. Chien JL, Sioufi K, Surakiatchanukul T, et al. Choroidal nevus: a review of prevalence, features, genetics, risks, and outcomes. *Curr Opin Ophthalmol* 2017;28(3):228-37.
12. Factors predictive of growth and treatment of small choroidal melanoma: COMS Report No. 5. The Collaborative Ocular Melanoma Study Group. *Arch Ophthalmol* 1997;115(12):1537-44.
13. Harbour JW, Paez-Escamilla M, Cai L, et al. Are Risk Factors for Growth of Choroidal Nevi Associated With Malignant Transformation? Assessment With a Validated Genomic Biomarker. *Am J Ophthalmol* 2019;197:168-79.
14. Shields CL, Lim LS, Dalvin LA, Shields JA. Small choroidal melanoma: detection with multimodal imaging and management with plaque radiotherapy or AU-011 nanoparticle therapy. *Curr Opin Ophthalmol* 2019;30(3):206-14.
15. Shields CL, Cater J, Shields JA, et al. Combination of clinical factors predictive of growth of small choroidal melanocytic tumors. *Arch Ophthalmol* 2000;118(3):360-4.
16. Balaskas K, Gray J, Blows P, et al. Management of choroidal naevomelanocytic lesions: feasibility and safety of a virtual clinic model. *Br J Ophthalmol* 2016;100(5):665-70.
17. Ophthalmologists ·TRCo. Referral pathways for adult ocular tumours. 2020.
18. Augsburger JJ, Correa ZM, Trichopoulos N, Shaikh A. Size overlap between benign melanocytic choroidal nevi and choroidal malignant melanomas. *Invest Ophthalmol Vis Sci* 2008;49(7):2823-8.
19. Mashayekhi A, Siu S, Shields CL, Shields JA. Slow enlargement of choroidal nevi: a long-term follow-up study. *Ophthalmology* 2011;118(2):382-8.
20. Doro D, Kotsafti O, Cimatti P. Long-term echographic surveillance of elevated choroidal nevi. *Am J Ophthalmol* 2013;156(3):438-43.e1.
21. Kern C, Kortuem K, Hamilton R, et al. Clinical Outcomes of a Hospital-Based Teleophthalmology Service: What Happens to Patients in a Virtual Clinic? *Ophthalmol Retina* 2019;3(5):422-8.
22. Roelofs KA, O'Day R, Al Harby L, et al. Detecting Progression of Melanocytic Choroidal Tumors by Sequential Imaging: Is Ultrasonography Necessary? *Cancers (Basel)* 2020;12(7).

23. Lamis AlHarby. Distinguishing Choroidal Nevi from Melanomas using the MOLES Algorithm: Evaluation in an Ocular Nevus Clinic. *Ocular Oncology and Pathology* 2020,September (In press).
24. Roelofs KA, O'Day R, Harby LA, et al. The MOLES System for Planning Management of Melanocytic Choroidal Tumors: Is It Safe? *Cancers (Basel)* 2020;12(5).

Table Legend

Table 1: Management algorithm

Table 2: Summary of comparisons of Clinical Management Decisions between face-to-face (Gold Standard) and virtual pathways (Masked non-medical Graders)

Table 3: Comparison of Clinical Management Decisions between two versions of virtual pathway (Masked Ophthalmologist versus Masked non-medical Grader)

Table 4: - Comparison of Choroidal naevi or 'Other' diagnosis between face-to-face (Gold Standard) and virtual pathway (Masked non-medical Grader)

Table 5: Agreement between Gold Standard and Masked non-medical Grader for risk factors of growth

Supplementary Table 1: Comparison of Clinical Management Decisions between face-to-face (Gold Standard) and virtual pathways (Masked Graders) when refer and follow-up decisions are pooled

Supplementary Table 2: Comparison of Clinical Management Decisions between face-to-face (Gold Standard) and virtual pathways (Masked Graders) when refer and follow-up decisions are pooled (sensitivity analysis)

Figure Legend

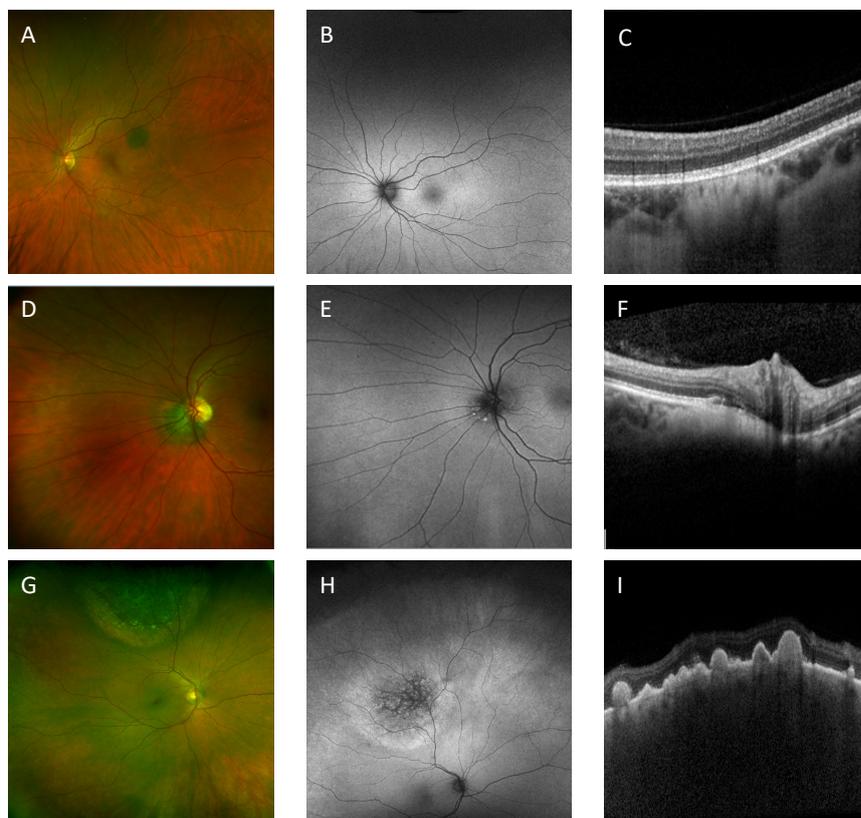
Figure 1: Color photographs, OCTs and fundus autofluorescence images of 3 cases:

(A-C) Total agreement case between gold standard vs masked reader and masked ophthalmologist, with basal diameter <3 DD, no SRF and US no thickness detected;

(D-F) Under-referral case, where masked graders decided to follow-up because of lesion within 1 DD of optic disc and no presence of subtle peripapillary SRF graded, while the gold standard advised referral to oncology;

(G-I) Over-referral case, where masked graders referred to ocular oncology due to physical findings and symptoms but gold standard advised follow-up in a naevus monitoring clinic.

Figure 1



Supplementary Tables

		Supplementary Table 1 - Comparison of Clinical Management Decisions between face-to-face (Gold Standard) and virtual pathways (Masked Graders) when refer and follow-up decisions are pooled	
		Gold Standard	
		Referred/F-up	Discharge
Masked non-medical Grader	Referred/ F-up	184	51
	Discharge	2	159
Masked Ophthalmologist	Referred/ F-up	180	49
	Discharge	6	161

		Supplementary Table 2 - Comparison of Clinical Management Decisions between face-to-face (Gold Standard) and virtual pathways (Masked Graders) when refer and follow-up decisions are pooled (sensitivity analysis)	
		Masked non-medical Grader	Masked Ophthalmologist
Gold Standard	Sensitivity in decision to refer or F-up, (95% CI)*	98.9% (96.2% - 99.9%)	96.8% (93.1% - 98.8%)
	Rate of under-referral, n (%) (95% CI)	2/186 (1.1%) (0.1% - 3.8%)	6/186 (3.2%) (1.2% - 6.9%)