

Development and validation of a patient based measure of outcome in ocular melanoma

Alexander J E Foss, Donna L Lamping, Sara Schroter, John Hungerford

Abstract

Background—Patients with uveal melanoma can be treated by a number of modalities. As none of the different treatments offer a survival advantage, a key factor in choosing among treatments is their differential impact on patients' quality of life. A short, patient based questionnaire was developed and validated for evaluating outcomes following treatment for uveal melanoma.

Methods—The 21 item measure of outcome in ocular disease (MOOD) assesses the patient's view of outcome in terms of visual function and the impact of treatment. The reliability and validity of the three MOOD scores (total, vision, impact) were evaluated in 176 patients who had been treated for uveal melanoma (75 brachytherapy, 78 proton beam radiotherapy, 23 enucleation). Of these, 165 patients also completed the SF-36.

Results—All three MOOD scales met standard criteria for acceptability, reliability, and validity. The proportion of missing data was low, and responses to all items were well distributed across response categories. Internal consistency, assessed by Cronbach's alpha coefficients, exceeded the standard criterion of 0.70 for all three summary scores. Item total correlations ranged from 0.22 to 0.77 (mean item total correlation 0.58), indicating good homogeneity. Test-retest correlations for all three summary scores exceeded 0.85. Scaling assumptions, assessed by item convergent and discriminant validity correlations, were met for the vision and impact scores. The MOOD showed good content validity, as assessed by review by ophthalmologists and patients. Construct validity was demonstrated by high intercorrelations between the vision and impact scores and the total scale; higher scores for patients who reported being very satisfied compared with those who were not very satisfied and for those who reported persistent red eye compared with those who did not have this complication (known group differences/hypothesis testing); moderate correlations between the MOOD and the SF-36 and visual acuity (convergent validity); and low correlations between the MOOD and age and sex (discriminant validity).

Conclusions—The MOOD is a practical and scientifically sound patient based

measure which can be used in research and audit to evaluate outcomes following treatment for uveal melanoma. It takes 5 minutes to complete and meets standard psychometric criteria for reliability and validity.

(Br J Ophthalmol 2000;84:347–351)

Enucleation was the standard treatment for uveal melanoma which has excellent local tumour control rates, but it involves sacrificing the eye. A number of techniques which spare the eye, such as brachytherapy and proton beam radiotherapy, have since been developed. No differences in survival have been reported for patients treated by these modalities.^{1 2} Primary tumour control rates are excellent: 100% for those without extraocular extension treated by enucleation, 95% for brachytherapy, and 98%³ or more for charged particle radiotherapy.¹ However, ocular radiotherapy is not without morbidity^{4 5} and charged particle beam radiotherapy^{6–12} is associated with side effects including keratitis, cataract, scleral and corneal necrosis, radiation retinopathy, radiation optic neuropathy, retinal detachment, phthisis, rubeosis iridis, and glaucoma. Furthermore, enucleation rates as high as 13% have been reported¹¹ for treatment associated morbidity after charged particle beam radiotherapy.

As none of these treatments has a survival advantage over the others, the treatment of choice should be guided by the effects on quality of life. The SF-36, the current gold standard generic measure of quality of life, has been shown to be sensitive to "blurred vision" as a symptom, but it may not be sensitive to other ocular symptoms or to a single eye disease.¹³ Most disease specific measures of outcome in ocular disease, such as the VF-14, cataract symptom score, and the vision related sickness impact profile,^{14 15} were developed mainly for use on cataract patients and have focused exclusively on visual function. These instruments do not assess other outcomes that are important to patients such as cosmesis or pain or discharge in the eye. The measure of outcome in ocular disease (MOOD) questionnaire was developed to assess visual function and the impact of treatment on the eye.

Patients and methods

The study received approval from the Moorfields ethics committee.

Queen's Medical Centre, University Hospital, Nottingham
A J E Foss

Health Services Research Unit, London School of Hygiene and Tropical Medicine, London
D L Lamping
S Schroter

Moorfields Eye Hospital, London
J Hungerford

Correspondence to:
Mr Alexander Foss, Queen's Medical Centre, University Hospital, Nottingham NG7 2HU

Accepted for publication
23 August 1999

PATIENTS

The MOOD was validated in a sample of 176 patients, recruited at the ocular oncology clinic at Moorfields Eye Hospital between December 1993 and June 1994, who had been treated for uveal melanoma 1–197 months (mean 33 months, median 23 months) before the clinic visit. The sample included 100 (57%) men and 76 (43%) women, ranging in age from 22 to 86 years (mean 58.2 years). All patients were white, as ocular melanoma is a disease found predominantly in white people and is rare in other ethnic groups. Seventy eight (44%) had been treated by proton beam radiotherapy, 75 (43%) by brachytherapy, and 23 (13%) by enucleation. A subsample of 32 patients who completed the MOOD twice formed the test retest sample.

METHODS

Patients completed two self completion questionnaires, including the SF-36,^{16,17} a gold standard, generic measure of quality of life, and the measure of outcome in ocular disease (MOOD), a newly developed, disease specific measure to assess the patient's view of outcome in ocular disease (see appendix).

The MOOD is a 21 item questionnaire which measures ocular outcomes in two domains, visual function and impact of treatment. It provides summary scores on three scales: total, vision, and impact. Low scores indicate good outcomes. Instructions for scoring the questionnaire can be obtained from the first author.

The 14 item vision scale includes eight items modified from the VF-14¹⁴ and five new items, all of which are measured on five point Likert scales, and one global question measured on a 100 point visual analogue scale. The seven item impact scale includes six new items measured on five point Likert scales and one new item measured on a four point Likert scale. In addition to the 21 items that are scored quantitatively, the MOOD also includes four open ended questions which are not scored, but which provide descriptive information about the patient's view of the best and worst points of treatment and how treatment could be improved.

Information was also collected for the following: age, sex, pretreatment, and post-treatment vision in both eyes (as measured by a Snellen chart at 6 metres with the vision recorded on an ordinal scale with 6/6 = 1, 6/9 = 2, 6/12 = 3, 6/18 = 4, 6/24 = 5, 6/36 = 6, 6/60 = 7, 3/60 = 8, 1/60 = 9, HM = 10, PL = 11, NPL = 12), treatment modality, months since treatment, time from treatment, whether the treated eye was requiring medication, and complications such as redness (either of the conjunctiva or the lids), presence of strabismus, inflammation, or glaucoma. Patients were given the MOOD on arrival at the clinic and returned the completed questionnaire either on leaving the clinic or by post. A subsample of 32 patients also completed a second MOOD questionnaire at home, sent by post after a 3 week interval.

The reliability and validity of the MOOD were evaluated using standard psychometric techniques.¹⁸ The acceptability of the question-

naire was evaluated through an examination of item non-response rate and the distribution of responses across response categories. Scaling assumptions were tested on the basis of item convergent and discriminant validity correlations. Reliability analyses included internal consistency (Cronbach's alpha) and test-retest reliability for all three summary scores. Item total correlations were used to evaluate the homogeneity of the questionnaire. Validity analyses included an evaluation of content and construct validity, including both within scale analyses (intercorrelations between scales, group differences/hypothesis testing) and analyses against external criteria, including correlations with other measures (SF-36, complications, visual acuity, sociodemographic factors).

Results

Preliminary psychometric analyses of the initial 22 item developmental version of the MOOD detected one item which failed to meet acceptability and reliability criteria. This item ("have you worn a patch to cover your treated/artificial eye?") was therefore eliminated from the final 21 item version of the questionnaire and excluded from further psychometric analyses.

ACCEPTABILITY

The proportion of missing data was low, ranging from 0–5% across items. Responses to all items were well distributed across response categories.

RELIABILITY

Cronbach's alpha coefficients for the three summary scores of the MOOD (Table 1) indicate excellent internal consistency. Alpha coefficients exceeded the standard criterion of 0.70 for all three summary scores. The removal of specific items did not substantially increase the internal consistency of any of the scales.

The homogeneity of the MOOD was evaluated on the basis of item-total correlations. These analyses compute the correlation between each item and the total score with the item of interest eliminated from the calculation of the total score. Item-total correlations below 0.20 are generally eliminated.¹⁸ As can be seen in Table 2, item-total correlations ranged from 0.22 to 0.77 (mean item-total correlation 0.58), indicating good homogeneity and no need for elimination of any items.

Test-retest correlations for the three summary scores, shown in Table 3, all exceeded 0.85, indicating good test-retest reliability.

Scaling assumptions were tested by comparing the correlations between each item and the scale to which it belonged (item own scale correlation) and to the other scale to which it did not belong (item other scale correlation). Item own scale correlations should be higher (item-

Table 1 Reliability: internal consistency

Scale	n	Cronbach's alpha
Total (21 items)	155	0.92
Vision (14 items)	167	0.95
Impact (7 items)	160	0.76

Table 2 Reliability: item-total correlations

Question (item)	n	Item-total r
Vision		
Reading newspapers (1a)	155	0.72
Reading numbers in a telephone directory (1b)	155	0.75
Reading prices (1c)	155	0.76
Reading labels (1d)	155	0.77
Recognising other people (1e)	155	0.66
Seeing steps (1f)	155	0.67
Seeing cracks in the pavement (1g)	155	0.74
Seeing road signs (1h)	155	0.75
Difficulty seeing at night (2a)	155	0.70
Hazy or blurry vision (2b)	155	0.62
Difficulty adapting to bright lights (2c)	155	0.64
Difficulty adapting to dim lights (2d)	155	0.73
In general would you say your vision is . . . (3)	155	0.63
Rating your current vision on a scale from 0 to 100 (4)	155	0.64
Impact		
Pain or discomfort in the treated eye (2e)	155	0.47
Watering of your treated eye (2f)	155	0.32
Stickiness or discharge from the treated eye (2g)	155	0.22
The appearance of your treated eye (2h)	155	0.34
Other people noticing your treated eye (2i)	155	0.29
Felt self conscious about your treated eye (5a)	155	0.46
Satisfied with the appearance of your treated eye (6)	155	0.24

Table 3 Reliability: test-retest correlations

Scale	n	Test-retest r
Total	32	0.88
Vision	32	0.91
Impact	32	0.86

convergent validity) than item other scale correlations (item-discriminant validity) by at least two standard errors.¹⁹ As shown in Table 4, all items in both scales met this criterion.

VALIDITY

Content validity was evaluated during the development of the MOOD. Five ophthalmologists and five patients reviewed the preliminary version of the questionnaire for completeness and appropriateness of content. The questionnaire was also pretested on a further five patients. The responses to open ended questions obtained during subsequent field testing in the full scale validation phase were also reviewed to determine whether patients identified any important new areas which had not been included in the questionnaire. Of the 153 patients who answered the open ended questions, 60% had nothing to add. The concerns raised by the remainder are presented in Table 5. Except for the responses from two patients about concerns about the loss of binocular function, all other responses that are relevant to treatment outcomes (for example, poor vision, pain, discomfort) are covered by the MOOD. The other responses to the open

Table 4 Tests of scaling assumptions: item-convergent and item-discriminant correlations

Question (item)	Vision	Impact
Vision		
Reading newspapers (1a)	0.85	0.14
Reading numbers in a telephone directory (1b)	0.85	0.21
Reading prices (1c)	0.88	0.13
Reading labels (1d)	0.88	0.18
Recognising other people (1e)	0.77	0.17
Seeing steps (1f)	0.74	0.23
Seeing cracks in the pavement (1g)	0.81	0.18
Seeing road signs (1h)	0.83	0.18
Difficulty seeing at night (2a)	0.76	0.27
Hazy or blurry vision (2b)	0.72	0.13
Difficulty adapting to bright lights (2c)	0.73	0.28
Difficulty adapting to dim lights (2d)	0.78	0.27
In general would you say your vision is . . . (3)	0.72	0.16
Rating your current vision on a scale from 0 to 100 (4)	0.64	0.16
Impact		
Pain or discomfort in the treated eye (2e)	0.36	0.55
Watering of your treated eye (2f)	0.22	0.60
Stickiness or discharge from the treated eye (2g)	0.09	0.60
The appearance of your treated eye (2h)	0.14	0.80
Other people noticing your treated eye (2i)	0.08	0.80
Felt self conscious about your treated eye (5a)	0.28	0.69
Satisfied with the appearance of your treated eye (6)	0.08	0.71

Table 5 Points raised in open ended questions. Numbers do not add up precisely as a patient sometimes complained of more than one matter

Response	n
Anxiety about the diagnosis	17
Poor vision in the treated eye	15
Found distance to travel for treatment either to London or to Liverpool, where the cyclotron is based, to be a problem.	15
Pain during treatment	11
Disliked the isolation required for brachytherapy	9
Disliked wearing the mask (part of the treatment for proton beam radiotherapy)	8
Doctors could be kinder	4
Lack of information	3
Complained of discomfort from bright lights during clinical examination	3
Found clinic impersonal	3
Loss of binocular vision	2
Use of eye drops	2

ended questions all pertain to service delivery and are thus more relevant to a measure evaluating the process of care, rather than a measure of the outcome of care such as the MOOD.

The construct validity of the MOOD is supported by three types of within scale analyses. Firstly, the high internal consistency of all three summary scales provides evidence for construct validity. Secondly, as shown in Table 6, intercorrelations between the three scales support the construct validity of the MOOD. The high correlations between the vision and impact scales and the total scale support the convergent validity of the questionnaire, while the moderate correlation between the vision and impact scales provide evidence of good discriminant validity. Thirdly, MOOD scores confirm hypotheses about differences expected between known groups defined on the basis of responses to individual questions on the MOOD. As shown in Table 7, MOOD total and impact scores were significantly higher for patients who reported being very satisfied compared with those who were not very satisfied.

Three types of between scale analyses also support the construct validity of the MOOD. Firstly, MOOD scores confirm hypotheses about differences expected between known groups defined on the basis of complications. As shown in Table 8, MOOD impact scores were significantly higher for patients who reported persistent red eye compared with those who did not have this complication. Secondly, correlations with other measures such as the SF-36 and a measure of visual acuity support the convergent validity of the MOOD. As shown in Table 9, correlations between the MOOD and the SF-36 are all in the moderate range. As both questionnaires measure quality of life, the MOOD at a more specific level than the generic SF-36, moderate correlations are expected. The moderate correlations between the total and vision scales of the MOOD and the measure of visual acuity in the better eye provide support for convergent validity. This compares favourably with correlations of 0.27–0.44 that have been reported between the widely used VF-14 and visual acuity.¹⁵ Thirdly, correlations between the MOOD and sociodemographic variables support the discriminant validity of the questionnaire. As shown in Table 9, correlations between the MOOD and age and sex are low,

Table 6 Construct validity: intercorrelations between scales

Scale	Total	Vision
Vision	0.93	
Impact	0.59	0.25

Table 7 Construct validity. Mean MOOD scores by satisfaction. Significance testing by two tailed *t* test

Scale	Very satisfied (n=131)	Not very satisfied (n=35)	p Value
Total	48.7	53.6	<0.001
Vision	49.2	51.2	0.10
Impact	47.7	58.7	<0.001

Table 8 Construct validity: mean MOOD scores by complications. Significance testing by two tailed *t* test

Scale	Red eye (n=18)	White eye (n=157)	p Value
Total	50.1	49.9	0.44
Vision	49.5	50.1	0.79
Impact	56.0	49.4	0.038

Table 9 Convergent and discriminant validity: correlations between the MOOD and other measures

Measures	MOOD total	MOOD vision	MOOD impact
SF-36 physical component summary score (PCS)	-0.41	-0.39	-0.26
SF-36 mental component summary score (MCS)	-0.36	-0.32	-0.26
Visual acuity	0.29	0.31	0.08
Age	0.05	0.13	0.21
Sex	0.19	-0.16	-0.03

suggesting that responses to the MOOD are not biased in terms of sociodemographic factors. The lower correlation of the MOOD impact scale with the measure of visual acuity compared with the total and vision scales also demonstrates discriminant validity.

Discussion

In recent years, there has been a great expansion in the number and use of instruments to measure quality of life and other patient based outcomes in health care. Rigorous scientific methods, derived from psychometric theory, are now being used by clinicians and health services researchers to identify "good" measures which are credible in scientific as well as clinical terms.

This paper describes the development and validation of the measure of outcome in ocular disease (MOOD). The aim was to develop a measure to assess outcomes in patients treated for ocular melanoma. As none of the treatment modalities has shown a comparative survival advantage, choice should be guided by the effects of treatment on ocular morbidity and quality of life. For example, we have previously reported the outcome in 127 patients treated by proton beam radiotherapy using rubeosis as an indicator of poor outcome.²⁰ Although the development of rubeosis correlated well with outcome (for example, subsequent ocular loss due to local morbidity), it is an indirect measure. Clinical outcome measures need to be supplemented with a more direct measure of greater relevance to patients. Most questionnaires developed for ocular conditions focus predominantly on vision and have been designed specifically

for cataract patients. For many ocular diseases, there are other important outcomes such as ocular appearance, pain and discharge which are not assessed by existing instruments. Although the MOOD was developed and tested on patients with ocular melanoma, it may be applicable to other ocular diseases. Future studies should investigate the validity of the MOOD in other ocular conditions.

The MOOD proved to be highly acceptable to patients. It was designed to be short and suitable for routine use, taking only 5 minutes to complete. The instrument met the standard psychometric criteria for reliability and validity and is thus a scientifically sound measure of outcome in patients with ocular melanoma. A companion paper describes outcomes as measured by the MOOD in groups defined on the basis of tumour stage and grade and treatment modality.

We wish to acknowledge the contribution of Dr Andrew Jaskowski who died shortly after this study began. He introduced AJEF to the field of psychometrics and DLL to ocular oncology and initiated this collaboration.

- Seddon JM, Gragoudas ES, Albert DM, et al. Comparison of survival rates for patients with uveal melanoma after treatment with proton beam irradiation. *Am J Ophthalmol* 1985;99:282-90.
- Char DH, Quivey JM, Castro JR, et al. Helium ions versus iodine 125 brachytherapy in the management of uveal melanoma. A prospective, randomized, dynamically balanced trial. *Ophthalmology* 1993;100:1547-54.
- Gragoudas ES, Egan KM, Seddon JM, et al. Intraocular recurrence of uveal melanoma after proton beam irradiation. *Ophthalmology* 1992;99:760-6.
- MacFaul PA, Bedford MA. Ocular complications after therapeutic irradiation. *Br J Ophthalmol* 1970;54:237-47.
- Char DH, Lonn LI, Margolis LW. Complications of cobalt plaque therapy of choroidal melanomas. *Am J Ophthalmol* 1977;84:536-41.
- Kim MK, Char DH, Castro JL, et al. Neovascular glaucoma after helium ion irradiation for uveal melanoma. *Ophthalmology* 1986;93:189-92.
- Guyet DR, Mukai S, Egan KM, et al. Radiation maculopathy after proton beam irradiation for choroidal melanoma. *Ophthalmology* 1992;99:1278-85.
- Ferry AP, Blair JP, Gragoudas ES, et al. Pathological examination of ciliary body melanoma treated with proton beam irradiation. *Arch Ophthalmol* 1985;103:1849-53.
- Crawford JB, Char DH. Histopathology of uveal melanomas treated with charged particle radiation. *Ophthalmology* 1987;94:639-43.
- Goodman DF, Char DH, Crawford JB, et al. Uveal melanoma necrosis after helium ion therapy. *Am J Ophthalmol* 1986;101:643-5.
- Char DH, Castro JR, Kroll SM, et al. Five-year follow-up of helium ion therapy for uveal melanoma. *Arch Ophthalmol* 1990;108:209-14.
- Meecham WJ, Char DH, Kroll S, et al. Anterior segment complications after helium ion radiation therapy for uveal melanoma. Radiation cataract. *Arch Ophthalmol* 1994;112:197-203.
- Lee PP, Spritzer K, Hays RD. The impact of blurred vision on functioning and well-being. *Ophthalmology* 1997;104:390-6.
- Brenner MH, Curbow B, Javitt J, et al. Vision change and quality of life in the elderly. Response to cataract surgery and treatment of other chronic ocular conditions. *Arch Ophthalmol* 1993;111:680-5.
- Alonso J, Espallargues M, Andersen TF, et al. International applicability of the VF-14. An index of visual function in patients with cataracts. *Ophthalmology* 1997;104:799-807.
- Ware JE Jr, Kosinski M, Keller SD. *SF-36 physical and mental component summary measures: a user's manual*. Boston: The Health Institute, New England Medical Center, 1994.
- Ware JE Jr, Snow KK, Kosinski M, et al. *SF-36 health survey: manual and interpretation guide*. Boston: The Health Institute, New England Medical Center, 1993.
- Streiner DL, Norman GL. *Health measurement scales: a practical guide to their development and use*. 2nd ed. Oxford: Oxford University Press, 1995.
- McHorney CA, Ware JE Jr, Lu JF, et al. The MOS 36-item short-form health survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care* 1994;32:40-66.
- Foss AJE, Whelehan IM, Hungerford JL, et al. Predictive factors for the development of rubeosis following proton beam radiotherapy for uveal melanoma. *Br J Ophthalmol* 1997;81:748-54.

Appendix: Measure of outcome in ocular disease (MOOD)

(1) The following questions ask about problems with your eyesight you might have had *during the last month*. For each question please give one answer. Do you have difficulty, even with glasses, in doing one of the following activities: (please circle one number on each line)

	<i>No difficulty</i>	<i>A little difficulty</i>	<i>Moderate difficulty</i>	<i>Quite a bit of difficulty</i>	<i>A lot of difficulty</i>
(a) Reading newspapers?	1	2	3	4	5
(b) Reading numbers in a telephone directory?	1	2	3	4	5
(c) Reading prices?	1	2	3	4	5
(d) Reading labels?	1	2	3	4	5
(e) Recognising other people?	1	2	3	4	5
(f) Seeing steps?	1	2	3	4	5
(g) Seeing cracks in the pavement?	1	2	3	4	5
(h) Seeing road signs?	1	2	3	4	5

(2) During the past month, how much have you been bothered by... (please circle one number on each line)

	<i>Not at all</i>	<i>A little</i>	<i>Moderately</i>	<i>Quite a bit</i>	<i>Extremely</i>
(a) Difficulty seeing at night?	1	2	3	4	5
(b) Hazy or blurry vision?	1	2	3	4	5
(c) Difficulty adapting to bright lights (eg, going out on a bright day)?	1	2	3	4	5
(d) Difficulty adapting to dim light (eg, entering a darkened room)?	1	2	3	4	5
(e) Pain or discomfort in your treated/artificial eye?	1	2	3	4	5
(f) Watering of your treated/artificial eye?	1	2	3	4	5
(g) Stickiness or discharge of your treated/artificial eye?	1	2	3	4	5
(h) The appearance of your treated/artificial eye	1	2	3	4	5
(i) Other people noticing your treated/artificial eye	1	2	3	4	5

(3) In general, would you say your vision is: (please circle one number)

Excellent	1
Very good	2
Good	3
Fair	4
Poor	5

(4) How would you rate your current vision on a scale from 0 to 100, if 100 is excellent vision and if 0 is blindness _____ (Please enter a number between 0 and 100)?

(5) During the past month, how often have you felt self conscious about your treated/artificial eye? (please circle one number on each line)

None of the time	1
A little of the time	2
Some of the time	3
Most of the time	4
All of the time	5

(6) How satisfied are you with the appearance of your treated/artificial eye? (Please circle one number)

Very satisfied	4
Somewhat satisfied	3
Somewhat dissatisfied	2
Very dissatisfied	1

(7) What were some of the best things about your treatment?

(8) What were some of the worst things about your treatment?

(9) How could we improve upon the treatment you received

(10) Is there anything else which has not been covered and you would like to say?