Efficacy Testing of Skin Whitening Products

John Staton

Dermatest Pty Ltd

Published in Personal Care Asia 2009

Skin Whitening — Lightening — Brightening ? These terms are used inter-changeably[(1)] when describing the claim for what is in effect a benefit of skin care products associated with attempting to arrest or reverse the effects of actinic light damage associated with time and sun exposure. Whilst local terms do appear to vary from country to country, skin "whitening" has been negatively associated with products which were previously marketed for the purpose of reducing the overall colour of the skin, utilising actives such as hydroquinone which can cause permanent depigmentation and sensitivity to sunlight. In most countries these are now considered to be therapeutics and banned from open sale. "Brightening" more or less describes the overall end effect of evening the skin appearance or tone. For the purposes of this paper, I will use the term "Lightening", interpreted as implying a focus on treating deeply pigmented areas.

Testing out vis Formulating in.

The skin lightening properties of any personal care product must obviously be designed into the formulation. As they are invariably a function of one or more "active" ingredients, this is the first place to look when deciding on the claim support protocol. Typically, the providers of these active raw materials provide claims supported by evidence of proven or potential efficacy. Some examples of these are set out below in Fig 1.

Fig 1 - Actives - Scientific Claims

Inhibition of tyrosinase	
Transcriptional control of tyrosinase expression	
Reduction of the melogenic mediators	
Reduction in melanosome transfer to keratinocytes	
Endothelin and alpha-Melanin stimulating hormone	

What is obvious from these claims is that they are mostly too scientific for consumer level, that they are based on in-vitro bio-chemical pathway experiments and that they imply therapeutic definition. For the marketer of a skin lightening preparation, this type of claim substantiation testing is probably not appropriate.

Levels of Claims

It is important to recognise that there are two ways to approach efficacy for these types of formulation – treatment and prevention. Claims also fall into the general classification of high level, low level or simple puffery.

When briefing the test lab for preparation of an appropriate test protocol, the easiest start point is to draft an intended label. This will provide a much closer match for

eventual claim support than simply requesting “to see if it works”. Some claims will be beyond substantiation. Consider for example “strongest formula available over the counter” or “stabilises the immune system of the skin”. In general though, anything relating to visual change, the primary performance expectation of the consumer, will be measurable.

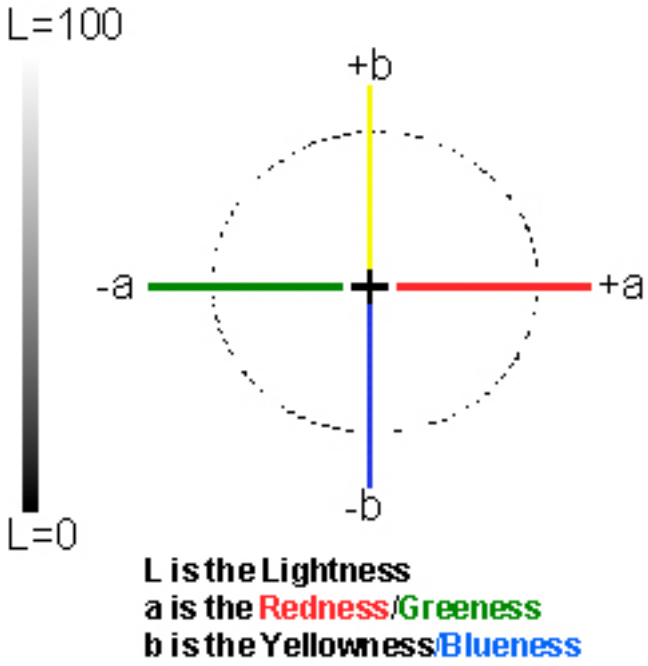
Treatment and Prevention Claims

Treatment testing will imply the use of human test subjects and an in-use study over a reasonably extended period of time and the measurement of skin colour change. Prevention testing is more likely to require in-vitro testing, with the common focus being on challenging the product with ultraviolet light.

Treatment Testing

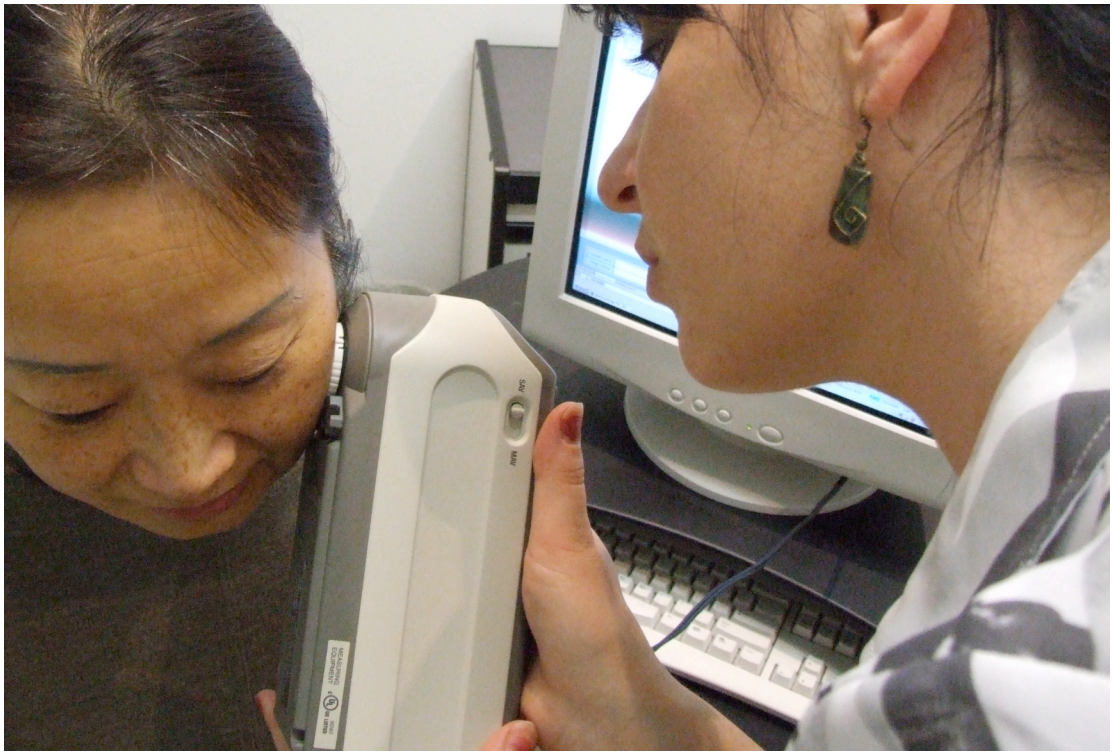
As the consumer’s expectation is the reduction in skin colouration of overly pigmented areas , such as ages spots and blotches, the most appropriate measurement is associated with defining colour change. For this purpose, a chromameter is used, typically in combination with a computer. The instrument most commonly used is a hand held spectrophotometer such as supplied by Konica Minolta. The instrument works by projecting tri-stimulus light [(2)] (similar to an overhead projector) and recording the colour values into an interpretable scale –L*a*b*. These values are derived from the C.I.E. standardised three dimensional colour space model, where L* is the “grey” scale i.e. whiteness - blackness, a* is the redness - greenness value and b* is the yellowness - blueness (see Fig 2). This is not dissimilar to the wheel for colour selection in computer software such as Microsoft Word.

Fig 2 Colour Wheel



For the L*a*b* scale, the L* value to a large degree represents the Lightening (clarity) measurement we are looking for in the test study, but it is modified by the a* and b* values according to skin type as seen in the background colour of an adjoining non-pigmented skin area. In general L* decreases whilst blueness b* and redness a* increase with chronological age[(3)]

Fig 3 Chromameter



Typically, the values obtained will form the quantifiable values for expressing change from the start point of the study. Fig 4 shows an example.

Fig 4 L* Values Charting

L* values															
ID No.	Age	Race	Hyper-pigmented spots												
			DAY 0			Mean	DAY 45			Mean	DAY 90			Mean
L1101	33	C	60.41	60.41	60.41	**60.41**	63.75	63.90	63.88	**63.84**	68.77	68.77	68.76	**68.77**
L1102	39	C	63.56	63.56	63.55	**63.56**	70.44	70.45	70.44	**70.44**	70.50	70.50	70.50	**70.50**
L1103	48	C	51.96	51.96	51.95	**51.96**	52.00	52.03	52.00	**52.01**	63.34	63.33	63.32	**63.33**
L1104	37	C	63.45	63.45	63.44	**63.45**	70.12	70.06	70.07	**70.08**	67.49	67.48	67.46	**67.48**
L1105	45	C	51.99	51.98	51.99	**51.99**	60.02	60.00	60.00	**60.01**	59.86	59.84	59.84	**59.85**
L1106	47	C	54.88	54.88	54.87	**54.88**	54.53	54.53	54.53	**54.53**	57.17	57.16	57.16	**57.16**
L1107	51	C	45.97	45.97	45.95	**45.96**	51.80	51.79	51.80	**51.80**	51.67	51.66	51.65	**51.66**
L1108	60	C	51.27	51.26	51.26	**51.26**	51.07	50.98	51.02	**51.02**	60.02	60.02	60.01	**60.02**
L1109	56	C	63.36	63.36	63.35	**63.36**	63.52	63.52	63.52	**63.52**	63.22	63.22	63.21	**63.22**
L1110	39	C	52.37	52.35	52.33	**52.35**	51.74	51.74	51.74	**51.74**	51.74	51.70	51.73	**51.72**
L1111	47	C	60.88	60.90	60.89	**60.89**	60.55	60.54	60.54	**60.54**	63.28	63.28	63.26	**63.27**
L1112	42	C	54.76	54.76	54.76	**54.76**	54.33	54.35	54.35	**54.34**	54.35	54.35	54.35	**54.35**
L1113	49	C	59.66	59.66	59.69	**59.67**	63.71	63.64	63.67	**63.67**	63.31	63.31	63.30	**63.31**
L1114	47	C	51.85	51.83	51.80	**51.83**	51.66	51.66	51.65	**51.66**	51.64	51.63	51.65	**51.64**
L1115	49	A	45.33	45.32	45.32	**45.32**	45.59	45.57	45.57	**45.58**	45.19	45.18	45.18	**45.18**
L1116	49	C	50.96	50.95	50.98	**50.96**	50.76	50.71	50.75	**50.74**	50.72	50.64	50.66	**50.67**
L1117	46	C	48.56	48.58	48.58	**48.57**	48.71	48.71	48.73	**48.72**	54.40	54.35	54.36	**54.37**
L1118	31	C	51.20	51.35	51.36	**51.30**	50.86	50.85	50.84	**50.85**	59.90	59.88	59.88	**59.89**
L1119	47	C	51.75	51.76	51.75	**51.75**	63.04	63.04	63.04	**63.04**	63.03	63.09	62.97	**63.03**
L1120	51	C	51.79	51.79	51.76	**51.78**	48.68	48.66	48.66	**48.67**	48.67	48.66	48.69	**48.67**
L1121	45	A	63.47	63.48	63.48	**63.48**	70.53	70.53	70.53	**70.53**	70.42	70.42	70.42	**70.42**
L1122	50	A	60.74	60.81	60.86	**60.80**	57.19	57.18	57.18	**57.18**	69.17	69.16	69.16	**69.16**
L1123	51	A	46.04	46.02	46.02	**46.03**	45.76	45.75	45.75	**45.75**	45.72	45.73	45.69	**45.71**
L1124	33	C	65.91	95.91	65.90	**75.91**	71.55	71.55	71.54	**71.55**	65.81	65.79	65.78	**65.79**
L1125	30	C	51.43	51.52	51.51	**51.49**	51.61	51.61	51.60	**51.61**	51.66	51.64	51.54	**51.61**
L1126	36	A	63.45	63.44	63.42	**63.44**	70.53	70.53	70.53	**70.53**	70.53	70.52	70.50	**70.52**
L1127	21	I	57.29	57.25	57.21	**57.25**	52.72	52.71	52.71	**52.71**	52.60	52.59	52.59	**52.59**
L1128	47	C	54.28	54.28	54.27	**54.28**	51.22	51.20	51.19	**51.20**	59.93	59.92	59.88	**59.91**
L1129	51	C	51.01	51.01	51.00	**51.01**	50.94	50.94	50.93	**50.94**	50.92	51.04	51.03	**51.00**
L1130	48	C	51.75	51.75	51.74	**51.75**	51.76	51.76	51.75	**51.76**	64.72	64.76	64.76	**64.75**
L1131	58	C	51.82	51.79	51.88	**51.83**	51.72	51.71	51.70	**51.71**	62.76	62.76	62.75	**62.76**
L1132	41	A	54.62	54.66	54.66	**54.65**	65.90	65.89	65.89	**65.89**	71.29	71.29	71.28	**71.29**

Protocol parameters which must be considered for any study are test subject age group, ethnicity and study duration. Whilst the target market might be, for instance,

age group 25 to 65, the most measurable change will typically be in the older segment of this group. So, is the objective to prove how effective the product is, or to show it works for a 25 year old as well as a 65 year old? The client will need to decide this.

What ethnicity mix should we select? Once again, this will depend on the balance between marketing and scientific considerations. Certainly, Asian skin tends to express age spots as blotching, particularly on the cheeks, whilst Caucasian skin expresses a more speckled appearance of very small blemishes. For darkly pigmented skin, it is much more difficult to measure change as the initial difference from the background is lower to begin with.

Whilst short term studies can be conducted, the slow acting nature of non-therapeutic formulations usually means that a study over 4, 8 or 12 weeks will be needed in order to show statistically significant change, for example - 90% of users showed more than 55% reduction in intensity of deeply pigmented spots after 60 days use.

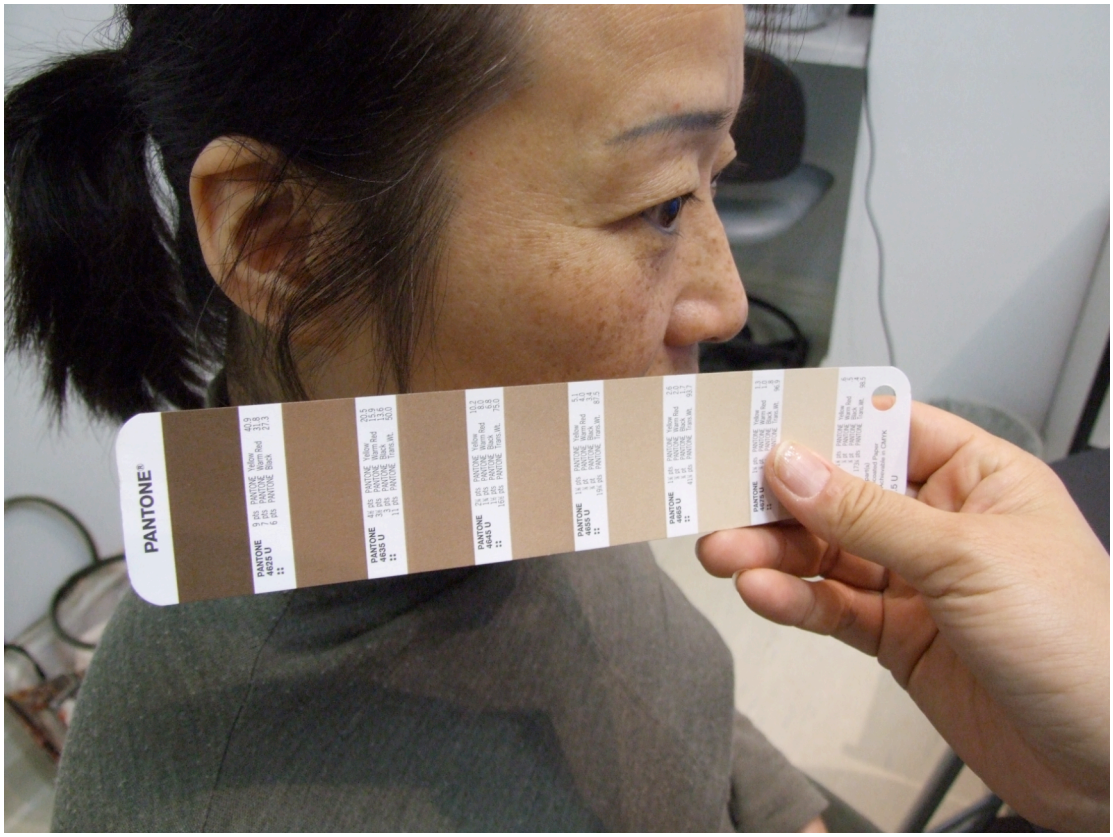
The preferred season for performing the testing is the Northern or Southern hemisphere winter, where actinic light is at its lowest. This gives the best opportunity to make the measurements without the interference of suntan, which will cause changes to the colour of the background unpigmented reference skin areas.

As it is unethical for the clinical tester to restrict the test participants from using their usual sun protection practices during the course of the study, this should be taken into account in preparation of the protocol.

Colour Charting

Standard colour swatches, such as the widely available PMS colour charts can be useful in expressing the colour of skin. They can be used for smaller pigmented areas below the minimum diameter limits of the chromameter. However, the matching is highly subjective and subject to the discriminatory abilities of the technician.

Fig 5 Colour Chart Matching.



Photographic
Photography is a very useful and convenient adjunct to colour measurement. Systems such as Visia provide highly reproducible mapping of progression during treatment testing.

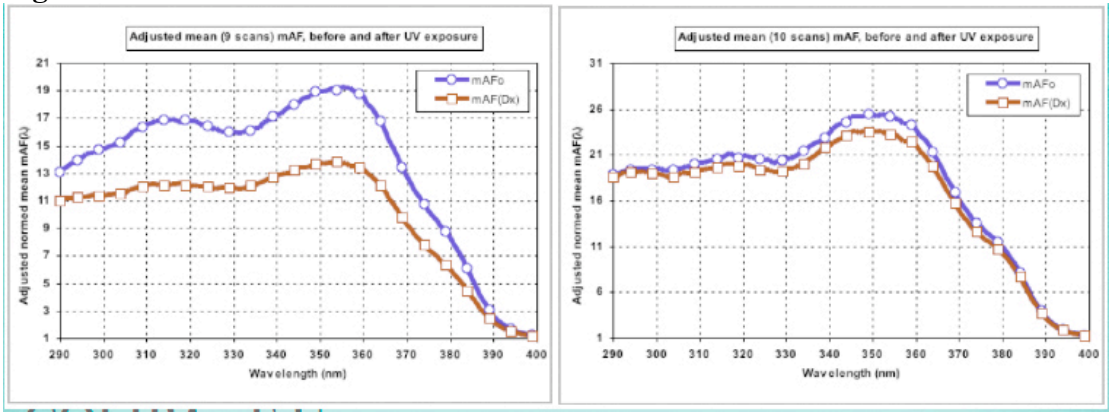
Prevention Testing
For practical purposes, establishing efficacy for prevention of further skin damage is restricted purely to measurement of the UV protective attributes of the product. Fortunately, these claims are now well supportable by recognised test methods. Both in-vitro and in-vivo measurements can be conducted and these can cover SPF and UVA broad spectrum testing. For SPF and UVA in-vivo, monographs exist for the major markets – E.U. , U.S.A., Japan, Korea and China. As well, corresponding ISO methods will be published around July 2010.
A ten subject study is accepted for all of these protocols and the test is a single point. The test is performed by exposing small areas of skin to simulated sunlight, filtered so as to remove the UVB wavelengths. The end point is Persistent Pigment Darkening (PPD) and this can be reported as a scalar range - PA +. PA++ etc. [4]

For in-vitro UVA, JCIA[4], COLIPA[5], Boots[6] in the U.K. and the Australia Standard [7] all have specific protocols. Testing to any of these will provide recognised evidence . Once again, ISO will shortly provide a UVA in-vitro UVAPF method, possibly by late 2010. The majority of these methods involve the preparation of a thin film of product to a transparent substrate, drying down of the film and measurement of light absorption over the UVA and UVB range.

An addition worthy of consideration is photo-stability. In-vitro testing of photostability can be performed for a very low cost, using thin film measurement methods similar to UV in-vitro measurements. Photostability is very important for

products intended to be applied in the morning and used on the exposed areas of the skin (typical for age spot treatment formulations.

Fig 6 Unstable vis Stable Sunscreen.

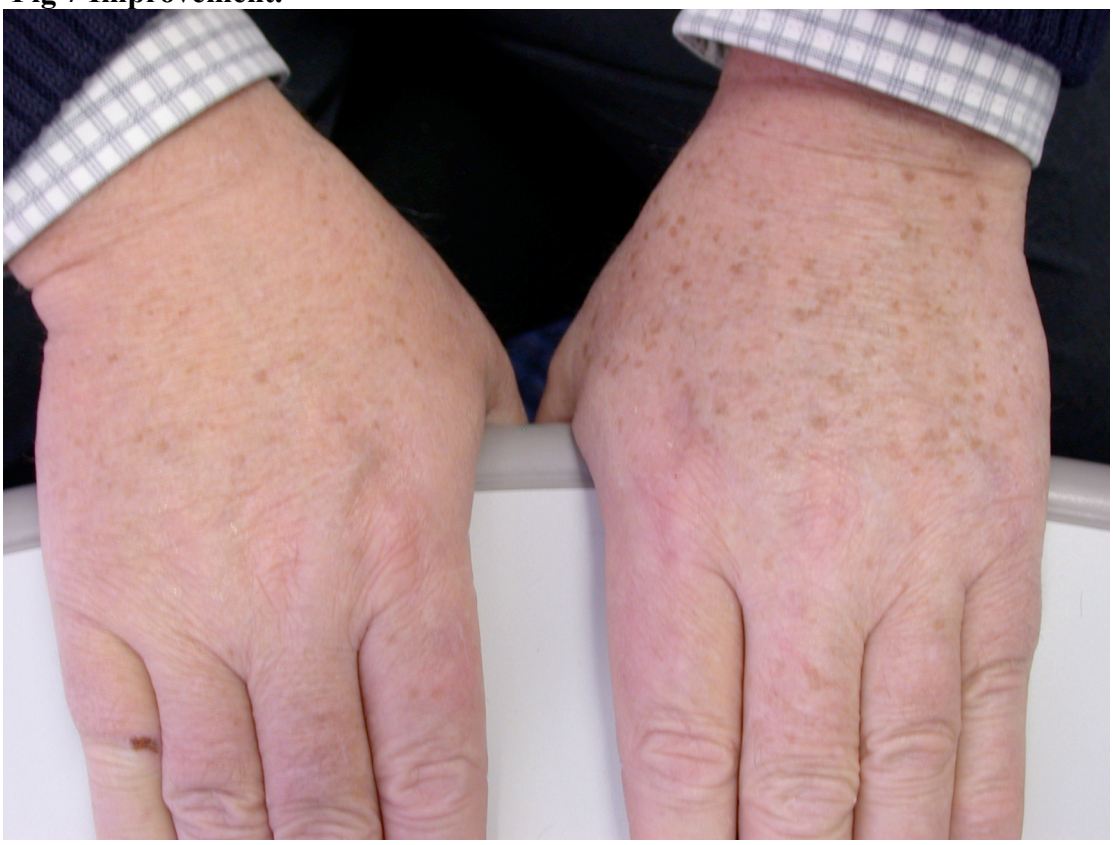


Summary

A product formulated for use as an effective skin lightening preparation should incorporate ingredients which act at both the treatment and the prevention level. Appropriate claim substantiation testing should be performed.

Whether performed on the product itself, or extrapolated from ingredient testing, it is important to ensure that the marketing claims reflect the results of testing. The final endpoint will still be the user's expectation of visual improvement.

Fig 7 Improvement.



(1) *Google* Search June 2009 "skin whitening" , "skin lightening" , "skin brightening".

(2) Jemec G.B.E., Serup J. Handbook of Non-Invasive Methods

(3) Nava D. Skin Ageing Handbook :An Integrated Approach to Biochemisty and Product Development William Andrew 2008

(4) Japan Cosmetic Industry Association – Measurement Standards for UVA Protection Efficacy (1999).

(5) Method for the *in-vitro* Determination of UVA Protection Provided by Sunscreen Products Guideline 2007 – COLIPA

(6) Measurement of UVA:UVB Ratios according to the Boots Star Rating System (2008 Revision).

(7) AS/NZS 2604 1998.

30th June 2009

.....