

# The Face Thymometer

## -A subjective measure of feeling.

Dr. Michael F. D'Souza MD FRCGP FFPHM FRSA  
Former Head of Unit of General Practice, NHLI, Royal Brompton Hospital UK

### Abstract:

The Face Thymometer is a simple scale for grading personal feelings. It consists of a series of emoticons with geometrically proportionate expressions that can be used to subjectively score a wide range of issues such as the quality of Life and the usefulness of treatments. This paper retrospectively reviews over 12,000 instances of its use. It examines how valid such a simple and brief scale is for measuring wellbeing and compares its performance with other more established instruments.

Face Thymometry has “face validity” and in many of the formal studies it was the least omitted data item suggesting it's meaning is fully understood. However, a few patients recorded smiles or frowns that were exaggerated when compared with what was known about them clinically. Interestingly many of these had known personality difficulties.

Face Thymometry scores were found to correlate significantly and positively with the gold standard EuroQol 5D when measuring Quality of Life and they also showed a strongly positive and significant correlation with the Zung depression scale ( $r_s = 0.7716$   $p < 0.00001$ ) however their correlation with McGill pain scores was less consistent. Variation in Hay fever symptoms, recorded by Thymometry, correlated significantly with ambient pollen counts, thus corroborating the biological validity of the scale.

In addition, this Faces scale works independently of language and culture and its brevity allowed it to be administered via a touch screen auto-receptionist. Thus attending patients may use it to routinely record their wellbeing and enable quality of life to be made before and after interventions. In this way Thymometry might be of use in obtaining *subjective outcome* information alongside Mortality, Morbidity and Service usage data. That Thymometry is so readily comprehensible to the general public might enable it to play a persuasive role in outcome-based commissioning.

### Introduction

Standardized, practical ways of measuring subjective outcomes are needed by Patients, Clinicians, Researchers, and Healthcare managers. Patients often like to know how well the treatments they are offered have symptomatically suited fellow sufferers. Clinicians generally feel this sort of subjective data is useful when advising on therapeutic choices. (Indeed because it is increasingly common for medical interventions to be given specifically to improve quality of life, having an easily applicable and understandable measure will help them prescribe with more confidence.) Furthermore those engaged in pragmatic health service research have a special need for tools that can be applied within the natural setting without disrupting normal clinical practice. Finally, Healthcare managers who have the task of allocating resources across a wide range of clinical interventions require patient-centered outcomes that are not specific to disease categories.

Unfortunately, most of our existing tools for measuring subjective outcomes have intrinsic credibility problems and are often too verbal, specialized and detailed. They have been designed for scientific explanatory studies where specialist statistical assistance is required<sup>1</sup>.

Hyland<sup>2</sup> discussed the selection of generic quality of life tools according to whether they were being used in a research or a clinical context and Ruta and colleagues<sup>3</sup> have suggested six features that the ideal clinical tool should possess:

1. Measure the aspects and effects of the illness that the **patient** decides are most important;
2. Enable the **patient** to score the chosen variables;

3. Be a sensitive measure of within person change over time;
4. Be applicable to the whole spectrum of illness seen in primary care;
5. Be capable of measuring the effects of a wide variety of care;
6. Be brief and simple enough to complete in a 7 to 10-minute consultation.

The COOP-WONCA charts<sup>4</sup> and the Andrews & Withey D-T scale<sup>5</sup> go some way in meeting these requirements and have generated a wide bibliography supporting their cross-cultural use. However, they have not become widely adopted in the clinical context, so it seems worthwhile to try to find alternatives that are more generally acceptable and might become standardized.

## Method:

The aim of this work is to examine our previous, often unpublished, studies to evaluate the strengths and weaknesses of Face Thymometry as a generic outcomes tool.

## The Settings and Populations studied

Most of the studies reported here took place in 120 general practices in the Kingston & Richmond Health Authorities in London UK. The catchment area population was about a third of a million. The participating GPs were members of the "Kingston & Richmond Multifund" which in the mid 1990s was the originating model for today's Clinical Commissioning Groups (CCGs)

## The Subjects

The major part of the analyses reported here was made on 11,230 routine consultations by patients from the above practices in 1995<sup>6</sup>.

Kingston Upon Thames, in common with other parts of London, has a very racially diverse population. During the periods of data gathering (one to six months) no patients reported difficulty understanding the scale but some individual patients made more than one consultation. However, due to analytic limitations, this repeat consultation was not quantified at the time. However, the U.K. average annual consultation rate in the mid-nineties was 3.9 per year.<sup>7</sup> A more detailed contemporary analysis in the author's own practice was done looking at very high consulters, (defined as those attending 12 times or more each year). It was found that while these represented only some 4% of the practice list, they caused 20% of the consulting workload. A further examination of this high consulting 4%, revealed that half merited the extra attention because they had serious medical problems such as cancer but the other half had no clear diagnoses and included many who might be classified as "Heart sink". This may assist in interpreting the populations in these studies.

The other populations examined here were either trial subjects or citizens of the Royal Borough of Kingston selected for a population study.

## The instrument

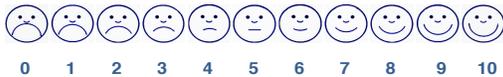
The Face Thymometer differs from the 5-point diagrammatic COOP-WONCA charts<sup>4</sup> in presenting a single eleven-point scale of geometrically precise faces. It retains the COOP-WONCA charts' independence from differences in culture, language and education but has greater discriminating power when used as a generic measure to compare the value of interventions in any symptomatic condition.

The scale starts with a full down-turned semi-circle representing a frown with a subjective score of **0** ("100%") i.e. "Maximum Unhappiness". In subsequent emoticons the arcs of the frown are reduced by intervals of 20% until it forms a straight line representing a subjective score of **5** ("0%") or

“neutral”. Thereafter it turns up into a smile increasing by the same 20% intervals until it becomes a fully upturned semi-circle representing a smile with a subjective score of **10** (“+100%”) or “Maximum Happiness” see **Fig 1**. Smaller increments of expression can be used but unless the scale is considerably enlarged any smaller differences are difficult to discriminate on touch-screens or paper questionnaires.

## The Thymometer

-A visual measure of feelings



Any question about how an individual feels may be inserted here  
The frowns and smiles vary by exact percentages

**Fig 1 The basic Thymometer design**

## Truncated Thymometer in App

The screenshot shows the BETME app interface. At the top, it displays the user's name 'BETME' and a 'Log out' button. Below this, the text 'In the last week...' is followed by a question: 'How happy/unhappy have you been?'. This question is accompanied by a truncated Thymometer scale with 11 faces, where the 5th face (score 5) is selected. Below this, another question asks 'How often have you felt moody?' with a dropdown menu set to 'Please select'. Further down, there are input fields for 'Weight' (set to 0 stone) and 'Cigarettes smoked' (set to 'Please select'). There are also dropdown menus for 'Alcohol units' and 'Spent on recreational drugs', both set to 'Please select'. At the bottom, there is a text input field for 'What have you done for yourself?'.

## Fig 1.

*Above is an example of how a truncated Face Thymometer is currently being used in a new App (2015) (<http://www.i-bet-me.com>.) to measure the effect of using Contingency Management to reduce harmful habits. Participants must try to earn more “CareCredits” each week. These are awarded for taking better care of their Communities, their families and themselves.*

Being a generic, rather than a condition-specific tool, the Face Thymometer requires users to *synthesize* how their current physical, psychological and social experiences are contributing to their wellbeing and to reflect this by their choice of a particular face.

However, unlike in metrics such as weight and height, individuals cannot measure their feelings by exact universally accepted intervals. Therefore, Thymometry and all similar scales behave intrinsically like “rubber rulers”. This situation is further complicated because a few subjects appear to rate their feelings in either hyperbolic or over-conservative ways. This will be discussed later.

## The study plan

There was no original study plan when these studies were started thirty years ago. The Thymometry scale was derived initially from the conventional 100mm analogue line being used by us in comparative studies of Hay Fever treatments<sup>13</sup>. We then compared it with some conventional

measures of symptomatology such as the Zung<sup>8</sup> Depression and the McGill Pain questionnaire.<sup>9</sup> However attempts to compare it with the best-known generic instrument the SF36<sup>10</sup> had to be abandoned because too few patients in these natural setting studies were prepared to fill in the very long SF36 forms provided. It is easy to agree with Jenkinson<sup>11</sup> "The SF-36 is not suitable for use within a consultation, which detracts from its clinical usefulness". Fortunately, EuroQol 5D has been validated against SF36<sup>12</sup> so the Euroquol has been employed as the comparison gold standard for Quality of Life. Examining old databases retrospectively has been challenging so occasionally additional specific studies have been done recently to fill in gaps.

### Using Thymometry to calculate a "Bentham Score"

Face Thymometry was used to assess concepts like "Treatment usefulness" and "Quality of Life". However, it was also used it to compare population differences; using a concept we are calling "the Absolute Bentham" after the great Utilitarian thinker Jeremy Bentham who argued for setting a target as "The greatest happiness for the greatest number."<sup>15</sup> An "Absolute Bentham" in terms of Thymometry is scored when every individual in a population records a score of 10 (100%) (i.e. as happy as possible). Obviously most populations contain individuals who have lower than absolute scores. If all these individual scores are summated and taken as a proportion of the "Absolute Bentham" we get a proportion that is amenable to non-parametric statistical testing. More importantly when graphically presented even non-statistically trained people can "see for themselves" the results of evaluations. See **Fig 5**. Such an approach can also be used to display trends visibly.

### Thymometry to measure Quality of Life and the "Usefulness" of interventions

In addition to these analytical issues, Thymometry has been applied to measure feelings about a wide range of verbal concepts as specified by appending different legends to the scale. Because using written English detracts from one of the scales virtues i.e. its cultural and literacy independence, legends are best delivered in a range of different written or spoken languages.

To capture the everyday concept of **Quality of Life** with Thymometry, we have used a touch screen in the waiting room administering, the simple social question,

#### **"How are you today?"**

*The word "today" was included in the legend as it was time specific and prompted a considered assessment. The vaguer "How are you?" could evoke a less considered polite reply.*

*We used this question in our trial of Antidepressants (1999) **Fig 4** and again in our study to compare Thymometry with EuroQol (2015) **Table 1** and in assessing what "face difference" patients thought "important".*

For comparing Thymometry with McGill Pain scores (See **Table 1**) we used the legend

#### **"Please tick the face which best shows how your pain is affecting your quality of life now"**

We specified the meaning of "Quality of life" (QoL) by reference to the WHO definition that Health e.g. used in surveys in Multifund General Practices **Fig 3**.

#### **"Health is social, mental and physical well-being. How happy are you with the way your health is affecting the quality of your life now?"**

We also developed the concept of treatment "**Usefulness**". This required each patient to record a Thymometry score for a treatment that synthesizes symptom relief, convenience of use and freedom from side effects viz.

**Overall, how useful\* has your prescription or therapy been so far?**

*\* Usefulness is a measure of symptom relief, convenience of use and freedom from side effects*

More recently (2011) Thymometry has been used in studies on over 50 severe heroin abusers. Seventeen of these were asked the baseline question

**“How happy have you been, in general, during the last month?”**

With the aid of modern digital delivery systems, it is now possible to include a language translation facility. Viz. our current (2012-) on-line **Confidential Enquiry** in the One Norbiton Community regeneration project. In this study we are also employing Thymometry to look at the baseline QoL for a urban regeneration project called “One Norbiton” 28 random adults in the area employing the question

**How happy or unhappy are you, in general, about living in Norbiton?**

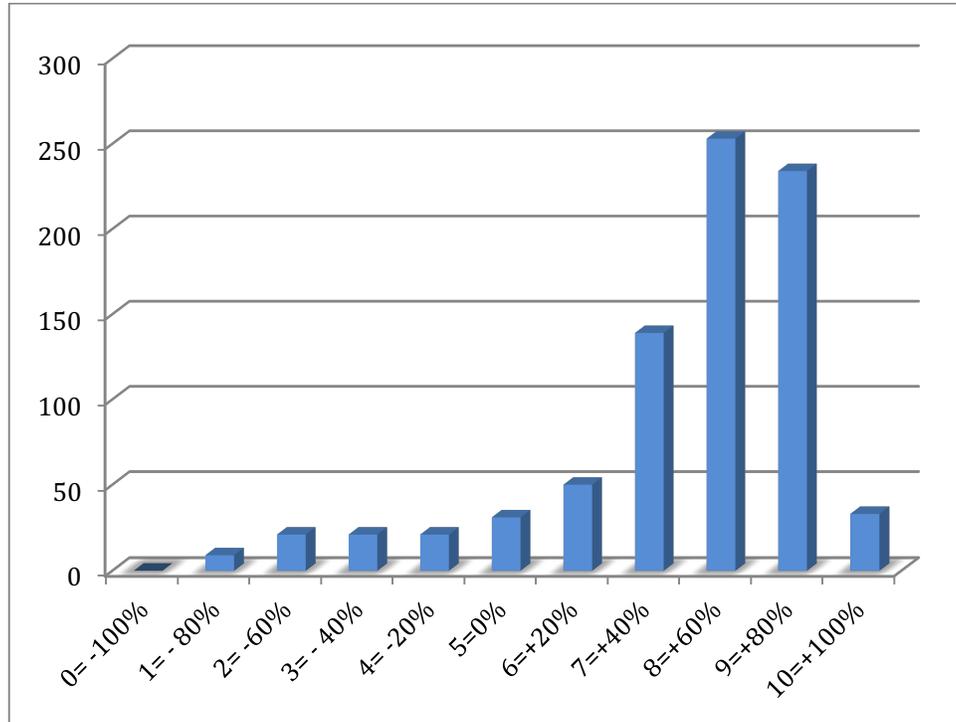
Finally we have used a truncated scale omitting the zero face in our current [App](#) (See **Fig 1a**) for a self-help service for addiction.

**Thymometry uses so far**

In summary Thymometry has been used with a wide range of questions on members of the general public, patients of all ages consulting GPs and on patients' family members and carers. Studies have included population surveys, Comparative trials of Hay fever treatments, Comparative studies of antibiotics for bronchitis, Validation studies against Euroqol, Studies on Depressed Patients, Specific intervention studies to support carers and finally in an on-line free App designed to help those with harmful habits. (See: [i-bet-me.com](http://i-bet-me.com))

## Results

**Fig 2 Population distribution of Face Thymometry**  
Agenda 21 on Quality of Life in Kingston May-Aug 2000



*This is a population survey done on 804 random citizens in Kingston Upon Thames U.K. (48/52 Male/Female, Age distribution: <18yrs 7.5%, 18-24 2.9%, 25-40 24.9%, 41-60 40.5%, >60 24.1%) The Question asked was: "Please give us your opinion about the Quality of Life in Kingston Where +100% = as happy as possible)". Most of this population was at least +60% Happy.*

All subjects studied appeared to understand the meaning of the Thymometry scale and how to use it. It was found to be both simple and brief enough to be applied in routine clinical care and its commonsensical and trans-cultural nature was an advantage.

Thymometry was also used in 1999 in Kingston's "Agenda 21 survey" in 1999. This was an international, UN inspired project, to measure what influenced their quality of life. (Fig 2) above. The faces scale was the least omitted data point (e.g. Age was omitted three times as often) The majority of this population rated themselves as Happy on the Thymometry Scale with a mode score of 8 (60%) and range 6 (20%) to (9.5) 90% in this fairly affluent part of the UK.

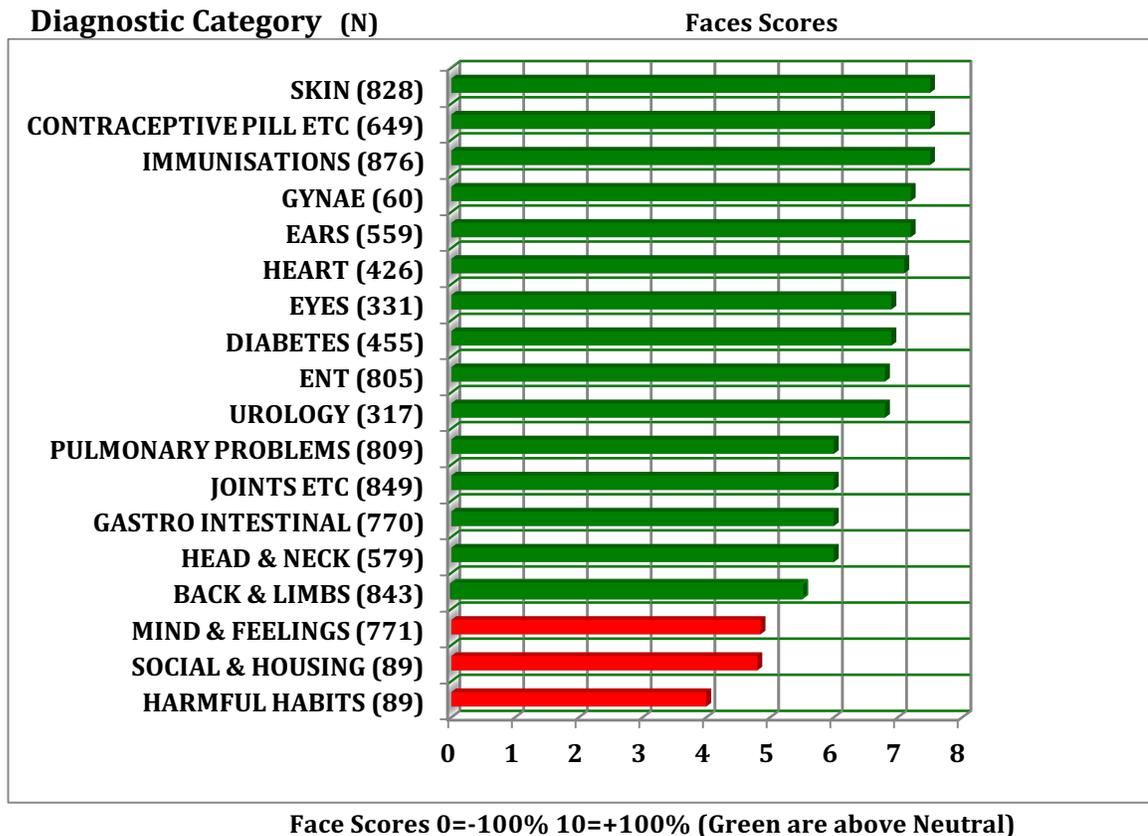
In early, 1995, touch-screen surveys were done on **GP Consultants**. Males recorded themselves as being slightly happier than Females and the younger (0-30yrs) and older (60+) age groups recorded higher happiness scores than those in middle age. These interesting observations showing a U-shaped happiness distribution has also been observed in more recent studies that employed many more measures<sup>14, 17</sup>; viz. The Gross National Happiness Index (33 different indicators). This further validates this much briefer tool.

### Surveys of 11,230 Consultations

Patients consulting their General Practitioners for any reason completed a brief questionnaire on the front of a mark-sense form on arrival and at subsequent visits to their surgery. GPs gave no prior guidance on how to complete the forms so the data consisted of only the patient's views. The Doctor or practice nurse completed the back of the form recording Diagnostic group and Treatment. See **Fig. 3**.

Follow up was not formalized but occurred on return for routine care (when the patients previous scores were *not* made available to them).

**Fig 3. Face Scores by Primary Care Diagnoses**



#### Construct Validity:

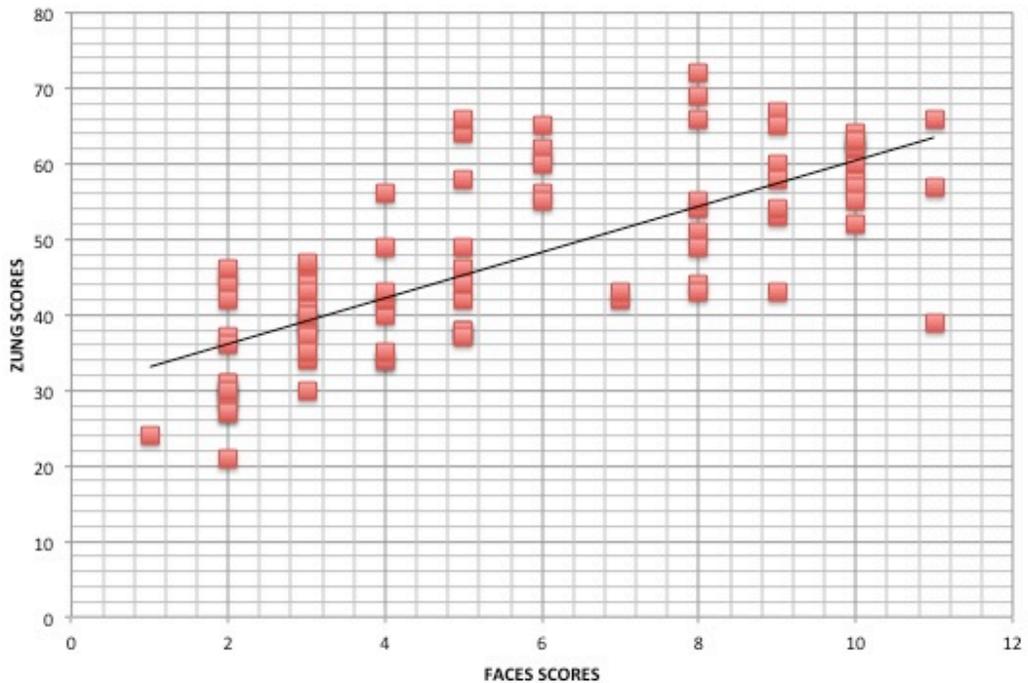
Not one of the subjects in our studies reported any difficulties in understanding what the faces meant. The Thymometry scale was usually completed in less than 10 seconds and its within-subject repeatability, after a 3-minute interval, averaged 99.7 %. Thymometry scores discriminated between this large group of GP patients in a logical and expected way; those attending for harmful habits recorded an average happiness score of only 3 (-40%) whereas in those attending for immunizations prior to going on holiday the average was 8.5 (+70%). See **Table 1**.

**Biological Validity:**

In comparative studies of Hay fever treatments, we found that changes in local pollen counts were reflected appropriately by changes in the analogue line scoring that developed into Thymometry.<sup>13</sup>

**Convergent Validity:**

This was shown for Depression where Zung scores correlated very well with Thymometry **Fig 4. below.** The change in Thymometry scores in depression at two weeks and at four weeks showed a consistent gradient across the spectrum of clinical change and this was supported by clinical observation. Additional support for Thymometry mirroring symptoms was shown in a study of painful conditions where McGill pain scores appropriately correlated with Thymometry, though correlation here was not as good as for depression, possibly due to the small numbers involved. See **Table 1**

**Fig. 4****Faces correlate well with Zung Depression scores**

**Zung scores in 84 patients in a Depression study in 1998 correlated significantly with Zung scores  $r_s = 0.7716$   $p < 0.00001$**

**Discriminant validity:**

Testing whether concepts or measurements that are supposed to be unrelated are, in fact, unrelated has not yet been formally checked for Thymometry.

**Sensitivity to Change:**

This is tested by appropriate responsiveness to situations. It was shown to occur in Depression studies, comparative trials of Hay Fever<sup>13</sup> and in Antibiotic studies where daily improvements in clinical condition were reflected by Thymometry scores. In our 1998 comparative trials of Antidepressant therapy involving 162 patients being randomized to receive either; Folic Acid (as a control), "Edronax", "Zispin", "Prozac", "Fluanxol", "Effexor", "Dutonin" or Counseling **Table 1**. The average starting Thymometry scores ranged from 3 (-40%) to 5 (0%). The outcomes showed that most patients showed improvement over three months with average scores rising to a range of 5 (0%) to 6 (+20%). This accorded with their observed improved clinical status but also confirmed that complete recovery from depression takes a long time.

**Concordance with Gold Standard**

When used to measure Quality of Life, Thymometry correlated highly significantly with the briefest currently used gold standard EuroQol 5D ( $r_s=0.7423$   $p=0.00005$ ) See **Table 1 below**

**Table 1 Some other studies using Face Thymometry**

| Study Topic        | Comparator   | Population                          | Result  | Reference           |
|--------------------|--------------|-------------------------------------|---|---------------------|
| 1. Pain            | McGill Pain  | 14 patients with painful conditions | $r_s = 0.392$<br>$p < 0.08$   | No prior publicatio |
| 2. Stress          | Counseling   | 24 relatives of substance abusers   | Significant and self-evident result using Thymometry                      | No prior publicatio |
| 3. Hay Fever       | Pollen Count | 114 Hay fever Sufferers             | Av. daily pollen count showed linear regression with patient diary scores | Reference (13)      |
| 4. Quality of Life | EuroQol 5D   | 50 GP patients                      | $r_s=0.7423$<br>$p < 0.00005$   | No prior publicatio |

*Face Thymometry scores correlated well with most other symptomatic measures.*

Analogue line tools rely on intellectual conceptualizations such as attributing the length of a line to the meaning appended verbal captions. Thymometry however uses common facial expressions to mirror subjective feelings.

Its eleven-point scale has been found to be more discriminatory than a binary or three point scale. On the other hand, it is uncertain whether increasing the number of points on such a faces scale would be even more advantageous. However, when a 21-point scale was placed under the faces it did not add any measurable discriminatory power. Altering size of the smiles might help e.g. by creating a 100-point scale. However very small differences in smile are difficult to distinguish visually in normal displays.

### **Importance:**

Significance and the likelihood of observations being real are routinely calculated, however assessment of "importance" is too often ignored. This is essentially a subjective view on what degree of change matters. This is critical when it comes to commissioning by outcome. Using our eleven-point faces scale, we asked 53 consecutive patients (Males 21, Females 32) "What is the smallest change in this faces score that you would consider important? E.g. indicate that you were unhappy to continue with your present treatment. The mode was "2 faces" the next highest was "1 face". Clearly the degree of freedom to move in either direction was altered as the extremities of the scale were reached. Further work is needed on this aspect.

### **Responsiveness to clinical change in Antibiotic Studies**

Thymometry seems to reflect a limited but adequate measure of subjective feelings associated with clinical change over time. In unpublished comparison studies of the symptom response of respiratory infections when treated with five different Antibiotics (Amoxicillin, Amoxil, Augmentin, Distachlor and Erythromycin) revealed similar regular improvements in daily Thymometry scores across the board. This matched the observed clinical recovery.

### **Responsiveness to clinical change in Addiction Studies**

Changes were observed over months in studies treating severe class A addicts with a contingency management regime; the mode for their initial Faces Score was low at 4 (-20%). This was in accord with their distressed clinical state at presentation and only 29% (5/17) scored above 0 (neutral). - This was in marked contrast with concurrent Norbiton community studies where the mode score was 6 (+20) with 83% (29/36) scoring higher than neutral. The contingency management regime we used awarded patients vouchers (CareCredits) if they took better care of themselves and or their friends and families. Earning [CareCredits](#) not only kept these notoriously nomadic patients under our care but, after nine months, 62.5% (15/24) recorded Thymometry scored above neutral. This was reflected in their observed unexpected and encouraging clinical improvement.

### **Mismatched Scores**

Of additional interest has been the observation that, particularly in our Depression and Addiction studies, a few individuals score themselves as extremely happy or extremely unhappy on the Faces scale despite this not being supported by our general clinical assessments. When these cases were

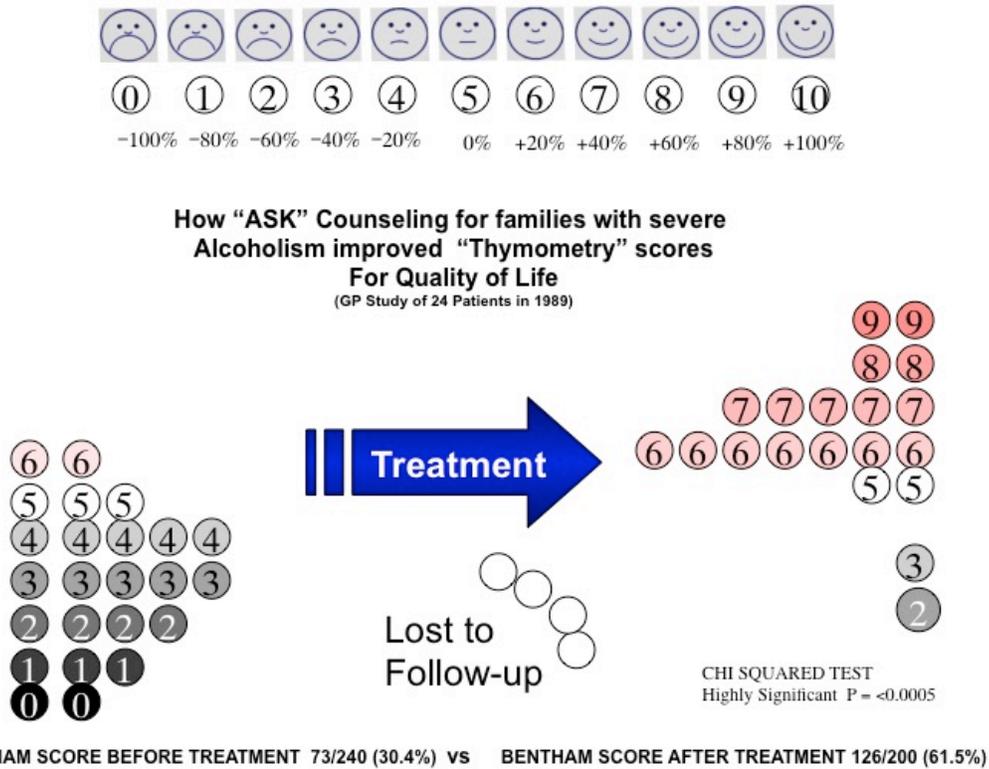
examined more closely many were often found to have personality difficulties. So rather than view this as a methodological weakness in the faces scale we realized that it might be revealing something of clinical use. E.g. Patients who score themselves as +100% happy despite being clinically depressed might be suffering from masked depression and require pharmaceutical intervention. While when the situation is reversed and patients record themselves as -100% unhappy yet have no evidence of depression, this might indicate personality disorder where drug treatment is best avoided.

**“Bentham Scores” and statistical testing with Thymometry**

Because of its subjective nature, Thymometry has important limitations as a metric. It is not a ratio scale and averaging is only of limited descriptive value and it is inappropriate to use T-test analyses. It may however be legitimately used to compare populations and the larger the group the better.

We have developed the concept of calculating a “Bentham score” for populations as one way of doing this. A “Bentham score” is the sum of every individuals happiness scores taken as a proportion of a notional “perfect score” in which every individual scores a maximum on the Thymometry scale i.e. exhibits the “greatest good for the greatest number of people.”<sup>15</sup> Then a simple non-parametric Chi-squared test can legitimately be used to estimate significance e.g. between the Bentham scores of population samples or between pre and post intervention scores in the same population. See **Fig 5** below.

**Fig. 5 An Evaluation of Counseling using Face Thymometry**



*The doubling of the “Bentham” scores in the relatives of severe alcoholics who had received counseling persuaded Kingston Health Authority to commission this service.*

*The low initial Quality of life in these relatives complemented the finding in Fig 3 that consulters with "harmful habits" score the lowest Quality of Life of all patients.*

*The doubling in the "Bentham" scores in the relatives of severe alcoholics after receiving counseling was convincing enough to persuade Kingston Health Authority to fund this service. The low initial Quality of life in these relatives matched the finding in Fig 3 that consulters with "harmful habits" seem to score the lowest Quality of Life of all patients.*

## Discussion

The term "Thymometry" was first used in 1967 in pain measurement<sup>16</sup>, however in this paper it is used generically as a "measurement of feelings".

Most studies reported here were done on GP consulters and cover quite diverse topics such as Quality of Life, Hay Fever, Pain, Depression, Respiratory infection, Counseling, Stress, Drug abuse and the Usefulness of Treatments. They exemplify *Clinical Epidemiology* where the investigator often personally knew the members of the groups studied. This has meant that direct clinical knowledge could refute or corroborate any conclusions and the delay in their publication has enabled a deeper reflection on the value of this simple instrument.

Like all such scales, Face Thymometry has been found to have both pros and cons. Most diseases, either immediately or eventually, generate a degree of unpleasant symptoms relevant to their severity. The use of diagrammatic facial expressions to self-grade severity and create a baseline and post-intervention scores seems an uncomplicated concept. However, whenever a person chooses what "Face" to record, they are making a synthesis of a wide range of factors. This synthesis might not always be appropriate e.g. when there are serious conditions, such as hypertension and malignancy that can be deceptively symptomless.

These studies suggest that Face Thymometer is a generally reliable measure. It has Construct, Convergent and Biological validity. It seems sensitive to short and long-term clinical change and to correlate well with the gold standard of Euroqol 5D as a measure of Quality of Life. In terms of appending a legend to the scale to measure Quality of Life the social enquiry "How are you today?" is preferred. A shift of one to two faces was considered "important" by patients in one of our studies, supporting the decision to create no more than an 11-point scale.

However we have also found that sometimes caution is required when interpreting the meaning of the faces scale, particularly in subjects with personality difficulties. Also, because of its very nature, it is statistically illegitimate to use T-tests to compare test averages of Thymometry data; instead the concept of comparing populations non-parametrically, by calculating "Bentham scores" is advised.

There is growing evidence that it would be invaluable to share existing databases from both clinical, social and voluntary services particularly to better understand complex individuals who may be blocking beds in hospitals or causing "heart-sink" in the community. Having a single metric like Thymometry that can be easily incorporated in the IT systems of all the caring agencies could greatly stimulate collaborative care and the interconnection of information systems.

Changes related to a whole variety of interventions or combinations of interventions could then be analyzed. This could complement RCTs or substitute for them when they cannot be done due to high costs or the ethical problems of giving sick people placebos etc. The development of such a unified system could facilitate the *ongoing* evaluation of new intervention policies such as merging NHS and Social service budgets and switching care from Hospital to the Community.

In conclusion, although there are many more detailed instruments available to measure disease-specific outcomes<sup>18</sup>, Face Thymometry despite its simplicity seems to be a valid *generic* tool. It is brief, transcultural and understandable even by the illiterate. It has been successfully used for thirty years to record wellbeing before and after interventions, including via unsupervised touch-screen auto-receptionists.

It can be argued that outcome-based commissioning might be best established when routine objective measurements of Mortality, Morbidity and Service usage data are augmented by the routine collection of validated subjective measures. However quantifying quality will always be a hazardous endeavor. As Bernard Shaw once remarked "Do *not* do unto others as you would that they should do unto you. Their tastes may not be the same." However, having a simple, tried and tested metric such as Face Thymometry that can continuously picture whether the greatest good is being delivered to the greatest number of people *in the recipients own opinion*, might be a helpful shared target, for most of our public services to aim at.

**Acknowledgements:** The author is grateful to Mr. David Gifford who drew the Face Thymometer to his specification and for the collaboration of the patients and General practitioners in Kingston upon Thames & Richmond. I am also grateful to Lilly and Upjohn for their financial support in the Depression studies.

**Conflicts of interest:** None.

**Limitations on use:** Face Thymometry's copyright is owned by TORCHE Clinical Consultancy Ltd. It is usually made freely available to non-profit making agencies and pro bono researchers. Contact [mdsouza@mac.com](mailto:mdsouza@mac.com).

## REFERENCES

1. **Schwartz D, Lellouch J.** Explanatory and pragmatic attitudes in therapeutic trials. *J Chronic Dis.* 1967 Aug; 20(8): 637-48. Republished in *J Clinical Epidemiology*
2. **Hyland ME.** "A brief guide to the selection of quality of life instrument" *Health and Quality of Life Outcomes* 2003, 1:24
3. **Ruta DA, Garratt AM, Leng M, Russell IT.** A new approach to quality of life: the patient-generated index. *Medical Care* 1994; 32: 1109-26.
4. **Kinnersley P, Peters T, Stott N.** Measuring functional health status in primary care using the COOP-WONCA charts: acceptability, range of scores, construct validity, reliability, and sensitivity to change. *Br J Gen. Practice.* 1994; 44:545-9
5. **Andrews FM, & Withey SB.** Social Indicators of wellbeing: *Americans perceptions of Life Quality.* New York, 19 Plenum Press
6. **D'Souza MF.** The Multifund and Outcome research *International Journal of Epidemiology* 1995 Vol. 24 No 3 (Suppl.1) S113 -S118
7. **Hippisley-Cox J, Vinogradova Y.** Trends in consultation rates in General Practice 1995/1996 to 2008/2009: Analysis of the QResearch® database. Final report to the NHS Information Centre and Department of health. Leeds: The NHS Information Centre for health and social care; 2009. Available from: <http://www.hscic.gov.uk/catalogue/PUB01077/tren-cons-rate-gene-prac-95-0...>
8. **Zung WK.** "A self-rating depression scale" *Archives of General Psychiatry* 1965 12: 63-70.
9. **Melzack R.** The short-form McGill Pain Questionnaire. *Pain.* 1987 Aug 30(2): 191-7.
10. **Garratt A, Ruta D, Abdulla MI, Buckingham JK, Russell IT.** The SF-36 health survey questionnaire: an outcome measure suitable for routine use in the NHS? *BMJ* 1993; 306:1440-4.
11. **Jenkinson, C.** MYMOP, a patient generated measure of outcomes. *BMJ* 1996 313: 626-626
12. **Luo, N., Chew, LH., Fong, KY., et al.** A comparison of the EuroQoL-5D and the Health Utilities Index mark 3 in patients with rheumatic disease. *Journal of Rheumatology*, 2003 30, 2268-2274.

13. **D'Souza MF, Tooley M, Charlton JRH.** Hay fever Treatments - Which should be Tried First? *J.Royal College of General Practitioners* 1987; 37: p.296-30.

14. **Helliwell J, Layard R. and Sachs J.** 2011: **World Happiness Report**

15. **Bentham J.** An introduction to the principles of morals and legislation 1789

16. **Peck. RE** Headache: 1967 The Journal of Head and Face Pain - Wiley Online Library

17. **Kahneman D and Krueger AB** Developments in the Measurement of Subjective Well-Being. *Journal of Economic Perspectives* 2006; 20 (1): 3-24

18. **Bowling A,** Measuring Health 1992 Open University Press ISBN 0-335-15436 Murphy E, Dingwall R, Greatbatch D, et al. Qualitative research methods in health technology assessment: a review of the literature. *Health Technol Assess* 1998; 2(16): 1-13.