Turning the page on paper-based assessments:

Three techniques and one technology to transform patient-reported assessments

Chris Gibbons

Director of Health Assessment and Innovation
The Psychometrics Centre

NIHR Research Fellow
Cambridge Centre for Health Services Research
University of Cambridge
TODAY’S TALK

Describe

• Three cutting-edge techniques and one open-source technology to improve patient reported outcome measures

Demonstrate

• How modern psychometrics can transform patient outcome and experience assessment

Discuss

• Where we might be heading in this world of big data and computational social science?
THE TERMS

• Patient-reported outcome and experience measures
• Psychometrics
• Item response theory
• Item banks
• eHealth
• Machine learning
• Computer adaptive testing
• Computational social science
• Item response theory
• Item banks
• eHealth
• Machine learning
• Computer adaptive testing
• Computational social science
ePsychometrics

- Item response theory
- Item banks
- eHealth
- Machine learning
- Computer adaptive testing
- Computational social science
PROMS AND PREMS

• Thousands, suitable for use in diverse conditions

• Many applications (trials, clinical practice, epidemiology, audit, quality, improvement)

• Patient-centred

• Ad hoc use (motivated clinicians)

• Take a long time to administer, score and interpret
• Scores can be difficult to understand
• Not much incentive
PROMS EVIDENCE

• Systematic reviews (e.g., Valderas et al., 2008)
• Effect of PROM feedback to clinicians on processes and outcomes of care.

• 2/3rd of studies show impact on processes
• Half show impact on outcomes

• Heterogeneity (of course)
WHAT SEEMS TO WORK?

- Specific PROM training/supported implementation
- Timely feedback
- Well aligned with practice
- Information integrated into available systems
- Information that is linked to specific clinical action
- Interventions with a formal, structured, feedback process perform better (Krageloh, 2014)
WHAT DO PATIENTS THINK?

- **PROMS** for Depression in primary care in the UK
- Patients favoured the PROMS
- They saw them as an efficient and structured supplement to medical judgement
- They saw them as evidence that clinicians were taking their problems seriously through full assessment
Routine provision of information on patient-reported outcome measures to healthcare providers and patients in clinical practice (Protocol)

How good is your GP at taking your problems seriously?

<table>
<thead>
<tr>
<th></th>
<th>Very poor</th>
<th>Poor</th>
<th>Neither poor nor good</th>
<th>Good</th>
<th>Very good</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage</td>
<td>9.1%</td>
<td>21.2%</td>
<td>69.7%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
How good is your GP at taking your problems seriously?

<table>
<thead>
<tr>
<th>GP Patient Survey</th>
<th>Very poor</th>
<th>Poor</th>
<th>Neither poor nor good</th>
<th>Good</th>
<th>Very good</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>9.1%</td>
<td>21.2%</td>
<td></td>
<td>69.7%</td>
<td></td>
</tr>
</tbody>
</table>
People who said “Good” or “Very good”

“Taking your problems seriously? Well no he didn’t.”

“No, he didn’t listen to me.”

“Well no, he didn’t really ask about symptoms.”
IMPROVING MEASUREMENT IN CLINICAL PRACTICE

• Tackle the barriers!

• Make measurement more efficient

• Engage patients and clinicians with instant feedback

• Align feedback more closely with care

• Address response biases
IMPROVING MEASUREMENT IN CLINICAL PRACTICE

• Tackle the barriers!
• Make measurement more efficient
• Engage patients and clinicians with instant feedback
• Align feedback more closely with care
• Address response biases

Hypothesis: All possible using ePsychometrics
IMPROVING MEASUREMENT IN CLINICAL RESEARCH

ePsychometrics can also transform..

• Clinical trials (recruitment, reliability, attrition, comparability)
• Comparative research
• Patient-centred Big Data analytics
WHAT IS ‘PSYCHOMETRICS’?

- Item response theory
- Classical test theory
The Psychometrics Centre
Psychometrics

Science
Maths
Technology

Psychometrics

Implementation
Assessment
Feedback
The Psychometrics Centre

Research, leadership, innovation, support

Science  
Maths  
Technology

Implementation  
Assessment  
Feedback

Psychometrics
THREE TECHNIQUES & ONE TECHNOLOGY

- Computer adaptive testing
  (Efficiency and precision)
- Prediction and feedback
  (Validation and engagement)
- Machine learning
  (New forms of data)
DEVELOPING A PREDICTIVE COMPUTER ADAPTIVE TEST FOR THE WHOQOL QUALITY OF LIFE MEASURE
ITEM RESPONSE THEORY

- Probabilistic psychometric theory
- The more you have the more you agree...
- Sample free and test free
- Individual reliability
- Permits computer adaptive testing
- Allows predictions
COMPUTER ADAPTIVE TESTING

• “Computer system which iteratively ‘learns’ about the test taker and selects the best item from a large ‘bank’ of items”

• Interfaces with item response theory

• Could be efficient, more reliable and better targeted

• Increasing use

• Used in international high stakes educational assessments (USA, UK and Australia)
COMPUTER ADAPTIVE TESTING

Item Bank

1  2  3  4  5  6  7  8

Computer adaptive test

Participant
COMPUTER ADAPTIVE TESTING

Item Bank

1  2  3  4  6  5  7  8

Computer adaptive test

Participant

Best item
COMPUTER ADAPTIVE TESTING

Item Bank

2
3
4
5
6
7
8

Computer adaptive test

Participant

Item
COMPUTER ADAPTIVE TESTING

Item bank

Participant

Computer adaptive test

Response
Quality of life estimate = 22
COMPUTER ADAPTIVE TESTING

Item Bank

2
3
4
5
6
7
8

Computer adaptive test

Participant

Best item

Quality of life estimate = 22

Reliability

70 .80 .90
COMPUTER ADAPTIVE TESTING

Item Bank

2
3
4
5
7
8

Computer adaptive test

Quality of life estimate = 47

Participant

Item

Response

Reliability

70
.80
.90
Computer Adaptive Testing

Item Bank

2
3
4
5
7

Computer adaptive test

Participant

1
6
8

Quality of life estimate = 55

Reliability

70 .80 .90
COMPUTER ADAPTIVE TESTING

Item Bank

2
3
5
7

Computer adaptive test

Quality of life estimate = 60

Participant

Response

Item

Reliability

70 .80 .90
COMPUTER ADAPTIVE TESTING
The WHOQOL

- Multi-dimensional measure of global quality of life
- Physical, psychological, social and environmental domains
- Fifteen field centers
• IRT analysis of the WHOQOL-100 on UK sample (n = 320)
• Four domains fit the Partial Credit Model (p>0.5)
• 52 items removed
• Mean 11 items per bank
• Banks were suitable for patients with long-term conditions
WHOQOL can be 82% shorter and more reliable

(Gibbons et al., In Press. J Med Int Res)

WHOQOL-CAT assessments are accurate and comparable between countries

(Gibbons et al., 2015 Qual Life Res)
Computer adaptive test

Personalised Quality of Life (QoL) feedback
Unique Identifier - S5QHJ

<table>
<thead>
<tr>
<th>Scale</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical QoL</td>
<td>60</td>
</tr>
<tr>
<td>Psychological QoL</td>
<td>71</td>
</tr>
<tr>
<td>Social QoL</td>
<td>53</td>
</tr>
<tr>
<td>Environmental QoL</td>
<td>76</td>
</tr>
</tbody>
</table>

The scores above are worked out from the answers that you gave to the questions you have just completed. Your results are given on the chart above. A higher score means that you have a higher quality of life. The table below gives you a little bit more information about what each of the scores mean.

**Physical Quality of Life**

Your physical quality of life includes things like how well you are able to move around, how much energy you have or how much you are in pain.

Your score of 60 on this scale indicates that your physical quality of life is normal. The majority of people in the United Kingdom report a similar quality of life to you.

If you are worried about your physical quality of life then you should contact your doctor. You can [click here to find a local doctor](http://example.com) or [click here to access local emergency services](http://example.com).

Adaptiveqol.com
PREDICATE STUDY

• Predictions in a Computer Adaptive Testing Environment

• Assess validity in the real-world

• CAT is completed until stopping rule is met

• An algorithm predicts what test-takers would answer to unseen questions

• Test-takers rate the predictions (Correct, close, wrong)

• PREDICTMYQOL.com
Here are the four questions and the answers we predict that you would have given, based on your previous responses.

<table>
<thead>
<tr>
<th>Question 1</th>
<th>Do you have enough energy for everyday life?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Answer</td>
<td>Moderately</td>
</tr>
</tbody>
</table>

We are 69% sure that we're right. What do you think?

- Exactly right, just what I was thinking!
- Close, but it's not quite right
- The prediction is wrong. I was thinking of a very different answer
PREDICATE Study

- 699 completions (52 countries)
- 173 participants from the UK
- Average age = 45 ± 10
- 56% male
- 50 patients from UK report long-term conditions (diabetes, depression, arthritis, chronic obstructive pulmonary disease, coronary heart disease)
For the UK sample, results were, on average, 94% “right” or “close”

‘Exactly right’ 69% of the time (predicted accuracy was 78%)

“Close, but not quite right” 25% of the time

“Wrong” 6% of the time [non-UK 10% wrong]

Prediction performance was better in the UK (where item bank is validated) than in the rest of the world
The item bank is brief and can be administered in 120 seconds.

As much as a 82% increase in efficiency compared to paper-based versions.

Simulations translated well to ‘real world’ CAT.

Prediction study supports the validity of the measurement model and ‘backfilling’ unanswered item bank items.

Feedback well received and instantaneous.

People like the predictions.
USING MACHINE LEARNING TO MAKE SENSE OF GP PERFORMANCE DATA
Many questionnaires include open-text elements to add further information.

May contain important information missed by questionnaires.

Typically underused / ignored.

Time-consuming to use human analysts.
PATIENT EXPERIENCE

We can make sense of this data using machine learning and natural language processing.
GMC Colleague Questionnaire

360 Degree Feedback

1636 Comments

548 Doctors

20-item questionnaire

Free-text information
MACHINE LEARNING
MACHINE LEARNING

Data
- Images
- Text
- Numbers

Prediction
- Classification
- Regression

“The algorithm”
MACHINE LEARNING

Colleague reports of GP performance

“The algorithm”

Machine classifications

‘Professional’
‘Innovative’
‘Respected’
‘Interpersonal’
‘Popular’
Colleague reports of GP performance (50% of comments)

Classifications made by human raters

“The algorithm”
VALIDATING

Colleague reports of GP performance

“The algorithm”

Machine classifications

Compared with

Classifications made by human raters
VALIDATING

Colleague reports of GP performance

"The algorithm"

Machine classifications
Compared with Classifications made by human raters
DEPLOYMENT

Colleague reports of GP performance

"The algorithm"

Machine classifications

‘Professional’
‘Innovative’
‘Respected’
‘Interpersonal’
‘Popular’
<table>
<thead>
<tr>
<th>Theme</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Innovator</td>
<td>“It is clear from the advice he gives that he is aware of [the] current good practice, is highly motivated, very practical and very much a team player. His advice, when working with consultant colleagues was respected, and he recognised where practice/primary care limitations were and yet looked for opportunities for change and improvement.”</td>
</tr>
<tr>
<td>Interpersonal</td>
<td>“She has an admirable level of commitment and enthusiasm for her patients and her work. She has been instrumental in promoting change and improvement in her department. She is a great asset to the department and the hospital.”</td>
</tr>
<tr>
<td>Popular</td>
<td>“She is a very good, committed colleague always keen to improve, very liked by her patients and highly valued by all who work with her.”</td>
</tr>
<tr>
<td>Professional</td>
<td>“I find this doctor to be very efficient, caring, honest and very professional.”</td>
</tr>
<tr>
<td>Respected</td>
<td>“I find that he very easy and helpful to work with, he always has time for patients and staff.”</td>
</tr>
</tbody>
</table>

“Very approachable and professional.”

“Excellent well liked and easy working colleague.”

“Very popular doctor. Works to high standards.”

“A first class colleague.”

“Pleasant and valued colleague.”
<table>
<thead>
<tr>
<th>Code</th>
<th>Agreement (Kappa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Innovator</td>
<td>.98</td>
</tr>
<tr>
<td>Interpersonal</td>
<td>.80</td>
</tr>
<tr>
<td>Popular</td>
<td>.97</td>
</tr>
<tr>
<td>Professional</td>
<td>.82</td>
</tr>
<tr>
<td>Respected</td>
<td>.87</td>
</tr>
</tbody>
</table>
T-tests comparing doctors with a rating in a category vs those with no ratings

<table>
<thead>
<tr>
<th>Code</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
</tr>
<tr>
<td>Innovator</td>
<td>48</td>
</tr>
<tr>
<td>Interpersonal</td>
<td>435</td>
</tr>
<tr>
<td>Popular</td>
<td>107</td>
</tr>
<tr>
<td>Professional</td>
<td>643</td>
</tr>
<tr>
<td>Respected</td>
<td>264</td>
</tr>
<tr>
<td>Any category</td>
<td>1069</td>
</tr>
</tbody>
</table>
T-tests comparing doctors with a rating in a category vs those with no ratings

<table>
<thead>
<tr>
<th>Code</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
</tr>
<tr>
<td>Innovator</td>
<td>48</td>
</tr>
<tr>
<td>Interpersonal</td>
<td>435</td>
</tr>
<tr>
<td>Popular</td>
<td>107</td>
</tr>
<tr>
<td>Professional</td>
<td>643</td>
</tr>
<tr>
<td>Respected</td>
<td>264</td>
</tr>
<tr>
<td>Any category</td>
<td>1069</td>
</tr>
</tbody>
</table>
### T-tests comparing doctors with a rating in a category vs those with no ratings

<table>
<thead>
<tr>
<th>Code</th>
<th>Performance</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>T-test</td>
<td>Sig.</td>
</tr>
<tr>
<td>Innovator</td>
<td>48</td>
<td>.00</td>
<td>.99</td>
</tr>
<tr>
<td>Interpersonal</td>
<td>435</td>
<td>1.98</td>
<td>.04</td>
</tr>
<tr>
<td>Popular</td>
<td>107</td>
<td>-.88</td>
<td>.38</td>
</tr>
<tr>
<td>Professional</td>
<td>643</td>
<td>2.51</td>
<td>.01</td>
</tr>
<tr>
<td>Respected</td>
<td>264</td>
<td>3.75</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Any category</td>
<td>1069</td>
<td>.77</td>
<td>.001</td>
</tr>
</tbody>
</table>
## PREDICTING PERFORMANCE

T-tests comparing doctors with a rating in a category vs those with no ratings

<table>
<thead>
<tr>
<th>Code</th>
<th>Performance</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>T-test</td>
<td>Sig.</td>
</tr>
<tr>
<td>Innovator</td>
<td>48</td>
<td>.00</td>
<td>.99</td>
</tr>
<tr>
<td>Interpersonal</td>
<td>435</td>
<td>1.98</td>
<td>.04</td>
</tr>
<tr>
<td>Popular</td>
<td>107</td>
<td>-.88</td>
<td>.38</td>
</tr>
<tr>
<td>Professional</td>
<td>643</td>
<td>2.51</td>
<td>.01</td>
</tr>
<tr>
<td>Respected</td>
<td>264</td>
<td>3.75</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Any category</td>
<td>1069</td>
<td>.77</td>
<td>.001</td>
</tr>
</tbody>
</table>
• Machine learning algorithms can classify open-text reports of doctors’ performance with human-level accuracy

• Machine-made classifications can signal significant differences in doctors’ performance
SUMMARY SO FAR

• Computer adaptive testing is an acceptable way to make assessments shorter and more reliable

• Machine learning may be used to make sense of open-text data in questionnaires
The problem
Concerto

Open source tool for the development of online assessments utilising computer adaptive testing and tailored feedback
The solution

Research, tools, training, and support
Vision

- Adaptive and electronic measurement should be widely available to non-experts
- It should always be free and open-source
- Assessment experience should be made as enjoyable as possible
- Data security protocol should be dictated by the test-developer (your servers, not ours)
- Where expertise is needed, it be should be readily available as consultancy, training, or free guides online
Deployment of open-source software

- Collaborators
- Patient and stakeholder feedback
- Case-specific adaptation

Brand recognition
Academic collaboration
Consultancy
Skills and learning
Paid support

Better products
More widely used
More patient benefit!
In practice
Flexibility

• **Item type** (text, pictures, sounds, movies)
• **Test type** (Computer adaptive test, short-form questionnaire)
• **Scoring** (IRT scoring or sum scores)
• **Feedback** (patient/clinician/graphs/text/links/e-mail)
• **Longitudinal assessment**
• ‘Log on’ details
• **Facebook connection**
• **Data security**
• **Data storage**
Concerto

Installable on different platforms

- Cloud servers (free trial on Amazon)
- Local server
- (at your University/hospital)
- Docker
- Windows 10 tablet
- Linux Installation

- User-defined data security
- Flexible speed and data-storage for scalable tests (up to millions of users)
PREDICTING PERSONALITY FROM DIGITAL FOOTPRINT
• Example of Big Data analytics in social science, using personality questionnaires

• Big Data widely used by Big Business

• Google, Facebook, Apple, Amazon, Barclays…

• “Inevitable application in health care…” (Murdoch & Detsky, 2013, JAMA)

• What might big data analytics using questionnaires or PROMs look like?
Psychometric tests (BIG-5) hosted on Facebook

Developed and managed by psychometrics centre members

More than 3,100,000 people have completed the Big-5 questionnaire

Large proportion of users shared their Facebook ‘Like’ information
Question:

Can you build a predictive model of a psychometric test (BIG-5) using Facebook ‘likes’
Do you like University of Birmingham? Yes | No
Private traits and attributes are predictable from digital records of human behavior

Michal Kosinski*,†, David Stillwell*,†, and Yaroslav E. Golosov*

Turner School, The Psychometric Society, University of Cambridge, Cambridge CB2 8PQ, United Kingdom;
†Microsoft Research, New York, New York 10019, United States.

We show that readily accessible digital records of behavior, Facebook data, can be used to automatically and accurately predict a range of personality traits (e.g., openness, conscientiousness, extraversion, agreeableness, and neuroticism) from a user’s Facebook profile data. The predictive power of these models is validated by cross-validation, leaving-one-out, and jackknife tests. Additionally, we find that these traits can be accurately predicted even when the traits are unknown. Finally, we show that these traits can be used to accurately predict a user’s age, gender, and political affiliation.
# MY PERSONALITY

**Private traits and attributes are predictable from digital records of human behavior**

Michai Keski**1,** David Stillwell**, and Thore Graepel**

*1EVE Online, The Psychometrics Centre, University of Cambridge, Cambridge CB2 8QZ United Kingdom; and \*Microsoft Research, Cambridge CB2 1JR, United Kingdom

Edited by Kenneth Wark, University of California, Berkeley, CA, and approved February 12, 2011 (received for review October 26, 2010)

We show that easily accessible digital records of behavior, Facebook Likes, can be used to automatically and accurately predict a range of existing personality traits (14). Similarly, it has been shown that personality can be predicted based on the contents of personal Web sites (14).

<table>
<thead>
<tr>
<th>Trait</th>
<th>Selected most predictive Likes</th>
</tr>
</thead>
<tbody>
<tr>
<td>IQ High</td>
<td>Jason Aldean, Tyler Perry, Sephora, Chiq, Bret Michaels, Clark Griswold, Bebe, I Love Being A Mom, Harley Davidson, Lady Antebellum</td>
</tr>
<tr>
<td>IQ Low</td>
<td></td>
</tr>
<tr>
<td>Satisfaction With Life Satisfied</td>
<td>Hawthorne Heights, Kickass, Atreyu (Metal Band), Lamb Of God, Gorillaz, Science, Quote Portal, Stewie Griffin, Killswitch Engage, Ipod</td>
</tr>
<tr>
<td>Satisfaction With Life Dissatisfied</td>
<td></td>
</tr>
</tbody>
</table>

**Curly Fries**
Private traits and attributes are predictable from digital records of human behavior

Michal Kosinski1,2,*, David Stillwell1,*, and Thore Graepel3
1School of Data Science, University of Cambridge, Cambridge CB2 8PB, United Kingdom; and 2Microsoft Research, Cambridge, CB2 1FB, United Kingdom

Edited by Kenneth Wark, University of California, Berkeley, CA, and approved February 12, 2013 (received for review October 26, 2012)

We show that easily accessible digital records of behavior, such as Facebook likes, can be used to automatically and accurately predict a range of psychological traits (11–15). Similarly, it has been shown that personality can be predicted based on the contents of personal Web sites (16).

<table>
<thead>
<tr>
<th>Trait</th>
<th>Selected most predictive Likes</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Godfather</td>
<td>Jason Aldean</td>
</tr>
<tr>
<td>Mozart</td>
<td></td>
</tr>
<tr>
<td>Thunderstorms</td>
<td></td>
</tr>
<tr>
<td>The Colbert Report</td>
<td></td>
</tr>
<tr>
<td>Freeman's Voice</td>
<td></td>
</tr>
<tr>
<td>The Scream</td>
<td></td>
</tr>
<tr>
<td>The King</td>
<td></td>
</tr>
<tr>
<td>The Noon</td>
<td></td>
</tr>
<tr>
<td>The Prodigy</td>
<td></td>
</tr>
<tr>
<td>Killswitch Engage</td>
<td></td>
</tr>
<tr>
<td>Being Conservative</td>
<td></td>
</tr>
<tr>
<td>Pride and Prejudice</td>
<td></td>
</tr>
<tr>
<td>Satisfaction with Jesus</td>
<td></td>
</tr>
<tr>
<td>Satisfaction with Christ</td>
<td></td>
</tr>
<tr>
<td>Satisfaction with God</td>
<td></td>
</tr>
<tr>
<td>Satisfaction with Christianity</td>
<td></td>
</tr>
<tr>
<td>Satisfaction with God</td>
<td></td>
</tr>
</tbody>
</table>
Computer-based personality judgments are more accurate than those made by humans

Wu Youyou1,2, Michal Kosinski3,4, and David Stillwell5

1Department of Psychology, University of Cambridge, Cambridge CB2 3EB, United Kingdom; and 2Department of Computer Science, Stanford University, Stanford, CA 94305

Edited by David Funder, University of California, Riverside, CA, and accepted by the Editorial Board December 2, 2014 (received for review September 28, 2014)

Accumal Accuracy (self-other agreement)

Number of Facebook Likes (log scaled)
• API which allows companies to make personality predictions at scale

• Hilton, Wrigleys, Barclays

• Using psychological information to tailor information improves conversion rate and satisfaction (Matz, 2015)
LEARNING POINTS

On feedback

• 3,100,000 people completed a personality test that had up to 336 items!
• No monetary incentive - just feedback
• People want to learn about themselves!
• Research planned – how does feedback affect response rates, reliability etc

On data sharing and collaborative / open science

• All these data (apart from Likes) are available for free to anyone (since 2012)
• No regrets!
WHY MIGHT THIS ‘INFERENTIAL’ APPROACH BE USEFUL FOR HEALTH SCIENCES

• Data exist on a massive scale (and increasingly so)
• ‘Learning data cities’ are being planned (cf. Buchan)
• May be less biased than questionnaire items
• Real-time monitoring/assessment/feedback
• No measurement latency (can be done anywhere)
WHAT MIGHT WE USE FOR ‘INFERENTIAL PSYCHOMETRICS’?

- Facebook likes
- Twitter ‘follows’
- Twitter updates
- Images (image recognition – e.g., Facebook profile photos)
- Instagram
- ‘Wearables’
- Geo-location data
- Movement data
- Phone activity
- Interactions with friends (proximity of two devices)
INTERNET OF THINGS

2020

4 BILLION
Connected People

$4 TRILLION
Revenue Opportunity

25+ MILLION
Apps

25+ BILLION
Embedded and Intelligent Systems

50 TRILLION
GBs of Data

Source: Mario Morales, IDC
The following questions are about activities you might do during a typical day. In the past 1-week does your health limit you in these activities? If so, how much?

(Please circle one number on each line)

<table>
<thead>
<tr>
<th>ACTIVITIES</th>
<th>Yes Limited A lot</th>
<th>Yes Limited A little</th>
<th>No, Not Limited At All</th>
</tr>
</thead>
<tbody>
<tr>
<td>3a: Vigorous activities, such as running, lifting heavy objects, participating in strenuous sports</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3b: Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3c: Lifting or carrying groceries</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3d: Climbing several flights of stairs</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3e: Climbing one flight of stairs</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3f: Bending, kneeling, or stooping</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3g: Walking more than one kilometre</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3h: Walking half a kilometre</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3i: Walking 100 metres</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3g ww: Wheeling more than one kilometre</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3h ww: Wheeling half a kilometre</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3i ww: Wheeling 100 metres</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3j: Bathing or dressing yourself</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

*Modified from SF-36*: Items 3 (a to j) are the original SF-36 questions, while 3g ww to 3i ww (shaded area) comprise the supplementary SF-36ww modification.
MENTAL HEALTH AND GOOGLE DATA?

- Do you have thoughts of suicide?

- Have you been experiencing delusions?
Former head of NIMH in the USA

Moved to Google in late 2015.

“Technology can have greater impact on mental healthcare than on the care for heart disease, diabetes, cancer or other diseases...

It could transform this area in the next five years.” (Insel, 2015)
IS USING DATA LIKE THIS A BIT CREEPY*

*OFF-PUTTING TO PATIENTS
Assessment

• 27 Yes/No Questions on aspects of Big Data
• Measuring attitudes towards Personalisation, Ethics, Wearable Tech, Internet of Things, Finance, Policy-Making, Cloud Storage, etc.
• 20-item BIG5 Personality test at the end

Feedback

• Participants given real-time feedback on the similarity of their answers to all previous participants
### Sample (by Oct ‘15)

- **Average age**: 30
- **No. of test sessions in total (Sept)**: 33,937
- **Answered At Least One Question**
  - All Big Data Questions: 19,126
  - All Personality Questions: 10,411
  - 43% Europe
  - 15% South America
  - 9% Asia
  - 27% North America

### Sample Demographics

- **Male**: 54%
- **Female**: 49%
- **Intersex**: 30

### Map of Test Session Locations

- 9% Asia

---

[Image of world map showing test session locations]
Sample (by Oct ‘15)

No. of test sessions in total (Sept)
33,937

Answered At Least One Question
19,126

All Big Data Questions
10,411

All Personality Questions
8,871

Average age
30

49% Female
54% Male

Male
Female
Intersex

No. of test sessions in total (Sept)
33,937

Answered At Least One Question
19,126

All Big Data Questions
10,411

All Personality Questions
8,871
### Results: Audit of Global Public Opinion

- **Are those with access to your personal data able to accurately predict your future behaviour?** *(n=19,100)*
  - 62% NO
  - 38% YES

- **Should predictive technologies be used to improve the quality of healthcare, for example by helping doctors to recommend personalised nutrition and exercise plans?** *(n=14,187)*
  - 16% NO
  - 84% YES

- **Should predictive technologies be used to assess your eligibility for a mortgage?** *(n=13,725)*
  - 62% NO
  - 38% YES

- **Do you think your organisation ought to invest in predictive technologies?** *(n=2,830 CMOs)*
  - 62% NO
  - 73% YES

- **Is it important for you to understand the psychological attributes of your customers?** *(n=2,489 CMOs)*
  - 8% NO
  - 92% YES
71% of patients were happy to share social media information with their doctor.
CONCLUSION

• Modern psychometrics offers many opportunities for improved research and clinical practice

• Computer adaptive tests work well in the ‘real world’

• Predictions are engaging

• myPersonality showed what will/could be possible for ‘inferential psychometrics’ in health research and the power of feedback

• The future is exceptionally exciting in this field!
THANK YOU!

cg598@cam.ac.uk
concertoplatform.com
drchrisgibbons.com

@DrCGibbons
@concerto_cam
@camppsych