The Exeter Consensus Statement:
A Delphi study to define ‘Treatment Resistant Depression’

David A Richards

Acknowledgements
• Authors: David A Richards, Jacqueline J Hill, Samuel M Waldron, Holly V R Sugg, Alison Bethel, Willem Kuyken
• Funding: this presentation reports independent research funded by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care (CLAHRC) for the South West Peninsula.
• The views expressed in this presentation are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health in England.
• Declaration of interests: all authors assert that they have no interests to declare

Treatment of Depression
• Treatment for depression includes pharmacological, psychological and physical interventions, mostly focussed on the acute episode.
• A significant proportion of patients (10 – 20%) do not demonstrate any response to these treatments*

‘Resistant Depression’
• Coined by the World Psychiatric Association in 1974 to describe patients who showed no response to tricyclic antidepressant treatment at a suggested dose of 150mg/day of imipramine or equivalent given over a period of four to six weeks*

The Current Situation
• “Comparison of any of the potential interventions in the field, nonpharmacologic or otherwise, is hampered by the variability in TRD definitions” (p162)*
• Reviewers found it impossible to include all trials of pharmacologic and non-pharmacologic interventions in a single meta-analysis because of the lack of “a standard definition of TRD that investigators should use in their clinical trials research” (p162)*

A recent example*
• “Eligible patients [with treatment resistant depression] were those aged 18–75 years who had adhered to an adequate dose of antidepressant medication (based on the British National Formulary and advice from psychopharmacology experts) for at least 6 weeks and had a Beck depression inventory (BDI) score of 14 or more. They also met international classification of disease (ICD)-10 criteria for a depressive episode assessed with the revised clinical interview schedule.”

But....

• Several authors now suggest that the term ‘Treatment Resistant Depression’ (TRD) should be defined by two failed courses of antidepressant medication at adequate doses and durations

• However, there is no consensus on the number, type, dose, and duration of therapeutic agents, and the specific criteria for ‘resistance’ and ‘non-response’


Issues

• Although NICE and APA guidelines recommend that psychotherapy should be offered to patients experiencing depression, current definitions of TRD do not require psychotherapy to have been tried before the term is used

• There are also concerns that the term ‘treatment resistant depression’ is pejorative, and places the blame for ‘resistance’ on the patient

• Little evidence supporting the assertion that people who fail to respond to two ADMs should be placed in a distinct ‘hard to treat’ category regardless of other psychosocial and/or qualitative factors

• NICE rejected the use of the term ‘treatment Resistant Depression’ in their recent guidelines for the treatment of depression

Co-production and Consensus

• Suggested definitions have not been arrived at using formal consensus development methods such as Delphi or nominal group techniques

• In a review of the development of guidelines,* formal methods were seen to outperform informal ones and were subsequently recommended for situations where the objective is to arrive at a consensus within groups who may hold different opinions and where evidence is not cut and dried


Aims

• To derive an agreed term to use when describing people with depression who do not respond to treatment.

• To determine an agreed operational definition of the preferred term for use in clinical situations and research trials

Method (1)

• ‘Delphi’ consensus development method
  – a robust, transparent and empirical method of establishing consensus through an iterative process that identifies ‘central tendency’ among a polled group and measures the ‘level of agreement’ around it.

• Delphi has been used previously in the field of mental health to better understand self-help strategies for depression and to determine elements of a Mental Health First Aid training course

• Never been used to develop consensus around the definition of depression that does not respond to treatment

Population

• All corresponding authors of papers published in peer reviewed English language journals from 2005 onwards in the field of treatment resistant depression (TRD)

• Authors cited by the NICE guidelines in 2010 in the wider field of depression

• Authors published before 2005 sufficiently in the TRD field to be recognised as an expert

• Chief Executives of three UK-based mental health non-governmental organisations
Method (2)

- Three-round online Delphi consultation
  - Round one: we invited respondents to suggest different options for the definition of depression which does not respond to treatment and operational criteria for it
  - Round two: respondents were presented with definitions, operational criteria and standards derived from the thematic analysis of round one and asked to select or rate their preferred options from those suggested
  - Round three: respondents were presented with the same options as in round two, plus data on the preferences expressed by themselves and the whole respondent group and asked to re-select and re-rate these options

Round 1

- Narrative responses to five open-ended questions:
  - What term would you use to describe depression that does not subside despite treatment/therapy?
  - Is there another term you would prefer to use to describe depression that does not subside despite treatment/therapy?
  - What factors would lead you to use either of these terms for a patient?
  - When would you NOT use these terms for a patient whose depression does not subside despite treatment/therapy?
  - What do you think is the best treatment protocol for people whose depression does not subside despite treatment/therapy?

Round 1 Analysis

- Two authors analysed and reviewed the narrative open-text data from round one using a thematic analysis procedure
- We grouped both manifest and latent meaning statements together until a set of terms and criteria that incorporated the constituent individual responses was determined
- We then agreed and generated a list of the final terms, criteria and standards

Round 2

- We presented participants with a list of the terms suggested in round one to describe depression that does not subside despite treatment/therapy and asked participants to select their first, second and third preferences.
- We presented criteria generated from the analysis of round one data to participants in round two, defined these criteria, and asked participants to rate each criterion as not important, slightly important, markedly important or very important.
- We then asked participants to choose standards from a multiple-choice list against which each criterion should be judged

Example Criterion and Standards

- Criterion: ‘Treatment Adequacy’
- Potential standards
  - Evidence based
  - Delivered at the recommended dose or intensity
  - Delivered at the recommended duration
  - Adhered to or engaged with by the patient
  - Delivered in a high quality manner
  - Delivered at the maximum tolerated dose or intensity
  - Tailored to the patient’s needs and circumstances

Round 2 and 3 Analysis

- Calculated the percentages of participants rating each criterion as not important, slightly important, markedly important or very important
- Calculated the percentage of respondents who selected specific standards for each criterion from the lists presented
- Applied established criteria to calculate the level of consensus.
  - High consensus = 80% or more of respondents rated the criterion as markedly or very important or endorsed a standard
  - Moderate consensus = 60 – 79% of respondents rated the criterion as markedly or very important or endorsed a standard
  - No consensus = less than 60% of respondents rated the criterion as markedly or very important or endorsed a standard
Results

• Round 1: 104 respondents
• Round 2: 75/104 (72% of previous round)
• Round 3: 71/75 (95% of previous round)

<table>
<thead>
<tr>
<th>Research Experience</th>
<th>Round 1</th>
<th>Round 2</th>
<th>Round 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>90 (87%)</td>
<td>66 (88%)</td>
<td>63 (89%)</td>
</tr>
<tr>
<td>No</td>
<td>14 (13%)</td>
<td>8 (12%)</td>
<td>8 (11%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Clinical Experience</th>
<th>Round 1</th>
<th>Round 2</th>
<th>Round 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>62 (60%)</td>
<td>46 (61%)</td>
<td>42 (59%)</td>
</tr>
<tr>
<td>No</td>
<td>42 (40%)</td>
<td>29 (39%)</td>
<td>33 (41%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Psychiatric Experience</th>
<th>Round 1</th>
<th>Round 2</th>
<th>Round 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>93 (89%)</td>
<td>65 (87%)</td>
<td>63 (89%)</td>
</tr>
<tr>
<td>No</td>
<td>11 (11%)</td>
<td>9 (13%)</td>
<td>8 (11%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Psychology Experience</th>
<th>Round 1</th>
<th>Round 2</th>
<th>Round 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>48 (46%)</td>
<td>37 (49%)</td>
<td>33 (46%)</td>
</tr>
<tr>
<td>No</td>
<td>56 (54%)</td>
<td>38 (51%)</td>
<td>47 (54%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>General Practitioners/Family Doctors</th>
<th>Round 1</th>
<th>Round 2</th>
<th>Round 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>7 (7%)</td>
<td>7 (9%)</td>
<td>6 (8%)</td>
</tr>
<tr>
<td>No</td>
<td>97 (93%)</td>
<td>93 (91%)</td>
<td>95 (92%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Non-clinical academic researcher</th>
<th>Round 1</th>
<th>Round 2</th>
<th>Round 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>11 (11%)</td>
<td>10 (13%)</td>
<td>9 (13%)</td>
</tr>
<tr>
<td>No</td>
<td>89 (89%)</td>
<td>90 (87%)</td>
<td>86 (87%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Expert by Experience</th>
<th>Round 1</th>
<th>Round 2</th>
<th>Round 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>2 (2%)</td>
<td>1 (1%)</td>
<td>1 (1%)</td>
</tr>
<tr>
<td>No</td>
<td>98 (98%)</td>
<td>99 (99%)</td>
<td>99 (99%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Other (e.g. voluntary care provider)</th>
<th>Round 1</th>
<th>Round 2</th>
<th>Round 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>11 (11%)</td>
<td>5 (7%)</td>
<td>5 (7%)</td>
</tr>
<tr>
<td>No</td>
<td>93 (93%)</td>
<td>95 (93%)</td>
<td>95 (93%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>UK</th>
<th>Round 1</th>
<th>Round 2</th>
<th>Round 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>33 (32%)</td>
<td>24 (32%)</td>
<td>22 (31%)</td>
</tr>
<tr>
<td>No</td>
<td>71 (70%)</td>
<td>76 (76%)</td>
<td>77 (77%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>USA</th>
<th>Round 1</th>
<th>Round 2</th>
<th>Round 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>29 (28%)</td>
<td>15 (20%)</td>
<td>15 (21%)</td>
</tr>
<tr>
<td>No</td>
<td>75 (72%)</td>
<td>85 (86%)</td>
<td>85 (89%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Europe</th>
<th>Round 1</th>
<th>Round 2</th>
<th>Round 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>26 (25%)</td>
<td>22 (29%)</td>
<td>21 (30%)</td>
</tr>
<tr>
<td>No</td>
<td>74 (75%)</td>
<td>78 (77%)</td>
<td>79 (77%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Canada</th>
<th>Round 1</th>
<th>Round 2</th>
<th>Round 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>6 (6%)</td>
<td>6 (8%)</td>
<td>6 (8%)</td>
</tr>
<tr>
<td>No</td>
<td>94 (94%)</td>
<td>94 (92%)</td>
<td>94 (92%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Asia</th>
<th>Round 1</th>
<th>Round 2</th>
<th>Round 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>5 (5%)</td>
<td>4 (5%)</td>
<td>3 (4%)</td>
</tr>
<tr>
<td>No</td>
<td>95 (95%)</td>
<td>96 (95%)</td>
<td>97 (96%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>South America</th>
<th>Round 1</th>
<th>Round 2</th>
<th>Round 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>3 (3%)</td>
<td>3 (4%)</td>
<td>3 (4%)</td>
</tr>
<tr>
<td>No</td>
<td>97 (97%)</td>
<td>97 (96%)</td>
<td>97 (96%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Australia</th>
<th>Round 1</th>
<th>Round 2</th>
<th>Round 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yes</td>
<td>2 (2%)</td>
<td>1 (1%)</td>
<td>1 (1%)</td>
</tr>
<tr>
<td>No</td>
<td>98 (98%)</td>
<td>99 (99%)</td>
<td>99 (99%)</td>
</tr>
</tbody>
</table>

Preferred Term

• The most popular term was: ‘Treatment Resistant Depression’.
• The majority of respondents in round two (n=53, 71%) and three (n=59, 83%) selected ‘Treatment Resistant Depression’ as their preferred term to describe depression that does not subside despite treatment.
• This represents a high (>80%) level of consensus.
• Participants did not reach consensus on any other term.

Consensus Criteria: Importance

• High consensus (>=80% agreement) on the importance of eight out of 13 potential criteria:
  – ‘Treatment/therapy ‘non-response’’
  – ‘Treatment adequacy’
  – ‘Diagnosis’
  – ‘Number of antidepressant attempts’
  – ‘Number of treatment/therapy attempts’
  – ‘Disorder duration’
  – ‘Use of qualitatively different treatment/therapy attempts’
  – ‘Types of antidepressant attempts’

• Moderate consensus (60-79% agreement) on three criteria:
  – ‘Additional exclusion factors’
  – ‘Number of psychotherapy attempts’
  – ‘Combination/augmentation’

• No consensus (<60% agreement) on two criteria:
  – ‘Types of psychotherapy attempts’
  – ‘Other physical treatment’

Consensus Criteria: Most Important Standards

• High consensus on the standards needed to meet six of the eight criteria which respondents had agreed were the most important:
  – ‘Treatment non-response’ (patient must remain clinically depressed);
  – ‘Treatment adequacy’ (treatment must be evidence based);
  – ‘Diagnosis’ (patient must have major depressive disorder);
  – ‘Number of antidepressant attempts’ (two or more);
  – ‘Number of treatment attempts’ (two or more);
  – ‘Use of qualitatively different treatments/therapies’ (important to try different treatments).

• Moderate consensus for one further most important criterion:
  – ‘Type of antidepressant attempts’ (different classes of antidepressants).

• There was also further moderate consensus on:
  – ‘Treatment adequacy’ (treatment delivered at recommended dose and intensity; Treatment delivered for the right duration; Treatment adhered to/engaged with properly by the patient).

• No consensus for the final ‘most important’ criterion:
  – ‘Disorder duration’ (defined as the length of time a patient had not responded to treatment). No option was endorsed by more than 40% of respondents, with respondents split largely between ‘At least 2-3 months’, ‘At least 6-12 months with no periods of remission’ and ‘At least 4-6 months with no periods of remission’.
Consensus Criteria: Other Standards

- Moderate consensus on one moderately important standard
  - ‘Additional exclusion factors’ (doubtful diagnosis; treatment non-response could be related to substance abuse)
- No consensus on the other two moderately important criteria:
  - ‘Number of psychotherapy attempts’ where respondents were evenly split between ‘one or more’ and ‘none required’
  - ‘Treatment combinations required’ where respondents were evenly split between ‘antidepressants and psychotherapy’ and ‘none’

Consensus Criteria: Other Standards

- No consensus on:
  - ‘Types of psychotherapy attempts’
  - ‘Other physical treatments’ reached high consensus for the standard ‘none required’

The Exeter Consensus Statement for Treatment Resistant Depression

- People with treatment resistant depression are defined as:
  - ‘patients with a diagnosis of major depressive disorder who remain clinically depressed after two or more different evidence-based treatments, which should include two different classes of antidepressants, where all treatments have been delivered at an adequate dose, adhered to by the patient for the recommended duration for that treatment, and where non response is unrelated to substance abuse.’

Exeter Consensus Statement for Treatment Resistant Depression

Decision Algorithm

Discussion

- A rigorous, transparent and empirical process
  - but a consensus is only as good as the consensees
- Failure to respond to psychological therapy neither a sufficient nor a necessary condition
- No ‘disorder duration’ standard
  - more contextualised and nuanced approach where the duration criteria is titrated against the specific treatments being delivered

Contact

http://medicine.exeter.ac.uk/research/healthserv/complexinterventions/projects/trd/
d.a.richards@exeter.ac.uk