

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/5533546>

Efficacy of antidepressants

Article *in* BMJ (online) · April 2008

DOI: 10.1136/bmj.39510.531597.80 · Source: PubMed

CITATIONS

42

READS

389

2 authors, including:



[Erick H Turner](#)

Oregon Health and Science University

93 PUBLICATIONS 3,577 CITATIONS

SEE PROFILE

BMJ

Efficacy of antidepressants

Erick H Turner and Robert Rosenthal

BMJ 2008;336:516-517; originally published online 4 Mar 2008;
doi:10.1136/bmj.39510.531597.80

Updated information and services can be found at:
<http://bmj.com/cgi/content/full/336/7643/516>

These include:

References This article cites 8 articles, 1 of which can be accessed free at:
<http://bmj.com/cgi/content/full/336/7643/516#BIBL>

Rapid responses 12 rapid responses have been posted to this article, which you can access for free at:
<http://bmj.com/cgi/content/full/336/7643/516#responses>

You can respond to this article at:
<http://bmj.com/cgi/eletter-submit/336/7643/516>

Email alerting service Receive free email alerts when new articles cite this article - sign up in the box at the top left of the article

Topic collections Articles on similar topics can be found in the following collections

- [Mood disorders \(including depression\)](#) (706 articles)
- [Regulation](#) (733 articles)
- [Drugs: psychiatry](#) (436 articles)

Notes

To order reprints follow the "Request Permissions" link in the navigation box

To subscribe to *BMJ* go to:
<http://resources.bmj.com/bmj/subscribers>

Editorials represent the opinions of the authors and not necessarily those of the *BMJ* or BMA

For the full versions of these articles see bmj.com

EDITORIALS

Antidepressants and suicide

Treatment is probably too sporadic to affect overall suicide rates



ANTOINETTE GERVAUX/FOTOLIA

RESEARCH, pp 539, 542

Gregory Simon investigator, Center for Health Studies, Group Health Cooperative, Seattle, WA 98112, USA
Simon.g@ghc.org

Competing interests: During the past five years Dr Simon has received \$750 (£382;€507) from Wyeth Pharmaceuticals for consultation regarding antidepressant adherence and \$1137.50 from Bristol Myers Squibb for consultation regarding suicide during psychotropic drug treatment. He has no active or planned consulting arrangements and no other financial interests to declare.

Provenance and peer review: Commissioned; not externally peer reviewed.

BMJ 2008;336:515-6

doi: 10.1136/bmj.39482.666366.80

Two accompanying observational studies by Biddle and colleagues and Wheeler and colleagues add to the recent literature regarding changes in antidepressant use and changes in suicide rates.^{1 2} The current controversy began in 2003, when reanalyses of data from clinical trials raised concerns that antidepressants might precipitate suicidal thoughts or attempts at suicide. Children and adolescents starting treatment with several newer antidepressants had a 4% risk of developing suicidal ideation or behaviour, compared with 2% in those receiving placebo.³ However, clinical trials cannot determine whether antidepressants increase or decrease the risk of genuine suicide attempts or death from suicide because these outcomes are, fortunately, too rare. No deaths from suicide and few attempts at suicide have occurred to date in trials of antidepressants in adolescents,³ so any clinical trial that could reliably assess effects on death from suicide would require several times more participants than have been included in all such trials to date.

Any study of serious suicide attempts or death from suicide must fall back on observational analyses of data from large populations. What population based studies gain in sample size, however, they lose in ability to account for bias or confounding. For example, if we simply compare suicide rates between people who are using and not using antidepressants, we would erroneously conclude that antidepressants increase risk. That conclusion would ignore the fact that symptoms of depression (including suicidal ideation) are the reason that antidepressants are prescribed. Although more sophisticated observational studies try to account for bias or confounding, they are never as good as randomised trials.

Bearing in mind these limitations, several observational studies have suggested that antidepressants might actually reduce the risk of attempted suicide or death from suicide. In adolescents and adults, the risk of suicide declines sharply after starting treatment with antidepressants.⁴ Areas with higher rates of antidepressant prescribing tend to have lower suicide rates.⁵ As the use of antidepressants in adolescents increased in the United States between 1990 and 2003, suicide rates declined. And when warnings led to decreased antidepressant use between 2003 and 2004, the suicide rate in US adolescents increased for the first time in a decade.⁶

More recent observational studies, however, raise doubts about an association between suicide rates and changes in antidepressant prescribing. The most recent

US data suggest that adolescent suicide deaths began to decrease again between 2004 and 2005.⁷ The two accompanying observational studies show a lack of connection between antidepressant prescribing rates and suicide rates in adolescents and young adults in the United Kingdom.^{1 2} Suicide rates declined when antidepressant use steadily increased but continued to decline when the use of antidepressants fell sharply in 2003 and 2004.

In truth, it would be surprising if antidepressants had any effect—positive or negative—on the risk of suicide in the general population. In the US, only half of adults and a quarter of adolescents who have a major episode of depression in any given year start taking an antidepressant.⁸ Nearly half of all adults who start treatment discontinue after just a few weeks.⁹ Only 3% of adolescents dying by suicide in New York City had toxicology data showing recent use of antidepressants.¹⁰ Sustained use of antidepressants is probably too rare to have much overall effect on risk of suicide in people living with depression.

The shifting association between antidepressant prescribing rates and suicide rates argues for caution in interpreting other ecological associations. Although we can see that changes in suicide mortality are associated with changes in social conditions or population rates of substance use, ecological associations do not imply causality. We tend to search for or believe only those associations that fit in with our expectations.

Observational data do clearly show that warnings in the UK and the US reduced rates of antidepressant prescribing, especially in adolescents. Between 2003 and 2005, antidepressant use in adolescents declined by about 20% in the US, 30% in the Netherlands, and 40% in the UK.^{1 2 6}

Unfortunately, we can find no evidence that regulators' urgent recommendations for closer monitoring of treatment led to any improvement in practice. The US Food and Drug Administration urges weekly or biweekly follow-up after starting antidepressant treatment. National quality measures in the US indicate that only a fifth of patients starting antidepressants have even three follow-up visits over three months.⁹ Furthermore, follow-up rates have shown no improvement over the past five years. We can only hope that regulatory warnings will eventually have as much effect on the quality of treatment as on the quantity.

1 Biddle L, Brock A, Brookes ST, Gunnell D. Suicide rates in young men in England and Wales in the 21st century: time trend study. *BMJ* 2008 doi: 10.1136/bmj.39475.603935.25.

- 2 Wheeler BW, Gunnell D, Metcalfe C, Stephens P, Martin RM. The population impact on incidence of suicide and non-fatal self harm of regulatory action against the use of selective serotonin reuptake inhibitors in under 18s in the United Kingdom: ecological study. *BMJ* 2008 doi: 10.1136/bmj.39462.375613.BE.
- 3 Hammad T, Laughren T, Racoosin J. Suicidality in pediatric patients treated with antidepressant drugs. *Arch Gen Psychiatry* 2006;63:332-9.
- 4 Simon G, Savarino J, Operskalski B, Wang P. Suicide risk during antidepressant treatment. *Am J Psychiatry* 2006;163:41-7.
- 5 Olfson M, Shaffer D, Marcus S, Greenberg T. Relationship between antidepressant medication treatment and suicide in adolescents. *Arch Gen Psychiatry* 2003;60:978-82.
- 6 Gibbons RD, Brown CH, Hur K, Marcus SM, Bhaumik DK, Erkens JA, et al. Early evidence on the effects of regulators' suicidality warnings on SSRI prescriptions and suicide in children and adolescents. *Am J Psychiatry* 2007;164:1356-63.
- 7 Olfson M, Shaffer D. SSRI prescriptions and the rate of suicide. *Am J Psychiatry* 2007;164:1907-8.
- 8 Office of Applied Statistics. *2006 national survey on drug use and health: detailed tables*. Washington, DC: Substance Abuse and Mental Health Services Administration, US Department of Health and Human Services, 2007.
- 9 National Committee for Quality Assurance. *The state of health care quality 2007*. Washington, DC: NCQA, 2007.
- 10 Leon A, Marzuk P, Tardiff K, Bucciarelli A, Markham Piper T, Galea S. Antidepressants and youth suicide in New York City, 1999-2002. *J Am Acad Child Adolesc Psychiatry* 2006;45:1054-8.

Efficacy of antidepressants

Is not an absolute measure, and it depends on how clinical significance is defined

See also **FEATURE**, p 532

Erick H Turner assistant professor, Department of Psychiatry, Oregon Health and Science University, Portland, OR 97239, USA
turnere@ohsu.edu

Robert Rosenthal distinguished professor, Department of Psychology, University of California, Riverside, CA 92521, USA

Competing interests: None declared.

Provenance and peer review: Commissioned; not externally peer reviewed.

BMJ 2008;336:516-7

doi: 10.1136/bmj.39510.531597.80

In February 2008, Kirsch and colleagues reported a meta-analysis of the efficacy of antidepressants using data from clinical trials submitted to the Food and Drug Administration.¹ They provocatively concluded, "there seems little evidence to support the prescription of antidepressant medication to any but the most severely depressed patients."

In January this year, we published an article about the selective publication of antidepressant trials and its influence on apparent efficacy,² in which we also used FDA data. Our main finding was that antidepressant drugs are much less effective than is apparent from journal articles. From the FDA data we derived an overall effect size of 0.31. Kirsch and colleagues used FDA data from four of the 12 drugs we examined and calculated an overall effect size of 0.32.

Although these two sets of results were in excellent agreement, our interpretations of them were quite different. In contrast to Kirsch and colleagues' conclusion that antidepressants are ineffective, we concluded that each drug was superior to placebo. The difference in our interpretations stems from Kirsch and colleagues' use of the criteria for clinical significance recommended by the UK's National Institute for Health and Clinical Excellence (NICE).

Clinical significance is an important concept because a clinical trial can show superiority of a drug to placebo in a way that is statistically, but not clinically, significant. Tests of statistical significance give a yes or no answer (for example, $P < 0.05$ is deemed significant, $P > 0.05$ non-significant) that tells us whether the true effect size is zero or not, but it tells us nothing about the size of the effect.³ In contrast, effect size does, and thus allows us to look at the question of clinical significance. Values of 0.2, 0.5, and 0.8 were proposed to represent small, medium, and large effects, respectively.⁴

NICE chose the "medium" value of 0.5 as a cut-off below which they deem benefit of a drug not clinically significant.⁵ This is problematic because it transforms effect size, a continuous measure, into a yes or no measure, thereby suggesting that drug efficacy is either totally present or absent, even when comparing values as close together as 0.51 and 0.49. Kirsch and colleagues compared their effect size of 0.32 to the 0.50

cut-off and concluded that the benefits of antidepressant drugs were of no clinical significance.

But on what basis did NICE adopt the 0.5 value as a cut-off? When Cohen first proposed these landmark effect size values, he wrote, "The terms 'small', 'medium', and 'large' are relative . . . to each other . . . the definitions are arbitrary . . . these proposed conventions were set forth throughout with much diffidence, qualifications, and invitations not to employ them if possible." He also said, "The values chosen had no more reliable a basis than my own intuition." Thus, it seems doubtful that he would have endorsed NICE's use of an effect size of 0.5 as a litmus test for drug efficacy.

To illustrate Cohen's use of "relative" with a metaphor, imagine antidepressant efficacy measured in terms of litres of a fluid called "d-juice" (named after Cohen's "d"—the effect size measure described here). When our group measured 0.41 litres of d-juice in the "glass" representing journal articles, but 0.31 litres in the FDA glass, we concluded that the FDA glass was empty relative to the journal glass. Nevertheless, we acknowledged that 0.31 litres was an amount that was measurable and significant. Kirsch and colleagues measured 0.32 litres of d-juice, but because they did not consider the glass sufficiently full (defined arbitrarily as $P \geq 0.5$), they concluded that the glass contained virtually no d-juice whatsoever. To summarise, we agree that the antidepressant "glass" is far from full, but we disagree that it is completely empty.

Hypothetically, if antidepressants are not worth taking, then what should doctors and patients do? Kirsch and colleagues recommend that if antidepressants are to be used at all they should be used only when alternative treatments have failed to provide a benefit.¹ Although the authors did not specify a preferred first line treatment, they may have had psychotherapy in mind.^{6,7} It seems unfair that pharmacological, and not psychotherapeutic, treatment has become the usual first line approach to depression merely for economic reasons.⁷ But before we embrace any treatment as first line, it is prudent to ask whether its efficacy is beyond question. For psychotherapy trials, there is no equivalent of the FDA whose records we can examine, so how can we be sure that selective publication is not occurring here as well?

Our clinical recommendation is that when

considering the potential benefits of treatment with antidepressants, be circumspect but not dismissive. Efficacy measured in clinical trials does not necessarily translate into effectiveness in clinical practice.⁸ Patients' individual responses are like clinical trial effect sizes in that they are not all or none. Thus, when a patient is tried on his or her first antidepressant, a partial response should not be surprising or discouraging. Also, depression rating scales used in clinical trials seldom measure quality of life, which has been suggested to be a reasonable measure of clinical significance.⁹

With regard to policy, we reiterate our request in 2004 for drug regulatory authorities such as the FDA to make their reviews publicly available on the world wide web—retrospectively.¹⁰ Making this unbiased information more accessible will allow other researchers to move beyond antidepressants and ascertain the true efficacy of all marketed drugs.

- 1 Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS Med* 2008;5:e45.
- 2 Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med* 2008;358:252-60.
- 3 Cohen J. The earth is round (P<.05). *Am Psychologist* 1994;49:997-1003.
- 4 Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd ed. New York: Lawrence Erlbaum Associates, 1988.
- 5 National Institute for Health and Clinical Excellence. *Depression: management of depression in primary and secondary care*. Clinical practice guideline CG23. 2004. www.nice.org.uk/nicemedia/pdf/cg023fullguideline.pdf.
- 6 Antonuccio DO, Burns DD, Danton WG. Antidepressants: a triumph of marketing over science? *Prevent Treat* 2002;5:Article 25.
- 7 Antonuccio DO, Danton WG, DeNelsky GY. Psychotherapy versus medication for depression: challenging the conventional wisdom with data. *Prof Psychol Res Pract* 1995;26:574-85.
- 8 Revicki DA, Frank L. Pharmacoeconomic evaluation in the real world. Effectiveness versus efficacy studies. *Pharmacoeconomics* 1999;15:423-34.
- 9 Gladis M. Quality of life: expanding the scope of clinical significance. *J Consult Clin Psychol* 1999;67:320-31.
- 10 Turner EH. A taxpayer-funded clinical trials registry and results database. *PLoS Med* 2004;1:e60. <http://dx.doi.org/10.1371/journal.pmed.0010060>.

Acupuncture with in vitro fertilisation

May increase birth rates, but guidelines should await the results of ongoing trials

RESEARCH, p 545

Anja Pinborg postdoctoral research fellow
pinborg@nru.dk

Anne Loft consultant

Anders Nyboe Andersen

professor of reproductive medicine, Fertility Clinic, Copenhagen University Hospital, Rigshospitalet, 2100 Copenhagen, Denmark

Competing interests: None declared.

Provenance and peer review: Commissioned; not externally peer reviewed.

BMJ 2008;336:517-8

doi: 10.1136/bmj.39503.643727.80

According to the European Society of Human Reproduction and Embryology, more than 300 000 treatment cycles of in vitro fertilisation and intracytoplasmic sperm injection are performed each year in Europe.¹ Because in vitro fertilisation can affect physical and mental health, professionals delivering this treatment are trying to optimise the technology to increase birth rates.

About 90% of all assisted reproductive treatment cycles result in the transfer of at least one embryo, but only about 25% of all cycles end in implantation of the embryo and live birth. The main factor limiting the success of treatment is failure of implantation and not the lack of human embryos for transfer. Initiatives to improve rates of implantation have had varying success. In the accompanying systematic review, Manheimer and colleagues report improved pregnancy rates with in vitro fertilisation when acupuncture accompanies embryo transfer.²

As far as we are aware, this is the first systematic review and meta-analysis of the success rates of in vitro fertilisation with adjuvant acupuncture. It included seven methodologically sound randomised controlled trials with little clinical heterogeneity that studied 1366 women having in vitro fertilisation between 2002 and 2006. Adjuvant acupuncture at the time of embryo transfer significantly increased the rates of clinical pregnancy (odds ratio 1.65, 95% confidence interval 1.40 to 2.49) and live birth (1.91, 1.39 to 2.64). The number needed to treat with adjuvant acupuncture to achieve one more pregnancy was 10 and to obtain one more live birth this number was nine. A subgroup analysis restricted to the three studies with highest clinical pregnancy rates in the control groups, however, found no significant benefit of acupuncture (1.24, 0.86 to 1.77).

The pooled odds ratios on pregnancy rates after in

vitro fertilisation with adjuvant acupuncture are higher than reported odds ratios for drugs or other procedures given to enhance the success of this treatment. For example, gonadotrophin preparations with luteinising hormone activity versus preparations without such activity gave a relative risk of clinical pregnancy of 1.17 (1.03 to 1.34)³; the use of assisted hatching (which prepares the embryos before transfer) gave an odds ratio of clinical pregnancy of 1.63 (1.27 to 2.09)⁴; and transfer of day 5 embryos (blastocyst stage) versus day 3 embryos (cleavage stage) gave an odds ratio for clinical pregnancy of 1.27 (1.03 to 1.55).⁵

Manheimer and colleagues have extensively dealt with and discussed the limitations of their meta-analysis and the possibility of biased results. We consider the results to be just as solid as any other meta-analysis of adjuvant treatment or medical procedures in assisted reproduction. However, although meta-analyses provide the highest level of scientific evidence, caution should be used when using data from small meta-analyses with a risk of publication bias to make daily clinical decisions.

The current meta-analysis was based on seven randomised controlled trials, and only three relatively small trials included a sham control group. Results of meta-analyses should always be interpreted in terms of how biologically plausible they are. In this instance, the biological mechanism is difficult to explain. Adjuvant acupuncture was given immediately before or immediately after embryo transfer. The effects of acupuncture are therefore most likely to involve uterine contractility rather than uterine receptivity. Acupuncture is unlikely to have exerted a central effect by mediating the release of neurotransmitters because the hypothalamic secretion of gonadotrophin releasing hormone

would be “switched off” by the gonadotrophin releasing hormone analogues used during in vitro fertilisation and the high levels of oestradiol. Acupuncture may act by reducing the contractility of the uterus and thereby avoiding expulsion of embryos after transfer or through unknown effects on the blood flow to the endometrium.⁶ Animal and human studies investigating the effect of acupuncture through a direct or indirect mechanism on the uterus or the endometrium are needed.

National guidelines and recommendations should be based on systematic reviews and meta-analyses. Should adjuvant acupuncture now be included in national guidelines such as the National Institute for Health and Clinical Excellence guidelines from the British Fertility Society?⁷ We think that it is too early for such a recommendation. Publication bias may have influenced the results of the meta-analysis. A Danish randomised controlled trial on adjuvant acupuncture that includes more than 600 women having in vitro fertilisation (twice as many as in the largest randomised controlled trial included in the meta-analysis) is currently under way. Before adding

adjuvant acupuncture for in vitro fertilisation to any national guideline we must wait for the results of this and other studies to clarify the value of this treatment.

- 1 Nyboe Andersen A, Gossens V, Gianaroli L, Felberbaum R, de Mouzon J, Nygren KG. Assisted reproductive technology in Europe, 2003. Results generated from European registers by ESHRE. *Hum Reprod* 2007;22:1513-25.
- 2 Manheimer E, Zhang G, Udoff L, Haramati A, Langenberg P, Berman BM, et al. Effects of acupuncture on rates of pregnancy and live birth among women undergoing in vitro fertilisation: systematic review and meta-analysis. *BMJ* 2008 doi: 10.1136/bmj.39471.430451.BE.
- 3 Coomarasamy A, Afnan M, Cheema D, van der Veen F, Bossuyt PM, van Wely M. Urinary hMG versus recombinant FSH for controlled ovarian hyperstimulation following an agonist long down-regulation protocol in IVF or ICSI treatment: a systematic review and meta-analysis. *Hum Reprod* 2008;23:310-5.
- 4 Edi-Osagie E, Hooper L, Seif MW. The impact of assisted hatching on live birth rates and outcomes of assisted conception: a systematic review. *Hum Reprod* 2003;18:1828-35.
- 5 Papanikolaou EG, Kolibianakis EM, Tournaye H, Venetis CA, Fatemi H, Tarlatsis B, et al. Live birth rates after transfer of equal number of blastocyst or cleavage-stage embryos in IVF. A systematic review and meta-analysis. *Hum Reprod* 2008;23:91-9.
- 6 Fanchin R, Righini C, Olivennes F, Taylor S, de Ziegler D, Frydman R. Uterine contractions at the time of embryo transfer alter pregnancy rates after in-vitro fertilization. *Hum Reprod* 1998;13:1968-74.
- 7 British Fertility Society. *Key facts on infertility, IVF and NHS provision*. 2005. www.britishtfertilitysociety.org.uk/public/factsheets/keyfacts.html.

Funding for primary health care in developing countries

Money from disease specific projects could be used to strengthen primary care

ANALYSIS, p 536

Jan De Maeseneer professor of family medicine, Department of Family Medicine and Primary Health Care, Ghent University, Belgium

jan.demaeseneer@ugent.be

Chris van Weel professor of family medicine, Department of Family Medicine, Radboud University Medical Centre, Nijmegen, Netherlands

David Egilman clinical associate professor, Brown University, Providence, RI, USA

Khaya Mfenyana professor of family medicine, Department of Family Medicine, Walter Sisulu University, Mthatha, South Africa

Arthur Kaufman professor of community health, Department of Community Health, University of New Mexico Health Sciences Center, USA

Nelson Sewankambo

professor of medicine, Faculty of Medicine, Makerere University, Kampala, Uganda.

Maike Flinkenflögel

researcher, Department of Family Medicine and Primary Health Care, Ghent University, Belgium

Competing interests: None declared.

Provenance and peer review: Commissioned; externally peer reviewed.

BMJ 2008;336:518-9

doi: 10.1136/bmj.39496.444271.80

The World Health Organization's World Health Report 2007 deals with access to primary health care as an essential prerequisite for health.¹ It acknowledges the importance of the Alma-Ata declaration of 1978, which called for integrated primary health care as a way to deal with major health problems in communities and for access to care as part of a comprehensive national health system. Yet the mission of Alma-Ata—to provide accessible, affordable, and sustainable primary health care for all—has been implemented only partially in developing countries.² We have therefore instigated the “15by2015” campaign (www.15by2015.org), which proposes a funding mechanism for strengthening primary health care in developing countries.

In the accompanying analysis article, Gillam notes that most developing countries have failed to provide even basic primary healthcare packages. Weaknesses in primary healthcare services often result from a variety of forces, including economic crises and market reforms, which limit the range and coverage of services and thus their effect on health.^{3 4} On the positive side, between 1997 and 2002, financial support to improve health care in developing countries increased by about 26%, from \$6.4bn (£3.3m; €4.4m) to \$8.1bn.⁵ However, most aid was allocated to disease specific projects (termed “vertical programming”) rather than to broad based investments in health infrastructure, human resources, and community oriented primary healthcare services (“horizontal programming”).⁶

An example of vertical programming is the enormous donor response to the HIV epidemic. In 2006, although Zambia's entire Ministry of Health budget was only \$136m, the President's Emergency Plan for AIDS Relief

provided the country with an HIV targeted budget of \$150m. This unbalanced distribution of health funding occurs across sub-Saharan Africa. Thus, although HIV positive patients receive free care, others with more routine diseases receive poor care and still have to pay. Salaries of healthcare providers working for donor funded vertical programmes are often more than double those of equally trained government workers in the fragile public health sector. This lures government workers to the higher paying vertical programmes and creates an internal “brain drain.” But it is the underfunded primary care clinics and health centres that care for all diseases, including common illnesses such as diarrhoea, malnutrition, and respiratory tract infections, which take many more lives than HIV, tuberculosis, and malaria.

A new global strategy is needed to reinforce community focused primary health care in developing countries. This will require cooperation between ministries, universities, non-governmental organisations, and donors working on health to overcome severe resource constraints, including insufficient numbers of doctors, pharmacists, and other health personnel. Four international organisations—the World Organization of Family Doctors (www.globalfamilydoctor.com); Global Health through Education, Training and Service (www.ghets.org); the Network: Towards Unity for Health (www.the-networktufh.org); and the European Forum for Primary Care (www.euprimarycare.org)—have therefore set up the 15by2015 campaign to foster a better balance between vertical and horizontal aid. This campaign calls for major international donors to assign 15% of their vertical budgets by 2015 to strengthening horizontal primary healthcare systems so that all diseases can

be prevented and treated in a systematic way.

This campaign is not acting in a vacuum. A broad approach—orienting funds to governments for comprehensive provision of care—is being implemented in several countries in sub-Saharan Africa. The Global Fund to fight AIDS, Tuberculosis, and Malaria has called for investment to strengthen health systems and tackle social determinants by supporting strategies to reduce poverty.⁷ The United Kingdom's prime minister, Gordon Brown, in a joint statement with Germany's chancellor, Angela Merkel, announced the launch of the "International Health Partnership."⁸ The core idea is to encourage low income countries to create comprehensive country-wide health programmes, which would serve as the basis for all foreign assistance for health. Hopefully, other donors will follow these leads.

How would 15by2015 work? Take the example of Mozambique. In 2005, the total health expenditure in the country was \$356m. Foreign assistance accounted for \$243m, from which \$130m was channelled through disease specific vertical funds managed directly by donors.⁹ We propose that, 15% of the vertical funds from donor organisations (in this example, \$19.5m) should be diverted into the government's common health fund and be earmarked for strengthening primary health care through improvement of infrastructure, health education, and investment in human resources. This amount of money could support 65 health centres for a year. These centres could be staffed by primary care teams including family doctors, mid-level care workers, primary care nurses, pharmacists, and health promoters. If one primary healthcare centre covers a population of 20 000 people, then 65 health centres would give 1.3 million people access to improved primary health care.

Part of the 15by2015 fund could be allocated to sup-

port the training and upgrading of skills. It could also be used to provide better pay for health personnel to encourage them to stay in areas where they are needed and to pay for community health workers, mid-level care workers, and "African family physicians" who are a fledgling but emerging force.^{10 11} The Ministry of Health should monitor the accessibility and quality of this care in a transparent way to ensure that the 15by2015 fund is used most effectively to improve community health.

- 1 WHO. The World Health Report 2007. *A safer future. Global public health security in the 21st century*. 2007. www.who.int/whr/2007/whr07_en.pdf.
- 2 De Maeseneer J, Willems S, De Sutter A, Van de Geuchte I, Billings M. *Primary health care as a strategy for achieving equitable care: a literature review commissioned by the Health Systems Knowledge Network*. 2007. www.who.int/social_determinants/resources/csdh_media/primary_health_care_2007_en.pdf.
- 3 Gillam S. Is the declaration of Alma-Ata still relevant to primary health care? *BMJ* 2008 doi: 10.1136/bmj.39469.432118.AD.
- 4 Segall M. District health systems in a neoliberal world: a review of five key policy areas. *Int J Health Plann Manag* 2003;18:S5-26.
- 5 WHO. *Engaging for health. Eleventh General Programme of Work 2006-2015*. 2006. http://whqlibdoc.who.int/publications/2006/GPW_eng.pdf.
- 6 De Maeseneer J, van Weel C, Egilman D, Mfenyana K, Kaufman A, Sewankambo N. Strengthening primary care: addressing the disparity between vertical and horizontal investment. *Br J Gen Pract* 2008;58:3-4.
- 7 Global Fund to Fight AIDS, Tuberculosis and Malaria. *A strategy for the global fund. Accelerating the effort to save lives*. 2007. www.theglobalfund.org/en/files/publications/strategy/Strategy_Document_HI.pdf.
- 8 Brown G. *Joint statement with chancellor Angela Merkel on international health partnership*. 2007. www.primeminister.gov.uk/output/Page12904.asp.
- 9 Martinez J. *Implementing a sector wide approach in health: the case of Mozambique*. 2006. www.hlspinstitute.org/files/project/100615/Mozambique_SWAP.pdf.
- 10 Pust R, Dahlman B, Khwa-Otsyula B, Armstrong J, Downing R. Partnerships creating postgraduate family medicine in Kenya. *Fam Med* 2006;38:661-6.
- 11 De Maeseneer J, Flinkenflögel M. *Primafamed network for training family physicians in sub-Saharan Africa, an Edulink-ACP-EU project*. 2008. www.primafamed.ugent.be.

Assessing the ability to work

New UK test claims to be fair but lacks rigorous scientific evaluation

Jos Verbeek occupational physician, Cochrane Occupational Health Field, Knowledge Transfer Team, Finnish Institute of Occupational Health, Kuopio, Finland

j.h.verbeek@amc.uva.nl

Frank van Dijk occupational physician, Coronel Institute of Occupational Health, Academic Medical Centre, Amsterdam, Netherlands

Competing interests: None declared.

Provenance and peer review: Commissioned; not peer reviewed.

BMJ 2008;336:519-20

doi: 10.1136/bmj.39472.451134.80

Recently, the Department of Work and Pensions in the United Kingdom announced a renewal of the personal capacity assessment. The report states that the renewal is expected to result in 20 000 fewer people claiming sickness benefits each year. It also claims that the new test is more robust, accurate, and fair than the previous one.¹

Two issues are at stake here: firstly, the provision of work and a decent income for millions of people with disabilities and, secondly, the billions of pounds that society is willing and able to pay in disability benefits. In many European countries, the growing numbers of people claiming disability benefit and expenditure on these benefits is an important point of political interest.²

The personal capacity assessment lists 17 activities, each of which can be given a score according to the degree of limitation. People with a score of 15 or more are assessed as unable to work. Changes have been made to the 17 activities and limitations in the new

assessment; for example, not being able to walk more than 30 metres had the highest disability score in the old assessment but this has been changed to 50 metres in the new one. Other changes make it more difficult to reach the highest level of work disability.

It is generally agreed that the World Health Organization (WHO) model of functioning provides the best framework for the evaluation of disability.³ The basis of this model is that disability has three major components apart from having a disease: impairments in bodily or mental functions or structures, limitations in activities, and restrictions in participation in societal roles. Personal and environmental factors also play a role.

It is important for doctors to understand the essential difference between having a disease and having a disability.⁴ The ability to work depends on balancing the limitations in activities with the demands that participation in working life imposes. The personal capacity assessment does reflect the WHO model in that the



ULRIKE PREUSS/PHOTOFUSION

related to sickness certification and the equipment used to evaluate functional capacity.^{7 8} This may be because the way society deals with disability is deeply embedded in culture and disability is a politically sensitive matter.² Subjecting the topic to scientific evaluation might be too confrontational or politically risky because the results might fundamentally challenge current practice.

The lack of use of evidence in general in insurance medicine has been noted by the Dutch Health Council, which therefore promotes evidence based disability assessment and produces guidelines for doctors working for social insurance companies.⁹ In addition, the Dutch Employee Insurance Schemes Implementing Body has collaborated with universities to form the Research Centre for Insurance Medicine, with the intention of improving the quality of disability assessment and guidance on reintegration.

Clinicians and researchers are increasingly interested in their patients' ability to work and how to help them return to work, especially those with back pain, rheumatoid arthritis, and cancer.¹⁰ The expertise they have gained should be put to better use in evaluating disability.

Employers have a relatively negative image of people with disabilities and many are reluctant to offer jobs.¹¹ Once a disabled person leaves the work force it is difficult for him or her to return. The Organisation for Economic Cooperation and Development therefore advises tailored intervention during the first few weeks of sick leave. This approach has been evaluated and found to be successful in the Netherlands.¹² Doctors can therefore help their patients with disabilities the most by supporting them in retaining their jobs.³

1 Henderson M. *Transformation of the personal capability assessment*. Department of Work and Pensions, 2007. www.dwp.gov.uk/welfarereform/tpca.pdf.

2 Zeitzer I. The challenges of disability pension policy: three western European case studies of the battle against the numbers. In: Fultz E, Ruck M, eds. *Reforming workers' protections: disability pensions in transformation*. Geneva: International Labour Office, 2002.

3 Verbeek JH. How can doctors help their patients to return to work? *PLoS Med* 2006;3:e88.

4 Bilsker D, Wiseman S, Gilbert M. Managing depression-related occupational disability: a pragmatic approach. *Can J Psychiatry* 2006;51:76-83.

5 Bolderson H, Mabbett D, Hvinden B. *Definitions of disability in Europe, a comparative analysis*. Brussels: EU Directorate General for Employment and Social Affairs, 2002.

6 UWV. *Functionele mogelijkheden lijst*. 2007. http://cba.uwv.nl/cba/opencms/CBA/module4/Map_CBBS_Intro/11.html.

7 Gouttebarga V, Wind H, Kuijer PP, Frings-Dresen MH. Reliability and validity of functional capacity evaluation methods: a systematic review with reference to Blankenship system, Ergos work simulator, Ergo-Kit and Isernhagen work system. *Int Arch Occup Environ Health* 2004;77:527-37.

8 Soderberg E, Alexanderson K. Sickness certification practices of physicians: a review of the literature. *Scand J Public Health* 2003;31:460-74.

9 Dutch Health Council. *Verzekeringsgeneeskundige protocollen*. 2007. <http://www.gezondheidsraad.nl/pdf.php?ID=1532&p=1>.

10 De Buck PD, le Cessie S, van den Hout WB, Peeters AJ, Rondoy HK, Westedt ML, et al. Randomized comparison of a multidisciplinary job-retention vocational rehabilitation program with usual outpatient care in patients with chronic arthritis at risk for job loss. *Arthritis Rheum* 2005;53:682-90.

11 McDonald S, Davis A, Stafford B. *Report of the survey of job brokers*. Report CRSP484S. Leicestershire: Centre for Research in Social Policy, 2003.

12 Organisation for Economic Cooperation and Development. *Transforming disability into ability, policies to promote work and income security for disabled people*. Paris: OECD, 2003.

It is important for doctors to understand the essential difference between having a disease and having a disability

test items refer to limitations in activities and the cut-off scores refer to the demands that work in general imposes on us. The cut-off score provides an indication of the severity of the limitations.

The new test claims to be fair and accurate, but the report does not define what this means. The test was evaluated against expert opinion in 212 cases. Even though essential qualifiers like sensitivity and specificity of the test are not reported, it is implied that specificity is higher and sensitivity is lower in the new test. Instead of referring to the usual trade-off between specificity and sensitivity, the authors of the report argue that the decrease in sensitivity occurred by chance. This is important because a decrease in sensitivity means that more disabled workers would be denied benefits. Even though the Department of Work and Pensions claims that the personal capacity assessment is the best assessment of its type in the world, we found no scientific reports on its validity in Medline. Such an important test deserves better scientific underpinning and evaluation.

Even within Europe, disability assessment varies widely, with some countries relying totally on strict rules and others on doctors' expertise.⁵ In the Netherlands, doctors working for insurance companies use a Dutch variant of the personal capacity assessment called the (remaining) functional abilities list.⁶ The abilities on the Dutch list are more work oriented than those on the UK list, and the scores for severity of limitations are different. For example, the most severe Dutch limitation is not being able to lift 1 kg compared with 0.5 kg in the UK. However, neither the Dutch nor the UK system has been properly evaluated, so we do not know which is best.

Surprisingly, no scientific evidence is available on the diagnostic accuracy of these tests and questionnaires that are used for evaluating disability. The few reports that exist on disability assessment are mainly