

Components of placebo effect: randomised controlled trial in patients with irritable bowel syndrome

Ted J Kaptchuk,¹ John M Kelley,² Lisa A Conboy,¹ Roger B Davis,³ Catherine E Kerr,¹ Eric E Jacobson,⁴ Irving Kirsch,⁵ Rosa N Schyner,¹ Bong Hyun Nam,¹ Long T Nguyen,¹ Min Park,¹ Andrea L Rivers,¹ Claire McManus,¹ Efi Kokkotou,³ Douglas A Drossman,⁶ Peter Goldman,⁷ Anthony J Lembo³

EDITORIAL by Spiegel and Harrington

¹Osher Research Center, Harvard Medical School, 401 Park Drive, Boston, MA 02215, USA

²Endicott College, 376 Hale Street, Beverly, MA 01915, USA

³Beth Israel Deaconess Medical Center, 330 Brookline Avenue, Boston, MA 02215, USA

⁴Department of Social Medicine, Harvard Medical School, 641 Huntington Avenue, Boston, MA 02215, USA

⁵Department of Psychology, University of Hull, Hull HU6 7RX

⁶Center for Functional GI and Motility Disorders, University of North Carolina School of Medicine, Chapel Hill, NC 27699, USA

⁷Harvard Medical School, 25 Shattuck Street, Boston, MA 02115, USA

Correspondence to: T J Kaptchuk ted_kaptchuk@hms.harvard.edu

BMJ 2008;336:999-1003

doi:10.1136/bmj.39524.439618.25

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: *BMJ* 2008, doi: 10.1136/bmj.39524.439618.25

ABSTRACT

Objective To investigate whether placebo effects can experimentally be separated into the response to three components—assessment and observation, a therapeutic ritual (placebo treatment), and a supportive patient-practitioner relationship—and then progressively combined to produce incremental clinical improvement in patients with irritable bowel syndrome. To assess the relative magnitude of these components.

Design A six week single blind three arm randomised controlled trial.

Setting Academic medical centre.

Participants 262 adults (76% women), mean (SD) age 39 (14), diagnosed by Rome II criteria for and with a score of ≥ 150 on the symptom severity scale.

Interventions For three weeks either waiting list (observation), placebo acupuncture alone (“limited”), or placebo acupuncture with a patient-practitioner relationship augmented by warmth, attention, and confidence (“augmented”). At three weeks, half of the patients were randomly assigned to continue in their originally assigned group for an additional three weeks.

Main outcome measures Global improvement scale (range 1-7), adequate relief of symptoms, symptom severity score, and quality of life.

Results At three weeks, scores on the global improvement scale were 3.8 (SD 1.0) v 4.3 (SD 1.4) v 5.0 (SD 1.3) for waiting list versus “limited” versus “augmented,” respectively ($P < 0.001$ for trend). The proportion of patients reporting adequate relief showed a similar pattern: 28% on waiting list, 44% in limited group, and 62% in augmented group ($P < 0.001$ for trend). The same trend in response existed in symptom severity score (30 (63) v 42 (67) v 82 (89), $P < 0.001$) and quality of life (3.6 (8.1) v 4.1 (9.4) v 9.3 (14.0), $P < 0.001$). All pairwise comparisons between augmented and limited patient-practitioner relationship were significant: global improvement scale ($P < 0.001$), adequate relief of symptoms ($P < 0.001$), symptom severity score ($P = 0.007$), quality of life ($P = 0.01$). Results were similar at six week follow-up.

Conclusion Factors contributing to the placebo effect can be progressively combined in a manner resembling a graded dose escalation of component parts. Non-specific effects can produce statistically and clinically significant

outcomes and the patient-practitioner relationship is the most robust component.

Trial registration Clinical Trials NCT00065403.

INTRODUCTION

Aside from the provision of a specific therapeutic regimen, a medical encounter might elicit non-specific or contextual benefits or what are most often called placebo effects. Such non-specific effects in a clinical setting can theoretically be separated into three components: a patient’s response to observation and assessment (Hawthorne effects), the patient’s response to the administration of a therapeutic ritual (placebo treatment), and the patient’s response to the patient-practitioner interaction.¹⁻³ We tested this by determining whether these distinct potential contributions can be separated and then combined incrementally to produce progressive improvement in clinical outcomes.

We carried out the trial on patients with irritable bowel syndrome. Previous randomised controlled trials of treatments for irritable bowel syndrome have shown a large positive response (about 40%) in placebo groups.⁴

METHODS

Study design

We conducted this randomised controlled trial in a single centre in 262 participants over six weeks. Participants were randomised to one of three groups: a “waiting list” that controlled for effects of assessment and observation (Hawthorne effects) and the natural course of the disease; “limited interaction,” providing placebo treatment with minimal interaction with the practitioner; or “augmented interaction,” providing placebo treatment with a defined positive patient-practitioner relationship. Our placebo treatment was delivered with a validated sham acupuncture device chosen because acupuncture has high placebo effects.⁵ All participants were evaluated at entry to the trial and after three and six weeks.

At three weeks, we randomised patients in the sham acupuncture groups to continue sham acupuncture or to switch to genuine acupuncture as part of a nested study (reported elsewhere). Results at three weeks provided data for the primary end point; those who remained on

placebo for the additional three weeks served to provide observations on non-specific effects over time.

Recruitment

Participants were recruited from advertisements in the media, fliers, and referrals from health professionals, were all at least 18 years old, and met the Rome II criteria for irritable bowel syndrome.⁶ We excluded patients if they had unexplained findings or had previously received acupuncture. Participants were allowed to continue medications for irritable bowel syndrome taken before entering the study if this therapeutic regimen was constant for at least the previous 30 days and during the trial.

Intervention components

Group 1 (waiting list)—Participants had neither placebo treatment nor interaction with a healthcare practitioner.

Group 2 (limited interaction)—Participants received a placebo intervention twice a week with “limited” interaction with a practitioner. The limited patient-practitioner relationship was established at the initial visit (duration <5 minutes). Practitioners explained that this was “a scientific study” for which they had been “instructed not to converse with patients.” The placebo needles were placed, and the patient left alone for 20 minutes after which the practitioner returned to remove the “needles.” At week three, participants completed assessments and those randomised to

continue the placebo treatment received an additional six sham treatments.

Group 3 (augmented interaction)—Participants in group 3 (augmented) received six sessions of placebo acupuncture using the same procedure as for group 2. In addition, they received an augmented patient-practitioner relationship that began at the initial visit (45 minutes’ duration). Content included questions concerning symptoms, relationships and lifestyle, non-gastrointestinal symptoms, and how the patient understood the “cause” and “meaning” of his or her condition. The interviewer incorporated a warm, friendly manner; active listening; empathy; and communication of confidence and positive expectation.

Informed consent and blinding

All participants gave written informed consent, but were completely unaware of the study’s primary aim to examine non-specific effects. The trial design included a nested acupuncture substudy that allowed potential participants in the “treatment” arms to be told, truthfully, that they had a 50% chance of receiving genuine acupuncture during the trial. All study personnel, except the practitioners, were blinded to participant assignment.

Outcome assessments

Our primary outcome was a change from baseline at three weeks in the global improvement scale which assesses change in symptoms over the past week.^{7,8} Our

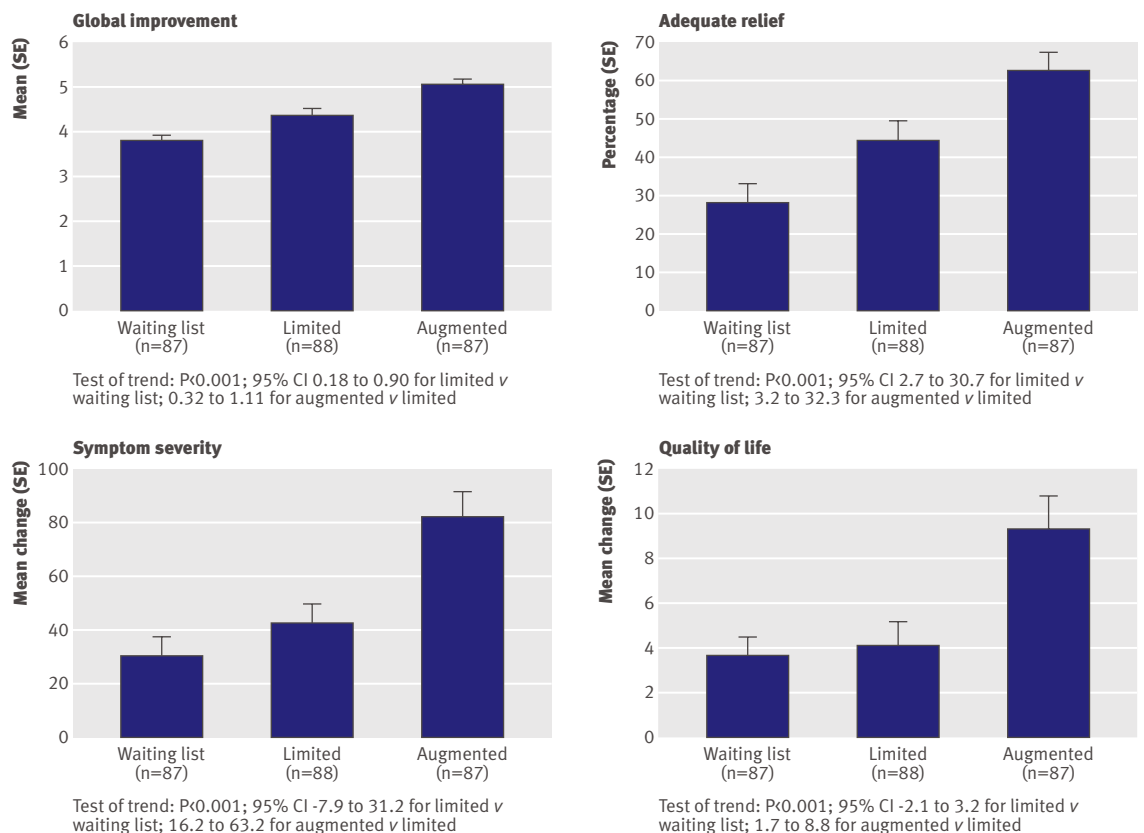


Fig 1 | Outcomes at three week end point

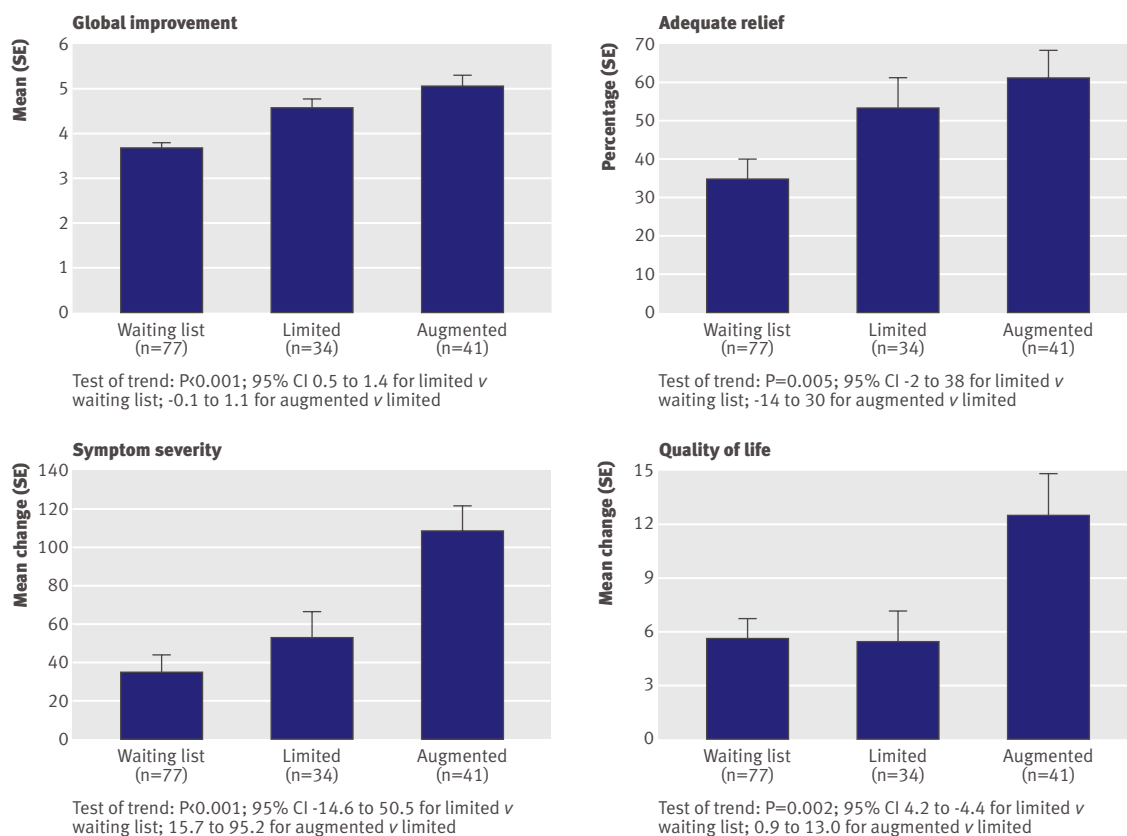


Fig 2 | Outcomes at six week follow-up

other main outcome was adequate relief, which assesses adequate relief of symptoms over the past week.^{9,10} Neither primary outcome was measured at baseline. Our other two outcomes were the symptom severity scale and the quality of life scale.^{11,12}

Statistical analysis

The primary test for each outcome measure was a test of trend examining the ordered alternative hypothesis, waiting list (group 1) < limited (group 2) < augmented (group 3). We considered $P < 0.0125$ (two sided) to be significant for each test of trend. If the trend test was significant we conducted pairwise comparisons of the groups—that is, augmented v limited and limited v waiting list. All analyses were carried out on an intention to treat basis.

Outcome measures in participants with irritable bowel syndrome at three weeks as mean (SD) unless stated otherwise

Outcome measure	Waiting list (n=87)	Limited (n=88)	Augmented (n=87)	P value for trend
Global improvement scale	3.8 (1.0)	4.3 (1.4)	5.0 (1.3)	<0.001
% with adequate relief of symptoms	28	44	62	<0.001
Change in symptom severity score	30 (63)	42 (67)	82 (89)	<0.001
Change in quality of life	3.6 (8.1)	4.1 (9.4)	9.3 (14.0)	0.001

RESULTS

Study population—Between December 2003 and February 2006, we screened 350 prospective participants of whom 289 were eligible. We randomised 262 people into the three groups. (Twenty seven people were randomised to an independent nested qualitative study.) At baseline the three groups were well balanced with regard to demographics, psychiatric symptoms, type of irritable bowel syndrome, and quality of life score.

Outcomes at three weeks—The observed values for all outcome measures were consistent with our prediction of a progressive improvement in symptoms among the three groups such that waiting list was less effective than limited, which was less effective than augmented (table). As indicated in figure 1 the test of trend for each of the outcome measures was significant ($P < 0.001$). For the global improvement scale and the adequate relief of symptoms, each of the pairwise comparisons (augmented v limited and limited v waiting list) was significant ($P < 0.001$). For the symptom severity score, the augmented group improved significantly more than the limited group ($P = 0.007$), but the limited and waiting list groups were not significantly different ($P = 0.20$). We observed the same pattern for quality of life ($P = 0.01$ and $P = 0.58$). The proportions of patients reporting moderate or substantial improvement on the global improvement scale were 3% (waiting list), 20% (limited), and 37% (augmented) ($P < 0.001$).

WHAT IS ALREADY KNOWN ON THIS TOPIC

In theory, the placebo effect of the clinical encounter can be divided into the response to three main components: assessment/observation, therapeutic ritual (placebo), and patient-physician relationship

WHAT THIS STUDY ADDS

Three components of the medical encounter can be progressively added to produce incremental improvement in symptoms

A therapeutic ritual (placebo treatment) has a modest benefit beyond no treatment

Placebo effects produce statistically and clinically significant improvement and the patient-physician relationship is the most robust component of the placebo effect

Outcomes at six weeks—As can be seen in figure 2, each of the tests for trend at week six was significant. Except for quality of life where improvement in the waiting list group was similar to that in the limited group, the observed values for all outcome measures were consistent with our a priori prediction of order of improvement.

Adverse effects—More than 80% of patients reported no side effects. The most common side effects included pain during needle placement (10%) and redness or swelling (6%) or pain (5%) after needle removal. Other adverse effects were reported rarely and were not serious.

DISCUSSION

In this large prospective study of placebo effects we found that such effects can be disentangled into three components that can then be recombined to produce incremental improvement in symptoms. We found that an enhanced relationship with a practitioner, together with the placebo treatment, provides the most robust effect. Placebo treatment with only limited interaction with practitioners was superior to staying on a waiting list with respect to only two of the four measures.

The magnitude of non-specific effects in the augmented arm is not only statistically significant but also clinically significant in the management of irritable bowel syndrome. A decrease in the symptom severity score of 50 reliably indicates improvement in symptoms,¹¹ and our study indicates that 61% and 59% of patients in the augmented arm achieved this level of improvement at three and six weeks, respectively. Likewise, the changes we observed in quality of life indicate at least moderate clinical improvement in symptoms.¹² Finally, the percentage of patients reporting adequate relief (62% and 61% at three and six weeks, respectively) is comparable with the responder rate in clinical trials of drugs currently used in the treatment of irritable bowel syndrome.^{13 14} These results indicate that such factors as warmth, empathy, duration of time, and the communication of positive expectation might indeed significantly affect clinical outcome.

Limitations

One limitation of our study is that we could not separate the effects of observation and assessment from natural history or spontaneous remission. Our outcome measures were subjective rather than objective. None the

less, these measures are consistent with the recommendations by the Rome committees for use in trials of irritable bowel syndrome because no objective measures of severity are currently available.¹⁵ We chose irritable bowel syndrome for this study because we suspected that non-specific effects are most likely to be demonstrable in disorders defined by subjective symptoms rather than more objective measures of disease.¹⁶ Whether our findings apply to other illnesses, including those with biochemical or other objective outcome measures, awaits further study. None the less, our study has important implications for routine clinical care.

It is unclear whether our placebo outcomes correspond to biological changes in irritable bowel syndrome or have any of the biochemical, neuroendocrine, or neuroanatomical correlates of placebo response found in recent laboratory experiments^{17 18} or whether our outcomes are mainly related to shifts in selective attention to diffuse symptoms.¹⁹ In either case, our study represents an incremental step in placebo studies and shows that non-specific effects have a considerable clinical impact.

We thank Franklin Miller, Kate Stony, and Jongbae Park for scientific mentorship; Mary Quilty, Oriana Rodrigues, Gabriel Kapchuk, Elizabeth Morey, and Patricia Wilkinson for research assistance; and the research nurses at the BIDMC, under the direction of Mary Williams and Jamie Vickers, and acupuncturists Stephanie Prady, Bella Rosner, and Lisa Desrosiers for all their hard work. We also thank J Thomas LaMont for his thoughtful review of the manuscript and the university seminar on effective and affordable health care at Harvard University for input on study design and analysis.

Contributors: See bmj.com.

Funding: NIH grant No 1R01 AT001414-01 from the National Center for Complementary and Alternative Medicine (NCCAM) and the National Institutes of Digestive, Diabetes and Kidney Disease (NIDDK), grant No 1R21 AT002860-01 from NCCAM and the Office of Behavioral and Social Science Research (OBSSR), and grant No 1R21 AT002564 and 1K24 AT004095 from NCCAM. This research was also supported in part by grant RR 01032 to the Beth Israel Deaconess Medical Center (BIDMC) General Clinical Research Center from the NIH.

Ethical approval: Institutional review boards at the Beth Israel Deaconess Medical Center and Harvard Medical School.

Competing interests: TK is a consultant for Kan Herbal Company, Scotts Valley, CA. AL has served on the scientific advisory boards and served as a consultant for Novartis, Takeda, Sucampo, Schwarz, Salix, Microbia, and GSK. PG is a consultant for Tsumura.

Provenance and peer review: Not commissioned; externally peer reviewed.

- 1 Kapchuk TJ. Powerful placebo: the dark side of the randomized controlled trial. *Lancet* 1998;351:1722-5.
- 2 Hrobjartsson A. What are the main methodological problems in the estimation of placebo effects. *J Clin Epidemiol* 2002;55:430-5.
- 3 Miller FG, Kapchuk TJ. The power of context: reconceptualizing the placebo effect. *J Roy Soc Med* 2008 (in press).
- 4 Patel SM, Stason WB, Legedza A, Ock SM, Kapchuk TJ, Conboy L, et al. The placebo effect irritable bowel syndrome trials: a meta-analysis. *Neurogastroenterol Motil* 2005;17:332-40.
- 5 Kapchuk TJ, Stason WB, Davis RB, Legedza ATR, Schnyer RS, Kerr CE, et al. Sham device v inert pill: randomised controlled trial of two placebo treatments. *BMJ* 2006;332:391-7.
- 6 Thompson WG, Longstreth GF, Drossman DA, Heaton KW, Irvine EJ, Muller-Lissner SA. Functional bowel disorders and functional abdominal pain. *Gut* 1999;45(suppl 2):II43-7.
- 7 Mangel AW, Hahn BA, Heath AT, Northcutt AR, Kong S, Dukes GE, et al. Adequate relief as an endpoint in clinical trials in irritable bowel syndrome. *J Int Med Res* 1998;26:76-81.
- 8 Mangel AW. Personal view: adequate relief as a primary endpoint in irritable bowel syndrome. *Aliment Pharmacol Ther* 2006;23:879-81.
- 9 Lembo AL, Wright RA, Bagby B, Decker C, Gordon S, Jhingran P, et al. Alosetron controls bowel urgency and provides global symptom improvement in women with diarrhea-predominant irritable bowel syndrome. *Am J Gastroenterol* 2001;96:2662-70.
- 10 Gordon S, Ameen V, Bagby B, Shahan B, Jhingran P, Carter E. Validation of irritable bowel syndrome Global Improvement Scale: an

- integrated symptom end point for assessing treatment efficacy. *Dig Dis Sci* 2003;48:1317-23.
- 11 Francis CY, J. Morris J, Whorwell PJ. The irritable bowel severity scoring system: a simple method of monitoring irritable bowel syndrome and its progress. *Alimen Pharmacol Ther* 1997;11:395-402.
 - 12 Drossman DA, Patrick DL, Whitehead WE, Toner BB, Diamant NE, Hu Y, et al. Further validation of the IBS-QOL: a disease-specific quality-of-life questionnaire. *Am J Gastroenterol* 2000;95:999-1007.
 - 13 Camilleri M, Northcutt AR, Kong S, Dukes GE, McSorley D, Mangel AW. Efficacy and safety of alosetron in women with irritable bowel syndrome: a randomised, placebo-controlled trial. *Lancet* 2000;355:1035-40.
 - 14 Novick J, Miner P, Krause R, Glebas K, Bliesath H, Ligozio G, et al. A randomized, double-blind, placebo-controlled trial of tegaserod in female patients suffering from irritable bowel syndrome with constipation. *Aliment Pharmacol Ther* 2002;16:1877-88.
 - 15 Camilleri M, Mangel AW, Fehnel SE, Drossman DA, Mayer EA, Talley NJ. Primary endpoints for irritable bowel syndrome trials: a review of performance of endpoints. *Clin Gastroenterol Hepatol* 2007;5:534-40.
 - 16 Kaptchuk TJ. The placebo effect in alternative medicine: can the performance of a healing ritual have clinical significance? *Ann Intern Med* 2002;136:817-25.
 - 17 Benedetti F, Mayberg HS, Wager TD, Stohler CS, Zubieta JK. Neurobiological mechanisms of the placebo effect. *J Neurosci* 2005;25:10390-402.
 - 18 Pacheco-López, G, Engler H, Niemi MB, Schedlowski M. Expectations and associations that heal: immunomodulatory placebo effects and its neurobiology. *Brain Behav Immunol* 2006;403-446.
 - 19 Allan LG, Siegel S. A signal detection theory analysis of the placebo effect. *Eval Health Prof* 2002;25:410-20.

Accepted: 2 March 2008

Diagnostic accuracy of urinary spot protein:creatinine ratio for proteinuria in hypertensive pregnant women: systematic review

Anne-Marie Côté,¹ Mark A Brown,² Elaine Lam,¹ Peter von Dadelszen,³ Tabassum Firoz,¹ Robert M Liston,³ Laura A Magee³

EDITORIAL by Chappell and Shennan

¹BC Women's Hospital and Health Centre, Vancouver, BC, Canada

²St George Hospital and University of New South Wales, Kogarah, Sydney, NSW, Australia

³Child and Family Research Institute, University of British Columbia, Vancouver, BC, Canada V6H 3N1

Correspondence to: L A Magee, BC Women's Hospital and Health Centre, Vancouver, BC, Canada V6P 1S8 L.Magee@cw.bc.ca

BMJ 2008;336:1003-6

doi:10.1136/bmj.39532.543947.BE

This article is an abridged version of a paper that was published on bmj.com. Cite this article as: BMJ 2008, doi: 10.1136/bmj.39532.543947.BE

ABSTRACT

Objective To review the spot protein:creatinine ratio and albumin:creatinine ratio as diagnostic tests for significant proteinuria in hypertensive pregnant women.

Design Systematic review.

Data sources Medline and Embase, the Cochrane Library, reference lists, and experts.

Review methods Literature search (1980-2007) for articles of the spot protein:creatinine ratio or albumin:creatinine ratio in hypertensive pregnancy, with 24 hour proteinuria as the comparator.

Results 13 studies concerned the spot protein:creatinine ratio (1214 women with primarily gestational hypertension). Nine studies reported sensitivity and specificity for eight cut-off points, median 24 mg/mmol (range 17-57 mg/mmol; 0.15-0.50 mg/mg). Laboratory assays were not well described. Diagnostic test characteristics were recalculated for a cut-off point of 30 mg/mmol. No significant heterogeneity in cut-off points was found between studies over a range of proteinuria. Pooled values gave a sensitivity of 83.6% (95% confidence interval 77.5% to 89.7%), specificity of 76.3% (72.6% to 80.0%), positive likelihood ratio of 3.53 (2.83 to 4.49), and negative likelihood ratio of 0.21 (0.13 to 0.31) (nine studies, 1003 women). Two studies of the spot albumin:creatinine ratio (225 women) found optimal cut-off points of 2 mg/mmol for proteinuria of 0.3 g/day or more and 27 mg/mmol for albuminuria.

Conclusion The spot protein:creatinine ratio is a reasonable "rule-out" test for detecting proteinuria of 0.3 g/day or more in hypertensive pregnancy. Information on use of the spot albumin:creatinine ratio in these women is insufficient.

INTRODUCTION

Urine collection over 24 hours is considered the traditional comparator for quantification of proteinuria

in pregnancy,¹⁻⁴ but it has limitations: the urine requires refrigeration; collection is cumbersome, time consuming, and potentially misleading if done inaccurately; and collection may not be possible during delivery. Timed collections also delay diagnosis and may result in prolonged hospital stay for investigations.

Alternatives for the diagnosis of proteinuria in pregnancy have been considered. These include urinary dipsticks, urine collections over a shorter period, the urinary spot protein:creatinine ratio, and the urinary spot albumin:creatinine ratio. Australasian and international guidelines advocate use of the urinary spot protein:creatinine ratio as an alternative to 24 hour urine collection.^{3,4}

We carried out a systematic review to assess the accuracy of the spot protein:creatinine ratio and spot albumin:creatinine ratio compared with 24 hour urinary collection for the detection of significant proteinuria in hypertensive pregnant women.

METHODS

We did a search (see bmj.com) of diagnostic studies that compared the urinary spot protein:creatinine ratio or albumin:creatinine ratio with urinary protein excretion over 24 hours among hypertensive pregnant women.

We assessed study quality using the quality assessment of studies of diagnostic accuracy in a systematic reviews tool.⁵

Information was recorded on characteristics of the study and participants, how the diagnostic tests were carried out and the results, and methods for assessing the diagnostic accuracy of the tests. When positive and negative likelihood ratios were missing we calculated them from sensitivity and specificity (see bmj.com).

For each study we analysed sensitivity, specificity, likelihood ratios, and area under the receiver operating curves, according to reported cut-off points (we used a conversion factor of 1.13 to transform results for the spot