Research methodology in rheumatology

Vivian Robinson
MSc., Centre for Global Health, Institute for Population Health, University of Ottawa, 1 Stewart Street, Room 312, Ottawa, Ontario

Peter Tugwell
MD, MSc, FRCPC, Chair, Centre for Global Health, Institute for Population Health, University of Ottawa, 1 Stewart Street, Room 312, Ottawa, Ontario

Maria Judd
BSc PT, Centre for Global Health, Institute for Population Health, University of Ottawa, 1 Stewart Street, Room 312, Ottawa, Ontario

Beverley Shea
BScN, MSc, Institute for Population Health, University of Ottawa, 1 Stewart Street, Room 312, Ottawa, Ontario

George Wells
MSc, PhD, Director, Department of Community Health and Epidemiology, University of Ottawa, 451 Smyth Road, Ottawa, Ontario

The evidence-based medicine research community is increasingly recognizing that the translation of evidence into clinical practice is an important research agenda. This paper describes research methods related to transferring knowledge to different audiences: clinicians, consumers and policy-makers. The iterative measurement loop (IML) described in 1987 (Tugwell 1987) highlights the 6 steps of evidence-based research that leads to the translation of evidence into practice and re-evaluation of the impact of evidence-based medicine (Figure 1). This paper describes research methods for assessing steps of the IML with the goal of improving knowledge translation and uptake of evidence-based medicine.

Evidence-based medicine has been defined as the process of systematically finding, appraising, and using contemporaneous research findings as the basis for clinical decisions. Evidence-based medicine asks questions, finds and appraises the relevant data, and harnesses that information for everyday clinical practice. Evidence-based medicine follows four steps: formulate a clear clinical question from a patient's problem; search the literature for relevant clinical articles; evaluate (critically appraise) the evidence for its validity and usefulness; implement useful findings in clinical practice (Rosenberg 1995)

The challenge of evidence-based medicine is to allow clinicians and consumers to quantify the impact of an intervention on individual patients. The steps described in this paper facilitate the acknowledgment of patient values in evidence-based decision making and the transfer of this knowledge to patients and clinicians.

Step 1: Burden of illness

The first step in the IML is to evaluate the burden of illness and assess the need for evidence-based tools. Arthritis and musculoskeletal disease are among the top 10 contributors to global burden of illness (World Bank Report 1996).

Step 2: Assessing effectiveness

To assess efficacy and toxicity in a clinical setting agreement on a core set of outcomes is essen-

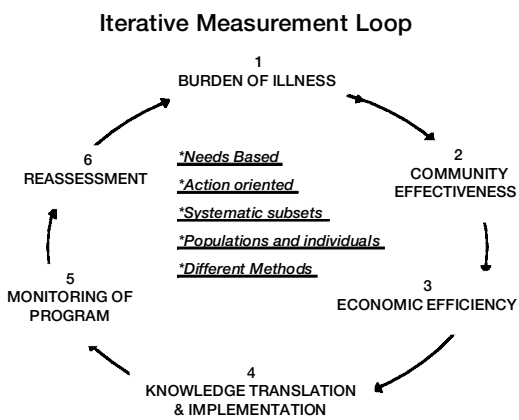

Figure 1. Iterative Measurement Loop

OMERACT conferences

- **OMERACT 1**: WHO/ILAR core set for rheumatoid arthritis trials
- **OMERACT 2**: cost-effectiveness
- **OMERACT 3**: core sets osteoarthritis, osteoporosis; psychosocial measures
- **OMERACT 4**: core sets SLE, AS; core set for longitudinal & observational studies RA: response criteria, imaging
- **OMERACT 5**: Economics reference case, minimum clinically important difference (MCID), Imaging, Safety

Figure 2. OMERACT conferences

tial. The Outcome Measures in Rheumatology (OMERACT) conferences were initiated in 1987 to help achieve international consensus on a core set of outcome measures.

The OMERACT conferences assessed all available measures against the OMERACT Filter for validity, discriminate and feasibility (Boers 1998). Each OMERACT conference consists of preconference mailouts and surveys that are used to gather opinions from the research community. At the conference, the focus is on small group discussion about areas of contention. The conferences end with interactive voting on consensus outcomes. The five OMERACT conferences developed potential core sets for rheumatoid arthritis, osteoarthritis, osteoporosis, SLE, ankylosing spondylitis, longitudinal and observational studies and imaging criteria for rheumatology (Figure 2). OMERACT has Task Forces which are working on achieving consensus for measurement of safety, assessment of minimal clinically important difference, quality of life and health economics. OMERACT 6 (April 2002) had two modules: 1) Economics; and 2) Imaging/MRI.

To determine effectiveness, both randomized controlled trials and observational "real-life" studies are needed . Randomized controlled trials (or systematic review of all available RCTs) are considered the highest level evidence for assessing efficacy in a clinical trial setting. However, there is increasing recognition that observational "real-life" studies are an important source of real-life safety and community effectiveness.

The Cochrane Collaboration was created to help clinicians, consumers and policy makers stay abreast of the latest literature by creating a database of up-to-date systematic reviews of all available controlled clinical trials. The Cochrane Musculoskeletal review group has now published over 50 systematic reviews of efficacy and toxicity of treatments for musculoskeletal disorders. These reviews can be used to rank-order treatments in terms of their effectiveness (Figure 3) (Tugwell 2000).

The way in which evidence is presented affects the interpretation of results. A survey of physicians found that presenting relative risks made an intervention appear more favourable. This survey of 100 physicians presented the effects of a cholesterol-lowering drug as both absolute risk (prevents 1 out of 71 people from having a heart attack) and as relative risk (33% fewer heart attacks with the drug). Of those surveyed, 88% chose to use the drug based on relative risk, but only 31% chose to use the drug when shown absolute risk (Oxman 2000).

One way tu handle to differences in interpretation is to show both relative and absolute risk

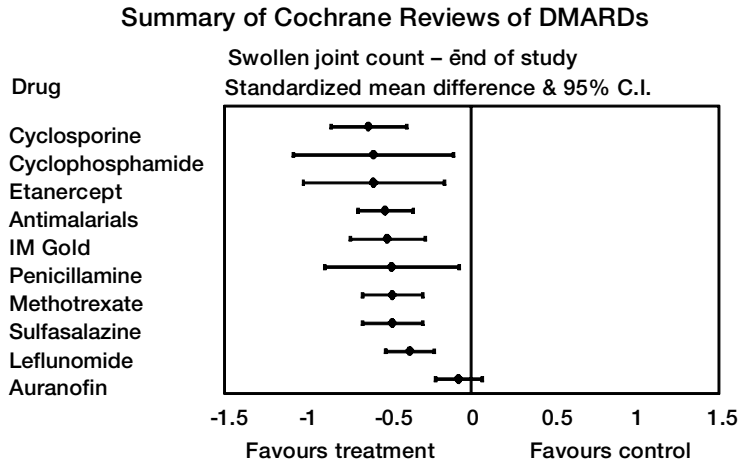

Figure 3. Rank-ordering of disease-modifying anti-rheumatic drugs using Cochrane review estimates.

Table 1. Presentation of both absolute and relative percentage change

Study	Treatment group	Outcome (scale)	No of patients	Baseline mean	End of study mean	Absolute benefit	Relative difference in change from baseline
Frost 95	E: Strength, stretch, aerobic exercises	Sensory Pain (0-100 VAS)	36	20.9	12.1	-5.30 (I) on 100 VAS	-23% (I)
	C: Control		35	25.6	22.1		
Hansen 93	E: Strength exercises	Pain (0-9 VAS)	44	5.0	4.1	-3.00 (I) on 9 point Likert	-60% (I)
	C: Control		28	5.0	7.1		
Risch 93	E: Strength, stretch exercises	Pain (West-Haven	31	3.4	2.9	-0.90 (I) on 25 point scale	-26% (I)
	C: Control	Yale, 0-25)	23	3.7	4.1		

(I): indicates improvement better in treated group than control, negative or positive depending on anchors for the scales

for all estimates of efficacy. Our group recently participated in the development of evidence-based clinical practice guidelines for selected rehabilitation interventions. The evidence was identified and synthesized using Cochrane Collaboration methodology, then the results were presented as both relative percentage change and absolute change (Philadelphia Panel 2001) (Table 1).

The Omeract Minimal Clinically important difference task force (Wells 2001) is working on defining methods for measuring outcomes that include relative, absolute and definition of state. This involves defining a 'low disease activity state' that can be considered as a goal for therapy. This target state would be defined by targets for each outcome or a set or core outcomes (eg number of tender joints, pain on VAS, ESR). This state could be used as a therapeutic goal and also as a discrete outcome in clinical trials (i.e. number of patients meeting the criteria for low disease activity). This topic has been discussed in workshop format at OMERACT 6.

Step 3: Assessing cost-effectiveness

Cost-effectiveness is increasingly important for regulatory approval of new, costly therapies for rheumatic diseases. It is essential to develop standards for economic studies in order to facilitate between study comparisons. The Omeract Economics Task Force is working towards developing a 'reference case' or standard methods for rheumatology economic studies (Gabriel 2001). The components of this reference case are: model horizon, duration of therapy, modeling beyond trial or therapy duration, synthesis of head to head data, outcome measures, valuation of health (including QALYs), classification of adverse events, discontinuation and compliance and therapeutic strategies vs single drugs. A draft version of this reference case was developed at a working meeting in New York, February 2001 (Gabriel 2002). Omeract 6 (April 11–14, 2002) had a module on developing standards for economic studies. This module will focus on clinical outcomes and methods that are used to estimate the costs and benefits of rheumatology therapies. The goal is to allow greater comparability between different economic studies.

Step 4: Knowledge translation

Knowledge has been defined as information in the mind, in a context which allows it to be transformed into action. Clinical recommendations are one way to incorporate evidence or knowledge into clinical practice. The development of recommendations has evolved so that guidelines have been published on developing recommendations based on transparent, clear and rigorous methods (Shekelle 1999).

Making evidence-based recommendations involves classifying both the strength of the evidence and the quality of the evidence available. Several methods have been developed for assessing the quality and strength of evidence

Table 2. Philadelphia Panel grading system for quality and strength of evidence

	Clinical importance	Statistical significance	Study design
Grade A	≥15% relative difference in change from baseline	p<0.05	RCT (single or meta-analysis)
Grade B	≥15% relative difference in change from baseline	p<0.05	CCT or observational (single or meta-analysis), with a quality score of 3 or more the 5 point Jadad's methodologic quality scale
Grade C+	≥15% relative difference in change from baseline	Not significant	RCT or CCT or observational (single or meta-analysis)
Grade C	<15% relative difference in change from baseline	Unimportant*	Any study design
Grade D	<0% (favours control)		Well-designed RCT with >100 patients

and presenting it to clinicians, policy-makers and consumers.

The evidence-based BMJ books and Clinical evidence (www.clinicalevidence.org) use a 3 step rating system of A, B and C where A denotes evidence from a large RCT, systematic review or high quality "all or "none" cohort study (eg where all died/failed with conventional and some survived/succeeded; B denotes at least one high quality study of non-randomized cohorts who did/did not receive new therapy, high quality case control study or high quality case series; and C denotes opinions from experts (eg bench studies, first principles).

Another set of recent guidelines developed by our group for musculoskeletal disorders (Philadelphia Panel 2001) modified the grading system of the Canadian Task Force on Preventive Medicine (Canadian Task Force 1986). This system grades both the quality of the evidence and the strength of the evidence (Table 2). In this system, the grade of recommendation reflects the results as well as the quality of evidence on which this conclusion is based. For example, a grade A recommendation indicates that the intervention results in greater than 15% relative difference in change from baseline between treated and untreated groups, is statistically significant and is based on high-level evidence from at least one randomized controlled trial.

The Ottawa Health Research Institute is a leader in developing consumer decision aids (www.ohri.ca). Work is ongoing to develop systems of explaining to consumers the quality of evidence available. Cranney, O'Connor, Stacey and colleagues developed the "ribbons scale", depicted in Figure 4. This rates evidence as "gold" (randomized trials), "silver" (observational- weaker evidence) and "bronze" (expert opinion based on experience, knowledge of specific cases or reports) and "laboratory studies" (evidence that has not yet been tested in humans) (Figure 4).

Decision aids allow the target user (e.g. consumer or clinician) to judge the benefits and risks based on the seriousness of the outcome, magnitude of treatment effect, precision of treatment effect, risk of the target event, risk of serious adverse events and costs. Decision aids allow the user to incorporate their values into the decision making process. Risks and benefits are presented in the way preferred by the target user. For example, a survey of 370 women taking HRT found that over 60% of women preferred information about risks and benefits to be given as percentages (Michaud et al 1999).

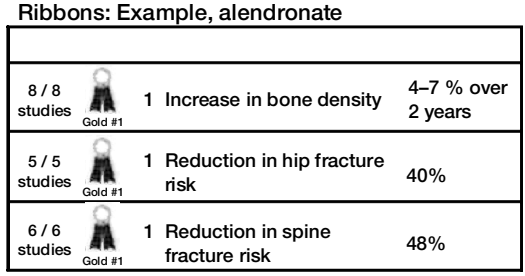

Figure 4. Ribbons scale example.

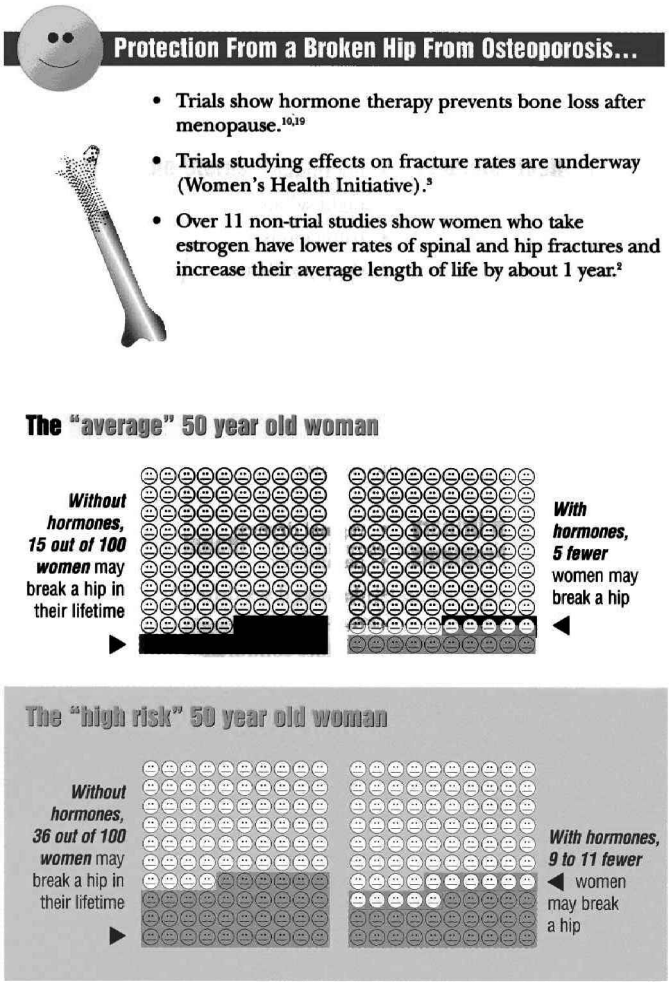

Figure 5. Hormone replacement therapy decision aid

"Making Choices: Hormones after Menopause" was one of the first Ottawa Health Research Institute decision aids. The risks and benefits are presented using figures that depict the risk of events (Figure 5). This decision aid includes a patient worksheet that asks the user the give weights to their value of benefits and risks as well as to answer questions about who should make decisions and their current health practices and context. Sixteen decision aids developed by members of the Ottawa Health Decision Centre (OHDEC) are available on the OHRI website (www.ohri.ca).

The Cochrane Musculoskeletal Review Group is working with a very active consumer group to develop consumer lay summaries of all of the Musculoskeletal Cochrane reviews. These lay summaries follow the format of a structured abstract. The numerical risks and benefits are presented as percentage improvement to allow consumers to gauge the amount of benefit they can expect with each therapy. These summaries are available on the web at www.arthritis.ca.

Dissemination and implementation are the two most important steps if evidence-based guidelines and decision aids are to be transferred into daily use. Cochrane reviews of strategies to change physician behaviour have shown that mailed instructional material and lecture based CME are weak methods of changing behaviour. Targeting opinion leaders and using audit and feedback are moderately successful. The best methods of changing physician behaviour are using a reminder system, academic detailing and using more than one of the above approaches. Several factors affect the likelihood of changing physician behaviour including: qualities of the guidelines, the health care professional, practice setting, incentives, regulation, and patient factors.

The Ottawa Ankle Rules are one example of a guideline that has been implemented with a structured approach to improve and assess its uptake by clinicians. The development team conducted a use and attitudes survey to determine the need (N=1,831) (Stiell, McDowell et al. 1992), a regression approach to develop the Rules (N=750) (Stiell, Greenberg et al. 1992), implementation trial (N=2,342) (Stiell 1994), multicentre implementation (N=12,777) (Stiell 1995) and cost-effectiveness analysis (Anis 1995).

Steps 5 and 6: Monitoring and reassessment

To determine whether these methods of knowledge translation have been effective, the fifth and sixth steps of the IML need to be conducted, i.e. moni-

toring and reassessment. Outcomes for such evaluation projects may relate to the process of care (e.g. use of hormone replacement therapy, prescribing patterns and use of radiography) or health care outcomes (e.g. fracture rates, cholesterol levels, costs). For example, an implementation study of the Ottawa Ankle Rules resulted in a 12% decrease in ankle radiography from 82.8% to 60.9% in community hospitals (Stiell 1995). Furthermore, a follow-up study of another implementation found that the reduced use of radiography was sustained after the implementation study (Verbeek 1997).

Conclusion

One of the greatest challenges for the development of evidence-based medicine is the continued development of the friendly-front-end. This involves developing useable decision aids for different audiences. Towards this end, we are working on a BMJ book of evidence-based rheumatology.

Methodology challenges include developing sound guidelines for the transfer of evidence-based knowledge into practice and the translation of evidence that is easily understood and put into practice by our clinical colleagues, consumers and policy makers. Another challenge is the development of rigorous methods for considering data from "real-life", observational studies of clinical effectiveness.

One of the greatest challenges to improving the development of the research is to involve transdisciplinary teams from the onset of the research from having consumers, clinicians, policy makers and health care decision makers at the table to eventually working with these target audiences in training initiatives that provide them with an environment where they are not only told results but actually taught to interpret and act upon these results and eventually turn evidence-based decision making into a reality for all.

References

Anis AH, Stiell IG, Stewart DG, Laupacis A. Cost-effectiveness analysis of the Ottawa Ankle Rules. Ann Emerg Med 1995; 26(4): 422-8.

Bellamy N, Kirwan J, Boers M, Brooks P, Strand V, Tugwell P, Altman R, Brandt K, Dougados M, Lequesne M. Recommendations for a core set of outcome measures for future phase III clinical trials in knee, hip, and hand osteoarthritis. Consensus development at OMERACT III. J Rheumatol 1997; 24(4): 799-802.

Boers M, Brooks P, Strand CV, Tugwell P. The OMERACT filter for Outcome Measures in Rheumatology. J Rheumatol. 1998; 25(2): 198-9.

Canadian Task Force on the Periodic Health Examination. The periodic health examination: 1986 update. Can Med Assoc J 1986; 134(721):729.

Clinical evidence. www.clinicalevidence.org

Cranney A, Tugwell P, Cummings S, Sambrook P, Adachi J, Silman AJ, Gillespie WJ, Felson DT, Shea B, Wells G. Osteoporosis clinical trials endpoints: candidate variables and clinimetric properties. J Rheumatol 1997; 24(6): 1222-9.

Felson DT, Anderson JJ, Boers M, Bombardier C, Chernoff M, Fried B, Furst D, Goldsmith C, Kieszak S, Lightfoot R, et al. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. The Committee on Outcome Measures in Rheumatoid Arthritis Clinical Trials. Arthritis Rheum 1993; 36(6): 729-40.

Gabriel S, Drummond M, Suarez-Almazor ME, Ruff B, Guillemin F, Bombadier C, Boers M, Maetzel A, Ruof J, Cranney A, Marentette M, Tugwell P. OMERACT 5 Economics Working Group: summary, recommendations, and research agenda. J Rheumatol 2001; 28(3): 670-3.

Gabriel SE, Tugwell P, Drummond M. Progress Towards An OMERACT-ILAR Guideline For Economic Evaluations In Rheumatology: Working Group Report. Annals Rheum Dis, In press 2002.

Michaud et al., 1999

O'Connor AM, Tugwell P, Wells GA, Elmslie T, Jolly E, Hollingworth G, McPherson R, Bunn H, Graham I, Drake E. A decision aid for women considering hormone therapy after menopause: decision support framework and evaluation. Patient Educ Couns 1998; 33(3): 267-79.

O'Connor AM, Tugwell P, Wells GA, Elmslie T, Jolly E, Hollingworth G, McPherson R, Drake E, Hopman W, Mackenzie T. Randomized trial of a portable, self-administered decision aid for postmenopausal women considering long-term preventive hormone therapy. Med Decis Making 1998; 18(3): 295-303.

Oxman A., personal communication.

Philadelphia Panel, Ottawa Methods Group. Philadelphia Panel evidence-based clinical practice guidelines on selected rehabilitation interventions: overview and methodology. Phys Ther 2001; 81(10): 1629-40.

Rosenberg W, Donald A. Evidence based medicine: an approach to clinical problem-solving. BMJ 1995; 310(6987): 1122-6.

Shekelle PG, Woolf SH, Eccles M, Grimshaw J. Clinical guidelines: developing guidelines. BMJ 1999; 318(7183): 593-6.

Stiell I, Wells G, Laupacis A, Brison R, Verbeek R, Vandemheen K, Naylor CD. Multicentre trial to introduce the Ottawa ankle rules for use of radiography in acute ankle injuries. Multicentre Ankle Rule Study Group. BMJ 1995; 2;311(7005): 594-7

Stiell IG, Greenberg GH, McKnight RD, Nair RC, McDowell I, Worthington JR. A study to develop clinical decision rules for the use of radiography in acute ankle injuries. Ann Emerg Med 1992; 21(4): 384-90.

Stiell IG, McDowell I, Nair RC, Aeta H, Greenberg G, McKnight RD, Ahuja J. Use of radiography in acute ankle injuries: physicians' attitudes and practice. CMAJ 1992; 147(11): 1671-8.

Stiell IG, McKnight RD, Greenberg GH, McDowell I, Nair RC, Wells GA, Johns C, Worthington JR. Implementation of the Ottawa ankle rules. JAMA 1994; 271(11): 827-32.

Tugwell P, Bennett KJ, Sackett DL, Haynes RB. The measurement iterative loop: a framework for the critical appraisal of need, benefits and costs of health interventions. J Chronic Dis 1985; 38(4):339-51.

Tugwell P, Welch V, Suarez-Almazor M, Shea B, Wells G. Efficacy and toxicity of old and new disease modifying antirheumatic drugs. Ann Rheum Dis. 2000 Nov;59 Suppl 1:i32-5.

Verbeek PR, Stiell IG, Hebert G, Sellens C. Ankle radiograph utilization after learning a decision rule: a 12-month follow-up. Acad Emerg Med 1997; 4(8): 776-9.

Wells G, Anderson J, Beaton D, Bellamy N, Boers M, Bombardier C, Breedveld F, Carr A, Cranney A, Dougados M, Felson D, Kirwan J, Schiff M, Shea B, Simon L, Smolen J, Strand V, Tugwell P, van Riel P, Welch VA. Minimal clinically important difference module: summary, recommendations, and research agenda. J Rheumatol 2001; 28(2): 452-4.

World Bank Annual Report, 1996, Washington, DC.