

Evaluating the effectiveness of using PROs



Department of Public Health
University of Aberdeen, UK



NTNU
Trondheim, Norway

Daily diary card 1984



UNIVERSITY of ABERDEEN

2

NTNU Trondheim, Norway

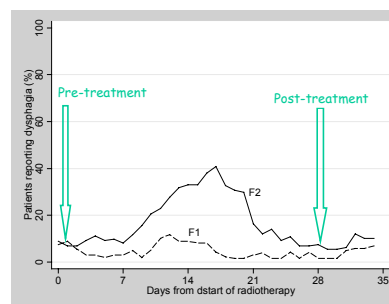
Daily diary card

UNIVERSITY of ABERDEEN

3

NTNU Trondheim, Norway

RCT of radiotherapy for non-small-cell lung cancer



MRC, *British Journal of Cancer*, 65 (1992) 934-941

UNIVERSITY of ABERDEEN

4

NTNU Trondheim, Norway

Diary cards

- Compliance varied hugely.
- Physician related compliance?
- Some physicians enthused about QoL cards, and started using for all their patients.
- Many patients used the "other" box to communicate that they were glad someone was interested in how they felt.
- Some patients asked to continue completing cards beyond the end of the study.

UNIVERSITY of ABERDEEN

5

NTNU Trondheim, Norway

Diary card

- In 1980s, QoL assessment was novel.
- In 1980s, physicians were mainly concerned with clinical observations, and less with patients' feelings.
- Assessment using diary card is intensive.
- To improve compliance, clinicians were encouraged to discuss results with patients ...
- In 1980s, much anecdotal evidence that some patients and some clinicians found QoL assessment rewarding.
- But statisticians don't do anecdotes.

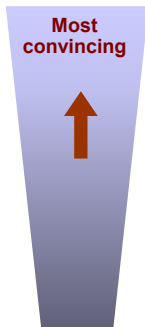
UNIVERSITY of ABERDEEN

6

NTNU Trondheim, Norway

Quality of evidence

- **Systematic reviews and meta analysis**
- **Randomised trials**
 - Two parallel groups
 - Cross-over trials
 - Cluster-randomised trials
- **Non-randomised studies**
 - Cross-sectional observational studies
 - Longitudinal studies
 - Cohort studies
 - Case-control studies
- **Anecdotal “evidence”**



Testable Outcomes

- **Improved health outcomes**
- **Improved patient satisfaction**
- **Clinical efficiency / use of resources**

(*communication* is a subsidiary endpoint for satisfaction and possibly health outcomes)



Randomized trials – systematic reviews

- **Randomized studies suggested health status reports improved communication but small impact on functional status.**

Velikova & Wright (2005), in: *Assessing Quality of Life in Clinical Trials: Methods and Practice*, 2nd edn., (ed. Fayers PM, Hays RD)

- **Systematic review of Greenhalg & Meadows 1999 found improved detection of psychological and functional problems, but little evidence of improved outcomes.**

Greenhalg J, Meadows KJ. *Eval Clin Pract* 5 (1999) 401–416

- **Systematic review of Espallargues et al 2000 found similar conclusions.**

Espallargues M, Valderas JM, Alonso J. *Med Care* 38 (2000) 175–186



Randomized trials

- **Feedback of HRQL scores to clinicians increases extent of discussion of HRQL issues.**
- **Little influence on decisions concerning treatment.**
- **Little impact on HRQL or satisfaction with care.**
 - Exception: Velikova et al., 2004

Greenhalg et al. *Social Science & Medicine* 60 (2005) 833–843



- **286 randomized oncology patients**
- **28 oncologists**
- **EORTC QLQ-C30 & HADS**
 - HRQL assessment + feedback to clinicians
 - HRQL assessment / no feedback
 - No HRQL assessment
- **Patients in both HRQL groups had better QoL. Routine assessment of HRQL improved physician-patient communication, and resulted in benefits for some patients.**
- **Nested hierarchical analysis was used with ‘doctor’ as a random effect.**

Velikova et al. *J Clin Oncol* 22 (2004) 714–724.



Cross-over study

- **10 physicians assigned to intervention or control.**
- **273 patients.**
- **Midway through the study the physicians crossed over between intervention / control.**
- **EORTC QLQ-C30.**
- **Intervention increased communication, had “only a modest effect on patient management” (mainly counselling), increased satisfaction with emotional support, and modest effect on QoL.**

Detmar et al. *JAMA* 288 (2002) 3027–3034.



Cross-over study

- Cross-over design intended to neutralize any effect that might be attributed to physicians' background characteristics.
- Also, since half of the physicians used the intervention first, order effects should largely cancel out.
- However, there is a risk of "carry-over" or contamination effect.
- The authors found some evidence suggesting that there was a learning effect from the intervention period.

Detmar et al. *JAMA* 288 (2002) 3027-3034.



Physician differences

- Clayton *et al.* randomized 174 patients. Their analysis used a **mixed model**, in which physicians were treated as a random effect.
- PROs (a 'prompt list') "stimulated discussion during the consultation, including about prognosis and end-of-life issues".
Clayton et al. *J Clin Oncol* 25 (2007) 715-723.
- Brown *et al.* randomized 318 patients to receiving 'prompt sheet' or not, and also randomized 9 doctors to be 'passive' or 'proactive'.
- "Prognosis was the only category in which question asking was influenced by the provision of a question prompt sheet."

Brown et al. *Br J Cancer* 85 (2001) 1273-1279.



Cluster-randomised trial

- Individual-patient randomisation
 - Impractical:
 - Patients talk to each other – those randomised to no assessment may complain.
 - Contamination:
 - Staff trained to use PRO data cannot simply forget their training.
 - Contamination dilutes observed intervention effects.
- Cluster randomisation:
 - Randomise "clusters of patients"
 - e.g., clinics or GPs
 - Cluster randomisation is widely used for evaluation of healthcare policies.
 - Cluster randomisation minimises contamination effects



Cluster randomisation

- Outcomes for patients within a cluster are likely to be more similar to each other than to outcomes from other clusters:
 - GPs and clinics may have different catchment areas => patient selection.
 - Physicians may vary in their enthusiasm and style of management.
 - Individuals within clusters may interact.
- "intracluster correlation".



Cluster randomisation

- Analysis of cluster randomised trials must allow for intracluster correlation.
- The primary unit of analysis = unit of randomisation = the cluster
- Sample size estimation, like analysis, must allow for the cluster design.
 - Ideal sample size: many clusters, possibly with few patients per cluster.
 - Less ideal: fewer clusters, each with many patients.



MECCA – Cluster-randomised trial

- European multi-centre trial.
- Schizophrenia & related disorders
- 134 clinicians randomised
- 451 patients recruited
- Intervention: Every 2 months for one year, patients rated their QoL, satisfaction with treatment, and requests for support.
 - Mental health, physical health, accommodation, job situation, leisure activities, friendships, relationship with family/partner, personal safety, practical help, psychological help and medication.
 - Clinicians trained to use ratings to facilitate a dialogue with the patients.
- Primary outcome: subjective quality of life at 12 months.
- Secondary outcomes: unmet needs, treatment satisfaction.
- Hypothesis: intervention will stimulate and promote a positive therapeutic dialogue and lead to a more favourable outcome.
- Analysis: Mixed effects model (random effects for centre and key-worker).

Priebe S, et al. *British Journal of Psychiatry* (in press)



MECCA – Cluster-randomised trial

- **Conclusions: Structuring patient-clinician dialogue to focus on patients' views positively influenced quality of life, needs for care and treatment satisfaction.**

Priebe S, et al. *British Journal of Psychiatry* (in press)

- **Analysis?**
- **The method of analyses described for this trial are unclear – ideally, presentation should follow the CONSORT guidelines for cluster RCTs.**

CONSORT statement: extension to cluster randomised trials. *BMJ* 328 (2004) 702–708



Cluster randomised trials

- **Aim to allow for contamination.**
- **Contamination is likely to weaken the effects observed in individual-patient RCTs.**
- **Allocation *concealment* is a problem in cluster RCTs – must use careful patient registration procedures.**
- **Cluster RCTs “less efficient” in as much as many clusters of patients are required.**

Torgerson D. *BMJ* 322 (2001) 355–357.



Other outcomes

- **Outcomes**
 - Improved health outcomes
 - Improved patient satisfaction
 - Clinical efficiency / use of resources
 - *Several studies have shown that using PRO data does not increase physician workload, but can PRO data improve efficiency?*



Summary, and ...

- **Effects may be weakened or even obscured by “contamination”**
- **Study design and analysis should make allowance for contamination**
- **Statistical design & analysis**
 - Mixed models / hierarchical mixed-effect models – e.g., physician as a random effect
- **Cluster-randomization (e.g., by physician) may offer a way forward**



... Comments

- **Many physicians are very positive about clinical usefulness of the data, and their patients are also in general positive.**
- **Despite this, many studies fail to find evidence of benefits.**
- **Many disease-specific HRQL instruments – such as EORTC QLQ-C30 – ask about symptom *severity* (“Have you had ...”), rather than *impact* (“How much are you troubled by ...”).**
- **Few of the instruments developed for RCTs provide an “other comments” box.**
- **Existing instruments may be suboptimal for individual-patient management**
- **Effects may be weak because good physicians probably collect relevant PRO data by asking the patient, anyway, ...**
- **... and modern medical education emphasizes communication skills and clinician-patient interaction, and involving patients in decisions.**
 - *The better the physician, the less the added value of formal HRQL instruments.*
- **New generation of patients expect their opinions and feelings to be considered.**
- **Maybe a diary – as widely used in allergy and asthma – would show clearer effects.**
- **Perhaps we were fortunate to be using a Dairy Card in 1980’s.**
- **Designs and analyses that allow for clinician effects may be needed.**
- **Weak effects may be obscured by contamination – cluster randomised trials, despite sample size and concealment issues – may be one way forward.**

