Experience with external review panels to validate a large clinical pharmacy intervention study

William McLean, Jeffrey Poston and Stephanie Tsao

ABSTRACT
Assessment of the true impact of pharmacists' interventions in pharmaceutical care is crucial to the justification of investment in resources for clinical pharmacy services. In a large study of clinical pharmacy interventions in hospitals, intervening pharmacists and attending physicians assessed the impact of the intervention on three aspects: therapeutic benefit, risks, and drug costs. The study showed that hospitals providing the highest level of pharmacotherapy monitoring made more interventions per patient and that, in these institutions, the impact was greater on therapeutic benefit and risk reduction. Both pharmacists and physicians caring for the patient had assessed the impact.

It was deemed important to validate these results. External panels of academic clinical pharmacists and clinical pharmacology physicians were chosen to review a sampling of cases to determine their level of agreement with their professional colleague (pharmacist or physician) at the original site. Differences were found in the assessments for both the pharmacist panel reviewing the site pharmacists and the physician panel reviewing the site physicians. In general, the review panels tended to be less positive about therapeutic benefit and risk reduction, but similar about the impact on drug costs.

Although validation exercises are desirable in such research, it remains to be established whether the expert panel approach is preferred to other methods such as the more arduous measurements of health outcomes, or even the subjective impression of the physician and pharmacist involved with the care of that patient.

Can J Hosp Pharm 1998;51: 200-208

INTRODUCTION
In these times, the idea of the profession seeking to evaluate the impact of its services seems not only valuable for its own development but essential for its self-preservation. Modern pharmacoeconomic assessment has come to include the cost-effective evaluation of not only drug therapies but also the pharmaceutical services that influence and deliver these drug therapies. Administrators of health care budgets, working under economic constraints, more than ever seek wisdom on where the limited budget will best be spent.

The profession is adopting the philosophy of pharmaceutical care (PC); however, in an era of critical evaluation of expenditures of health care dollars, pharmacy will have difficulty advancing PC if we are not able to justify it on a pharmacoeconomic basis. Clinical pharmacy services have been shown to have certain overall benefits such as a decreased incidence of drug reactions, decreased length of stay in hospitals and decreased drug costs,1-7 but much work remains to be done to explore which kinds of services provide the optimal cost-benefit.

The Canadian Society of Hospital Pharmacists' (CSHP) Clinical Pharmacy Advisory Committee (now the Pharmaceutical Care Advisory Committee) advocated in its "White Paper on the Establishment and Elaboration of Clinical Pharmacy Services"6 that of all the various clinical activities which pharmacists perform, patient pharmacotherapy monitoring (PPM) is the most valuable. A research arm of the Committee posed the question of whether increasing the intensity of PPM from a low level central drug order review (DOR), through a selective PPM,

William McLean, PharmD, FASHP, FCCP is Head, Pharmaceutical Outcomes Research Unit, Ottawa General Hospital/University of Ottawa.
Jeffrey Poston, PhD, is Director, Research and Practice Development, Canadian Pharmaceutical Association, Ottawa, Ontario.
Stephanie Tsao, BScPhm, is a pharmacist, London Health Sciences Centre Hospital, London, Ontario and former pharmacy resident at the Ottawa General Hospital.

Send correspondence to: William McLean, Head, Pharmaceutical Outcomes Research Unit, Ottawa General Hospital/University of Ottawa, 501 Smyth Road, Ottawa ON K1H 8L6. Phone 613 737 8161, fax 613 737 8891.
to a comprehensive PPM (C-PPM) model, made a difference in the perceived impact on patient care. To answer this question, the Clinical Pharmacy Services Study (CPSS) was developed. The results of this large study on the impact of interventions has recently been published.7

The CPSS collected data on pharmacist interventions from 17 randomly selected hospitals in the province of Ontario at 3 different levels of PPM and evaluated the impact of the interventions. In hospitals with the most intense level of monitoring — C-PPM — many more drug-related problems were identified and these interventions had more impact on therapeutic benefit and risk reduction than interventions at lower PPM levels. Pharmacists and physicians provided similar assessment of the impact; however, physicians tended to be less impressed by the risk reduction of the interventions than pharmacists. For example, hospitals with C-PPM made 9 times as many interventions as pharmacists with central Drug Order Review (DOR). Further, recommendations in C-PPM sites were assessed by pharmacists as having impact on therapeutic benefit 64% of the time, compared with 36% in DOR sites. Site or attending physicians found these recommendations had impact on therapeutic benefit 68% of the time in C-PPM sites versus 48% in DOR sites. In terms of risk reduction, pharmacists assessed their own recommendations as having impact 44% of the time in C-PPM sites, whereas physicians assessed this impact in 37% of the recommendations.

Members of the Pharmaceutical Care Advisory Committee of CSHP and an editorial in the Canadian Journal of Hospital Pharmacy8 suggested the need for validation of these results. A number of options were available. Review by pharmacy supervisors was considered but the modern, often autonomous, practice of clinical pharmacy does not readily lend itself to such peer review. Outcomes measures in patients were considered, but this would have considerably increased the resources needed. Using the patient as the judge would be interesting, but in most cases, patients still do not know what is intended by the various drug therapies and, even if they did, would not be in a position to judge their relative benefits, risks and costs. It was considered to ask another pharmacist to review the cases for a second opinion; however it was felt that this would add an institution bias. We considered asking the treating physician and included that as part of the study, but also considered having another physician at the site review the intervention; feasibility prevented it. Finally we considered the use of external panels to review a sample of interventions.

The use of expert panels or the “nominal group technique” is an increasingly popular method of seeking consensus, particularly on clinical interventions.9 Agreement is studied first among panellists and, second, between the panel and the issue (in our case the site recommendation).

The selection of panellists may lead to bias. Although it has been shown that physicians willing to participate on panels may be representative of their colleagues,10 the exact composition of the panel can affect the results.11–13 Expert panels have been satisfactorily used for a variety of consensus exercises, including the development of clinical guidelines.14 We developed a panel of physicians to review the physicians’ assessments and a panel of pharmacists to review the pharmacists’ assessments, to avoid interprofessional bias. Rather than comparing the assessments of the 2 professions, our goal was to validate the site pharmacists’ and physicians’ assessment independently by 2 separate external panels.

METHOD

Two panels of 3 members each were assembled, 1 panel of recognized clinical pharmacists to review the pharmacists’ assessment of their interventions; the other panel of recognized clinical pharmacologists, to review the site physicians’ assessment of the interventions. A panel of 3 was felt to be more potentially concordant than larger groups. In addition, it was desirable to find panellists from the general area of the study (Ontario) but who had not been part of the original study.

Criteria for the clinical pharmacy panel included nonparticipation in the original study (to avoid bias), a PharmD degree and at least 5 years’ clinical experience (to ensure substantial clinical experience and expertise), current clinical practice (to ensure practitioner relevance), a university affiliation (to provide academic acceptance), and geographic location in the metropolitan Toronto area (for ease of meeting).

Criteria for the clinical pharmacology panel included non-participation in the original study (to avoid bias), an MD with internal medicine specialization and training in clinical pharmacology (to provide experience with internal medicine cases), current clinical practice (to ensure practitioner relevance) and geographic location in the the Toronto–Hamilton area (for ease of meeting).
Table I—Distribution of cases by category of monitoring. DOR=Drug order review. S-PPM=Selective patient pharmatherapy monitoring. C-PPM=Concurrent pharmacotherapy monitoring.

<table>
<thead>
<tr>
<th>Type of monitoring</th>
<th>Cases by category (and %) by panel</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pharmacists n=260</td>
</tr>
<tr>
<td>DOR</td>
<td>24 (9)</td>
</tr>
<tr>
<td>S-PPM</td>
<td>46 (18)</td>
</tr>
<tr>
<td>C-PPM</td>
<td>190 (73)</td>
</tr>
</tbody>
</table>

Members of each panel were supplied with 100 case summaries (the anticipated workload limit of panellists) on Data Collection forms (see Appendix I), selected at random proportionately from each level of PPM. Twenty were common among panellists (to check intrapanelist variation), for a total of 260 case summaries per panel. Cases were selected randomly from the pool of case summaries, including those not accepted by the site physicians. (This latter information was not provided to the panel.) Selection of cases for the physician panel was done separately and only from those cases where the site physicians had evaluated the impact.

Table I shows the distribution of the cases by category and the distribution of the cases according to the type of monitoring involved. Table II shows the same distribution of the cases common to all panellists.

Table II—Distribution of common cases by category of monitoring. DOR=Drug order review. S-PPM=Selective patient pharmatherapy monitoring. C-PPM=Concurrent pharmacotherapy monitoring.

<table>
<thead>
<tr>
<th>Type of monitoring</th>
<th>Cases by category (and %) by panel</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pharmacists n=20</td>
</tr>
<tr>
<td>DOR</td>
<td>3 (15)</td>
</tr>
<tr>
<td>S-PPM</td>
<td>6 (30)</td>
</tr>
<tr>
<td>C-PPM</td>
<td>11 (55)</td>
</tr>
</tbody>
</table>

For purposes of orienting the panels to the review process and developing common definitions for the ratings of the interventions, study investigators met with each panel separately providing the study objectives, 8 case samples for a common understanding of the rating scale and definitions for the interpretation of the case summaries. Over a 2-week period, panellists completed their assessment forms independently and their assessment was compared to the original site pharmacist’s or physician’s assessment.

Statistical analysis

Percent agreement was calculated for the panellists’ co-agreement on the impact of the 3 assessed variables: therapeutic benefit, risk and drug costs. Percent agreement was calculated to indicate the percent of cases in which panellists were in agreement on the positivity or negativity of the therapeutic benefit, risk reduction and drug costs.

Concordance among panellists was evaluated using the Kappa statistic, while the degree of agreement on the impact rating of the 3 variables was evaluated using the weighted Kappa statistics, approximating the interclass correlation coefficient. The Wilcoxon Matched Pair Signed Rank statistic was used to measure the degree of agreement between the site and panel assessors. For purposes of interpreting Kappa values in this study, we adopted the scale of “strength of agreement” proposed by Brennan and Silman, in which a Kappa value of less than 0.2 indicates poor strength of agreement, 0.21-0.40 is fair, 0.41-0.60 is moderate, 0.61-0.80 is good and 0.81-1.00 is very good.

If a case was deemed not evaluable as judged by any one panellist, it was rejected from analysis.

RESULTS

The results are shown in Figures 1-6.

Agreement amongst panellists

Table III shows the percent agreement amongst panellists on the 2 panels. Eighteen (90%) of the 20 cases were deemed evaluable by the panellists. There was excellent inter-rater agreement on the 2 panels, with weighted Kappa values of 0.92, 0.86 and 0.90 for the pharmacist panel and 0.91, 0.88 and 0.94 for the pharmacologist panel on the 3 variables: therapeutic effect, risk reduction, and drug costs respectively (p<0.00001).
Figure 1: Pharmacist Panel vs. Site Pharmacist Assessment of Therapeutic Impact

Figure 2: Pharmacist Panel vs. Site Pharmacist Assessment of Risk Impact

Figure 3: Pharmacist Panel vs. Site Pharmacist Assessment of Cost Impact

Figure 4: Physician Panel vs. Site Physician Assessment of Therapeutic Effect

Figure 5: Physician Panel vs. Site Physician Assessment of Risk Impact

Figure 6: Physician Panel vs. Site Physician Assessment of Cost Impact
Table III—Agreement among panellists, by impact category. Percentage agreement refers to the % of cases in which panellists all agreed on either the absence or presence of the therapeutic impact in question.

<table>
<thead>
<tr>
<th>Factor or impact category</th>
<th>Pharmacists</th>
<th>Physicians</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accept recommendation</td>
<td>94</td>
<td>83</td>
</tr>
<tr>
<td>Therapeutic benefit</td>
<td>79</td>
<td>67</td>
</tr>
<tr>
<td>Risk reduction</td>
<td>72</td>
<td>44</td>
</tr>
<tr>
<td>Drug costs</td>
<td>72</td>
<td>72</td>
</tr>
</tbody>
</table>

Agreement between panels and sites

Pharmacists’ panel

Two hundred and fifty (96%) of the 260 cases were judged to be evaluable. The panel agreed with 94% of the recommendations. Yet, among panellists there were differences; these are summarized in Table IV. In general, while there was fair agreement, examination of Figures 1–3 does suggest that site assessors were slightly more positive about the impact of therapeutic benefit and risk reduction, but less positive about drug costs than were the panellists.

Table IV—Agreement between panel and sites. % agreement refers to % of cases in which panel members agreed with the site assessor for each particular variable. K=Kappa value, the degree of strength of that agreement. * p=statistical confidence for that K value. p≤0.001 except where noted.

<table>
<thead>
<tr>
<th>Type of impact</th>
<th>% agreement by panel (and Kappa value)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pharmacists</td>
</tr>
<tr>
<td>Therapeutic benefit</td>
<td>60 (0.27)</td>
</tr>
<tr>
<td>Risk reduction</td>
<td>70 (0.37)</td>
</tr>
<tr>
<td>Drug costs</td>
<td>70 (0.32)</td>
</tr>
</tbody>
</table>

* K=P(A) - P(E) where P(A) is the observed proportion of agreement, and P(E) is the expected proportion of agreement due to chance.

† p≤0.04.

Physicians’ panel

In this case, agreement was less strong for risk reduction (Table IV). The panellists tended to be less positive than the site physicians, except for cost impact, where they were slightly more positive (Figures 4–6).

Overall, there was good inter-rater agreement within each of the panels as evidenced by the high weighted Kappa values in all 3 assessment categories (therapeutic effect, risk reduction, and drug costs).

Regarding agreements between the panels and the site professionals, some differences in the perceived impact were found. Both panels rated the interventions as having less therapeutic impact than the site professionals. For example, the clinical pharmacist panel rated 45% of the interventions as having less impact on therapeutic effect than the site pharmacists; similarly, the physician panel rated 50% of the interventions less favourably. These differences (88%) were mostly by 1 rating degree on the 7-point scale, with the majority of interventions being rated as no impact or mild benefit.

In terms of risk assessment, the clinical pharmacist panel rated the impact of risk reduction less than the site pharmacists; the vast majority (92%) differed by only 1 degree. There was no difference on the physician side. Regarding impact on drug costs, the panels and site professionals judged the impact similarly, noting a mostly neutral or favourable impact. In general, the panel agreed with the site professional but to a lesser degree.

DISCUSSION

The problem of bias has limited the interpretation of many studies and external panels have often been proposed. Hatoum has found similar differences between site and external assessors; similar to our study, cases were rated somewhat less favourably by external assessors. In their study, an external panel of physicians rated pharmacists’ assessments; therefore some difference might be explainable by interprofessional differences. Pharmacists see their impact as particularly important in risk reduction. Physicians are perhaps less impressed by the amount of risk reduction claimed by pharmacists and are perhaps more accustomed to health risks than pharmacists. A recent survey has shown that physicians perceive the importance of pharmacists’ clinical activities differently than pharmacists depending on the particular activity.

The exercise of asking external panels to review the impact of pharmacist interventions was feasible and appeared to be an objective method of avoiding intra-site biases. Part of the success of panel agreement was the use of a test period to orient panellists to the process. Our live meetings with the panellists appeared to have
permited consensus development on the scale used. This was a crucial step before comparing results between the panel and the original site professionals.

There is suspicion that self-assessors, in this case, the site pharmacists, may have a subjective bias and may wish to present themselves as being valuable. Others might argue to the contrary because this was a public process, that is, that the rating had to be done in a research study where there was a site coordinator for each hospital reviewing the data collection. As well, it was known by the pharmacists that their intervention would also be sent to the treating physician for similar impact rating. These two influences may have offset some bias.

The blinded panelists should be free of bias. They have the necessary academic credentials to bring objective evidenced-based therapeutics to their judgment. Some would argue that they may be viewed as being too critical, especially if the evidence for the intervention is weak. However, we found that the greatest drawback for the panellists was the lack of information on the case. The case summary provided to the panellists, although the same as that provided to the site physicians, gives only basic data on the intervention with some patient demographics and diagnosis and problem list. Relatively little information on the seriousness of the need for the intervention or the outcome is provided; even concurrent pharmacotherapy was not available to them. The authors believe that this is the largest factor in causing panellists to modify conservatively their assessments of the impacts. From this point of view, the site physician who knew the case intimately may be in a better position to provide an assessment of the impact. In addition, the existence of agreement does not necessarily mean that everyone is right or correct.9

Based on this potential limitation of the study, it would have been interesting to provide more data to the external assessors. This would have required a much greater time commitment by the panellists — perhaps as much time as the original intervention took.

Better definition of the rating scales would also be ideal. Although the panellists arrived at consensus during the orientation period as to what constituted a “marked” versus a “moderate” improvement, the original study did not permit as much consensus building for site physicians and site pharmacists. Many of the pharmacists’ interventions fell into a category of “no impact”, for example, formulary changes, PO versus IV (no change in therapeutic effect, although some in risk and costs); it would be interesting to focus on those interventions that were felt to have more impact.

The differences in the ratings between the panels and the site practitioners is disturbing. It is important to develop a review process of the value of pharmacists’ impact that becomes accepted as credible. Until this methodology becomes available, we are prevented from making good pharmacoeconomic extrapolations. For instance, we could suggest a theoretical model where “marked” impacts reduce hospital stay by 1–2 days, while “moderate” impacts may only reduce morbidity without influencing length of stay. Even if the savings per case is only $4.75 in reduced drug costs per 24 hours (as shown in the CPSS1), a conservative estimate of 5 days (half of their length of stay) of reduced costs yields a savings of $640,000 per year for a hospital of 450 beds. This does not include savings from reduced hospital stays or decreased adverse reactions. If you extrapolate this drug cost savings alone to the approximately 145,210 acute care treatment beds in Canada (as provided by the Canadian Health Care Association, mid 1996), such interventions could result in an annual savings of $206.52 million to the Canadian health care system.

CONCLUSION

In an attempt to validate the assessments by the site professionals of the impact of the pharmacists’ interventions in the CPSS, 2 external review panels were established for each profession. Similar to other studies,4,20 it was found that site professionals rated these interventions more positively than the panelists. Of the 3 categories assessed, both therapeutic effect and risk impacts were rated higher by the site staff; impact on drug costs were rated similarly.

While the results confirm previous findings, the limitations of the methodology make an assessment of the actual impact of the interventions on patient outcomes difficult. Indeed this research area remains challenging.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the contributions of Elliot A. Loh, MSc, Canadian Pharmaceutical Association, Ottawa and Barbara G. Ogle, MScPhm,
206

The Canadian Journal of Hospital Pharmacy

consultant and Director, Professional Services, Pharmasave National, Langley, B.C.

We also thank the members of the Pharmacist Panel: Susan Bowles, PharmD, FCCP, Sunnybrook Health Sciences Centre, Toronto; Al Rasymus, PharmD, St. Michael’s Hospital, Toronto; and, Glen Pearson, PharmD, The Toronto Hospital. The Physician Panel consisted of Anne Holbrook, MD, PharmD, MSc, McMaster University, Hamilton; Paul Oh, MD, Sunnybrook Health Sciences Centre, Toronto; and B. Liu, MD, Sunnybrook Health Sciences Centre, Toronto.

Primary funding was provided by the Ontario Ministry of Health’s Hospital Incentive Fund. Additional research funding was obtained from Glaxo Wellcome Canada Inc. and Apotex Inc.

REFERENCES

Part Ill: PERCEIVED IMPACT

Do you think this recommendation(s) is likely to have an impact in terms of:

- THERAPEUTIC EFFECT for the patient?
  - [ ] Yes
  - [ ] No
  - If YES, rate your perception of the marked degree of impact: benefit, detriment

- RISK for the patient?
  - [ ] Yes
  - [ ] No
  - If YES, rate your perception of the marked degree of impact: decrease, increase

- COSTS of drug therapy?
  - [ ] Yes
  - [ ] No
  - If YES, rate your perception of the marked degree of impact: decrease, increase

SUBSEQUENT ACTION (within 48 working hours):

- Recommendation was: [ ] modified and accepted [ ] withdrawn [ ] action unknown

Other factors that could have influenced patient response:

- [ ] Yes
  - [ ] No
  - [ ] Unable to determine

If YES, specify:

- Monitoring continued until:  

- Date monitoring stopped:

PATIENT RESPONSE (to be completed once per patient):

- [ ] Yes
  - [ ] No
  - [ ] Not required
  - [ ] Not feasible
  - [ ] Unable to determine

If YES, describe patient response:

- [ ] Monitoring continued until:  

- Date monitoring stopped:

- Appropriate time for therapy: 

- Discharge/transfer:
  - [ ] Maximum of 7 days