

INTER-RATER AGREEMENT BETWEEN PALLIATIVE CARE CLINICIANS WHEN EVALUATING ABDOMINAL RADIOGRAPHS TO QUANTIFY THE DEGREE OF FAECAL SHADOWING VISIBLE

Katherine Clark, Lawrence Lam, Nick Talley, Jane Philips, Naomi Byfieldt, David Currow

Introduction

Even though the majority of cancer patients cared for with palliative intent are likely at some point to experience constipation symptoms, evidence-based approaches to diagnosing and assessing the severity of the problem are lacking. The highly subjective nature of the problem and current lack of a gold standard diagnostic approach explains why so many clinicians continue to rely on the appearance of plain radiographs, extrapolating from this both the presence and severity of constipation.

Previously, we have illustrated that there is very poor agreement between four clinicians with varying levels of clinical experience when reporting plain radiographs.

Study Aims

This current study was undertaken with the intent of exploring whether the level of a clinician's training and experience affected the levels of agreement between palliative care practitioners when they were asked to report their assessments of the amount of faecal shadowing visible on plain abdominal radiographs.

Patients and Methods

-Settings and participants

Six different palliative care services were approached to participate and with the approval of their human ethics committees, convenience samples of palliative care doctors drawn from three clinical sub-groupings were recruited; namely palliative medicine specialists, registrars and resident medical officers (RMO).

-Data collection

Each of the consenting participants individually reviewed 10 abdominal radiographs with the aim of summarising their opinions of the amount of faecal shadowing they observed in three sections of the colon: the right colon, the left colon and the sigmoid. This was according to previously published recommendations, where a numerical score of 0-15 is allocated, with higher scores reflective of more significant shadowing.

The 10 radiographs utilised for this work had all been previously taken from consenting patients with advanced cancers who had self-identified themselves as constipated and requiring laxatives.

The participants reporting the radiographs were not given this information or any clinical history so as not to prejudice their assessments, as this was an exercise reporting alone.

-Analysis

De-identified demographic details and clinician's opinions regarding the use of radiographs were summarised using summary statistics. Fleiss's Kappa (FK) was used to evaluate concordance between multiple raters within each group with this technique evaluating the reliability of agreement amongst > 1 rater, with 0-0.2 = slight association and 0.8-1 = almost perfect agreement. The regression coefficient was calculated of pair-wise comparisons among groups. This was performed with the registrar as the reference group and included Specialists versus Registrars as well as RMOs vs Registrars.

Results

Demographics

There were 46 doctors who agreed to report the radiographs. Of these, 62% were women (n=28) with nearly half less than 35 years of age (n=19). There were similar numbers of participants in each of the three groups. The doctors were asked to comment when was the last time they had ordered a radiograph to assess constipation with 40% (n=18) of the doctors leaving this section blank.

Address for correspondence:
Katherine Clark, Clinical Director of Palliative Care
Northern Sydney Local Health District, St Leonards, NSW, Australia
Katherine.clark@health.nsw.gov.au

Agreement regarding the degree of faecal shadowing in subgroups and between subgroups

Little agreement between doctors of similar experience and training faecal shadowing scores was seen as summarised below in table 1..

Table 1: Concordance between raters by clinical group

Clinical grouping	Right colon	Left colon	Sigmoid colon
RMO	0.05	0.05	0.05
Registrar	0.10	0.13	0.06
Specialist	0.13	0.11	0.11

The regression coefficient was calculated to assess whether the levels of agreement within a clinical subgroup was different to another subgrouping based on the expectation that more experienced doctors were more likely to agree with other. However, the data did not support this as detailed in table 2.

Table 2: Results on the comparison of X-ray viewing among different groups of participants with the registrars as the referent group

	Regression coefficient	S.E.	Z	p-value
Right Colon				
RMO	0.23	0.195	1.20	0.232
Specialist	-0.23	0.192	-1.21	0.228
Left Colon				
RMO	-0.08	0.22	-0.35	0.725
Specialist	-0.14	0.22	-0.64	0.521
Sigmoid				
RMO	-0.10	0.25	-0.40	0.690
Specialist	-0.25	0.24	-1.02	0.309
Likelihood				
RMO	-0.02	0.297	-0.07	0.946
Specialist	-0.27	0.288	-0.92	0.356
Symptoms				
RMO	0.28	0.311	0.88	0.377
Specialist	0.03	0.294	0.10	0.924

Discussion

The aim of this work was to explore whether there were higher levels of agreement amongst more experienced clinicians compared to their less experienced colleagues when reporting their opinions of faecal shadowing seen on plain radiographs. Although it was expected that the more senior clinicians opinions would be more closely aligned, the results suggest that only very faint levels of agreement within all three groups. Further when the levels of agreement were compared between groups there were no real differences seen.

It was not the main aim of this study to consider the diagnostic accuracy of the plain radiographs but instead to comment on the performance of the radiograph in interpretation alone.

The use of plain radiographs becomes more problematic when the lack of data to report the sensitivity and specificity of the investigation for this purpose is considered. In the absence a diagnostic gold standard, obtaining this information is not really possible.

There are strengths and weaknesses to this work. The strengths include the fact that this is the only study that has compared the opinions of a reasonable number of palliative care doctors drawn from different stages of experience.

The weaknesses include the fact that the clinicians do not have the patient's clinical history. While this was true, this was done with the intent of not influencing the reports and rather trying to make this an exercise that was assessing the radiographs appearance first and then interpreting it second.