Health Service Research

Agreement between patients and general practitioners on quality deviations during the cancer diagnostic pathway and associations with time to diagnosis

Henry Jensen\textsuperscript{a,b,*}, Cecilie Sperling\textsuperscript{c}, Mette Sandager\textsuperscript{c} and Peter Vedsted\textsuperscript{a}

\textsuperscript{a}Research Centre for Cancer Diagnosis in Primary Care, Research Unit for General Practice, Department of Public Health and \textsuperscript{b}Section for General Medical Practice, Department of Public Health, Aarhus University, Bartholins Allé 2, DK-8000 Aarhus C and \textsuperscript{c}Documentation & Quality, Danish Cancer Society, Strandboulevarden 49, DK-2100 Copenhagen OE, Denmark.

*Correspondence to Henry Jensen, Research Centre for Cancer Diagnosis in Primary Care, Research Unit for General Practice, Department of Public Health, Aarhus University, Bartholins Allé 2, DK-8000 Aarhus C, Denmark; E-mail: henry.jensen@feap.dk

Abstract

Background. High quality and minimal delay are crucial and anticipated elements in the diagnostic cancer pathway as delay in the diagnosis may worsen the prognosis and cause lower patient satisfaction.

Objective. The aim of this study was to describe agreement in reported quality deviations (QDs) between general practitioners (GPs) and cancer patients during the diagnostic pathway in primary care and to estimate the association between length of diagnostic interval and level of agreement on reported QDs.

Methods. The study was carried out as a Danish cross-sectional study of incident cancer patients identified in the Danish National Patient Registry. Data were collected by independent questionnaires from patients (response rate: 53.0%) and their GPs (response rate: 73.8%), and 2177 pairs of questionnaires were subsequently combined. Agreement between GP- and patient-reported QDs was estimated using Cohen’s Kappa, whereas the association between level of agreement and time to diagnosis was estimated using quantile regression.

Results. Patients reported QDs in 29.0% and GPs in 28.5% of the cases, but agreed only slightly on QD presence (Kappas between \textemdash0.08 and 0.26). Agreement on ‘QD presence’ was associated with a 54-day (95%CI: 44\textendash64) longer time to diagnosis than agreement on ‘no QD presence’. The association with a longer diagnostic interval was stronger when only GP reported a QD the association than when only patient reported a QD.

Conclusion. Included GPs and patients agreed only slightly on QD presence although they reported the same amount of QDs; this suggests that GPs and patients see QDs as two different concepts. QD presence had a stronger impact on time to diagnosis when reported by the GP (alone or in agreement with the patient) than when reported by the patient alone. The GP may thus be the most important source of information on QD and diagnostic interval, while the patient information tends to underpin this assessment.

Key words: Agreement, Denmark, delayed diagnosis, neoplasms, general practice, health care quality assessment.
Introduction

Identification of patients that may have serious disease is a particularly important task in general practice as they often consult the general practitioner (GP) for medical advice prior to the actual diagnosis. For example, more than 75% of all cancer patients will contact their GP because of symptoms (1). Only half will present well-known alarm symptoms of cancer with relatively low positive predictive values (2), while the other half will present with vague or uncharacteristic low risk but not no-risk symptoms and signs (3). Nevertheless, high quality and minimal delay in the diagnostic pathway are expected, as delay in the time to cancer diagnosis may affect treatment options or require more intensive treatment with more adverse effects, which may negatively impact the patient’s quality of life and affect the prognosis unfavourably.

Delays in cancer diagnosis can have many different causes (4), but have been shown to be the most common consequence of an event, that deviates from good quality in primary care (5). A quality deviation (QD) can be defined as an event ‘that should not have happened and that you don’t want to happen again’ (6). This definition has formerly been used in general practice (5,6). We have recently shown that GPs in Denmark reported at least one QD for a third of all cancer patients during the primary diagnostic phase and that these patients had a longer diagnostic interval than patients without a reported QD (7). This substantiates the importance of reducing QDs. Patient evaluations are an important source of knowledge about the diagnostic pathway. Therefore, the patients’ experience with the healthcare sector may also be used as a data source to provide detailed information on QDs during cancer diagnosis (8).

However, the overall findings of QDs in cancer diagnosis are so far diverse, and patients and doctors do not always seem to agree (9–15). Most former studies have been conducted among hospital doctors, why it is unknown how much GPs and their patients agree. Another important issue is whether agreement between GP and patient is particularly associated with a delayed diagnostic pathway, doctors, why it is unknown how much GPs and their patients agree (far diverse, and patients and doctors do not always seem to agree (16,17)). If such association does exist, it would be crucial to identify of patients, data collection and response analyses have been reported in detail elsewhere (1,7,17). In brief, we identified incident cancer patients of 18 years or older who were listed with a GP and given a C00.0–C99.9 [except nonmelanoma skin cancer (C44)] diagnosis coded according to International Classification of Diseases, version 10 (ICD-10) during the period from 1 May 2010 to 31 August 2010. All patients were identified in the Danish National Patient Registry. A patient was defined as incident when registered with cancer as the primary diagnosis during the inclusion period, and with no prior history of cancer recorded in the Danish Cancer Registry. We excluded 177 (2.2%) patients as their diagnosis could not be verified in the Danish Cancer Registry, leaving 7562 eligible patients (1).

Methods

This study was a population-based nationwide cross-sectional study among consecutive incident cancer patients in Denmark.

Setting

The data were collected in the summer and autumn of 2010 in Denmark. Denmark’s publicly funded healthcare system ensures free access to diagnostics and treatment for all citizens. More than 98% of Danish citizens are registered with a specific general practice, which they have to consult for medical advice. The GP initiates diagnostic investigations and acts as a gatekeeper to the specialized healthcare system. Since 2008, Danish GPs have been provided with the opportunity to refer patients with suspected cancer to standardized fast-track cancer pathways.

Identification of patients

Identification of patients, data collection and response analyses have been reported in detail elsewhere (1,7,17). In brief, we identified incident cancer patients of 18 years or older who were listed with a GP and given a C00.0–C99.9 [except nonmelanoma skin cancer (C44)] diagnosis coded according to International Classification of Diseases, version 10 (ICD-10) during the period from 1 May 2010 to 31 August 2010. All patients were identified in the Danish National Patient Registry. A patient was defined as incident when registered with cancer as the primary diagnosis during the inclusion period, and with no prior history of cancer recorded in the Danish Cancer Registry. We excluded 177 (2.2%) patients as their diagnosis could not be verified in the Danish Cancer Registry, leaving 7562 eligible patients (1).

Data collection

Data were obtained separately at patient and GP level by asking both the patient and the patient’s GP to complete a questionnaire. These two data sets were subsequently linked through the civil registration (CPR) number, a unique personal identifier assigned to all Danish citizens at birth or immigration.

GP questionnaire

A questionnaire for each individual patient was sent to the patient’s GP, who was asked to complete the questionnaire on the basis of the electronic medical records. The GP received no remuneration for participation. Non-responders received a reminder, including a new questionnaire, after approximately 4 weeks.

The questionnaire was based on 16 previously used and tested measures of the diagnostic pathway, including questions on the extent of GP involvement (e.g. symptom-based diagnosis, breast cancer screening or emergency admission) (1). Part of the questionnaire addressed whether the diagnostic pathway had given rise to QDs; if positive, the respondent was asked to state which types of QDs. The QD definition was based on the initial work of Dovey et al., who defined QDs (6) as ‘events that should not have happened and that you do not want to happen again’. The GP could choose from 10 different QD categories (including one free-text option) based on level 1.1 (office administration), 1.2 (investigations), 1.4 (communication) and 2 (knowledge and skills errors) from Dovey et al.’s taxonomy (6). Hence, the GP QDs was based on the GP’s subjective interpretation. The free-text coding created two new major categories (‘waiting time in secondary care’ and ‘patient wanted to postpone diagnostics’), which were most often stated by the GPs (7).

The questionnaire also included previously tested items about specific dates during the diagnostic interval. In accordance with the Aarhus Statement, the diagnostic interval was defined as the time from the patient’s first presentation of symptoms of cancer to a doctor until the time of diagnosis (4).

Patient questionnaire

Each patient received a questionnaire including a prepaid return envelope. Questionnaires were distributed in two batches approximately 3 months after the patient had been registered in the National Patient Registry. Non-responders received a reminder after 4 weeks.

The patient questionnaire was developed by the Danish Cancer Society to explore cancer patients’ experiences with health care services during diagnosis and treatment. A review of the literature and 6 focus group interviews with 31 cancer patients were conducted to identify substantial needs and problems for the patients. A questionnaire containing 104 items was developed to determine the patient’s...
entire pathway. The content validity of the questionnaire was tested by assessing the respondents’ understanding of the included items through interviews with 13 cancer patients of different age and gender who were diagnosed with different diseases. The content validity of the questionnaire was also assessed by health care professionals from both clinical environments and research settings.

Patient-reported QDs were confined to include only the patients’ experiences in the pre-diagnostic phase (eight tick-off questions and six free-text options). We defined a QD to be present when the patient reported the worst category in any question regarding: information given, communication with health care professionals, missed test (results), adverse events in primary care, waiting time and collaboration between GP and secondary care (i.e. medical specialists).

Analyses

The overall agreements are presented descriptively by numbers and percentages. We calculated Cohen’s Kappa statistics to determine the agreement between patients and GPs to account for agreement by chance. The following categories have been proposed for evaluating Kappa values: less than 0.00 = poor, 0.00–0.20 = slight, 0.21–0.40 = fair, 0.41–0.60 = moderate, 0.61–0.80 = substantial and 0.81–1.00 = almost perfect. As we allowed for both patient and GP to respond ‘unknown’, we calculated a weighted Kappa; agreement on ‘no QD present’ or agreement on ‘QD present’ were both weighted 1, while cases in which either the GP or the patient reported ‘unknown’ were weighted 0.5. The overall agreement in percentages and appertaining Kappa values were stratified by gender, age groups and the five most common cancers (colorectal, breast, prostate, lung and malignant melanoma cancer), two cancers often described as clinically difficult to diagnose (ovarian and bladder cancer), and uterus cancer. Bootstrapping procedure was applied to obtain bias-corrected estimates of 95% confidence intervals (CIs) for the Kappa values.

To investigate the impact of the ‘unknown’ QD category on the agreement level, the analyses were also performed on data for which all ‘unknown’ QD categories were coded as ‘no QD’. This, under the assumption, that if no QD was noticed, the QD was not severe enough to be significant.

Quantile regression was used to estimate the association between QD agreement and diagnostic interval. We defined GP-patient agreement in four distinct categories: (i) agreement that no QD was reported, (ii) GP reported a QD (but not the patient), (iii) the patient reported a QD (but not the GP) and (iv) agreement that at least one QDs was present (both GP and patient reported a QD). For increasing centiles, we estimated additional days in the diagnostic interval compared with the diagnostic interval for which both GP and patient reported no QDs. Differences were adjusted for categorical variables for patient gender and age, geographical region and cancer site. The 95% confidence intervals were computed by the standard error (SE) estimated by 500 bootstrap replications. The statistical level of significance was set at 0.05 or less.

All analyses were done using Stata v. 13.1 (StataCorp, College Station, TX).

Results

GP questionnaires were returned for 5581 cases (response rate: 73.8%). Patients from responding GPs were more likely to be female, to be diagnosed with breast cancer, or to have high educational level (1).

Patients returned 4010 questionnaires (response rate: 53.0%; response rate of those alive and with publically available address: 64%). Respondents were more likely to be women, to be aged 45–74 years, to be diagnosed with breast cancer or malignant melanoma, to have higher 1-year survival rates, to have more localized tumours, to have higher educational background and to have higher disposable income (1).

Of the 2984 combined cases, we excluded 694 (23.3%) patients for whom the GP was not involved in the diagnosis (e.g. patients diagnosed by screening or emergency) and 113 (3.8%) patients due to missing information on QD presence (Fig. 1). In total, 2177 cases were included in the analyses (Fig. 1). The characteristics of the included cancer patients are shown in Table 1.

GPs reported that QDs were present in 620 (28.5%) of the cases and that no QD was present in 1495 (68.7%) of the cases. The patients reported that QDs were present in 632 (29.0%) and that no QD was present in 771 (35.4%) of the cases. In the remaining cases, QD presence was reported as ‘unknown’ (Table 1). When ‘unknown QD’ was recoded into ‘no QD’, the number of cases with no QD reported rose to 1557 (71.5%) for GP-reported cases and to 1545 (71.0%) for patient-reported cases.

The overall agreement on reported QDs was 36.9% (95%CI: 34.6–39.0%), with a corresponding Kappa value of 0.09 (95%CI: 0.05–0.13). Minor differences in agreement were observed across the patient characteristics, with the highest Kappa value of 0.18 (0.03–0.35) among bladder cancer patients and 0.18 (0.05–0.28) among patients aged 18–44 years (Table 2). The agreement increased when ‘unknown QD’ was recoded as ‘no QD’ (Table 2). Here the overall agreement was 64.9% (62.9%–66.9%), with a Kappa value of 0.14 (0.10–0.19) with the highest agreement observed among cases with rectal cancer displaying a Kappa value of 0.29 (0.11–0.47) (Table 2).

The agreement on reported QDs were of the same calibre when stratified by subcategories; e.g. agreement on whether the communication between patient and GP had a QD present was 70.2% (95%CI: 68.2–72.3) with a corresponding Kappa Value of 0.03 (0.00–0.07). The agreement on communication QDs increased to 90.2% (95%CI: 88.6–91.6) with a corresponding Kappa value of 0.06 (0.00–0.13), when ‘unknown QD’ was recoded as ‘no QD’.

The amount of reported QDs, agreements and Kappa values were the same when stratified by gender, age and diagnoses (data not shown). The diagnostic interval was associated with the level of agreement. Compared to the median diagnostic interval of 21 days for cases in which both patient and GP agreed on ‘no QD’, the adjusted median diagnostic interval was 32 (95%CI: 23–41) days longer when only the GP reported a QD, 9 (95%CI: 4–14) days longer when only the patient reported a QD, and 54 (95%CI: 44–64) days longer when patient and GP agreed on QD presence. The association between agreement and diagnostic interval was strongest at all percentiles when GP and patient agreed on QD presence and when only the GP noted a QD (Fig. 2).

Discussion

GPs and patients both reported that 29% of the diagnostic pathways included one or more QDs. However, they only agreed slightly on whether it was the same pathway, with more than one in three of the included patients reported that they did not know if any QD occurred. This indicates that GPs and patients may regard QDs differently. In addition, data on QDs were more complete for GPs. When GP and patient agreed on QD presence in the diagnostic pathway, the median time to diagnosis was 54 days longer than when they agreed on no presence of QDs. When only the GP reported a QD, the median
interval was 32 days longer, while it was 9 days longer when only the patient reported a QD. The combined GP and patient information on QDs was thus a strong indicator of a delayed cancer diagnosis, although the GP's reported assessment was the strongest indicator.

**Strengths and weaknesses of study**
The greatest strength of the study was the possibility for linking a high number of responses from patients to responses from their GP. The study population was well-defined and complete with minimal

---

**Table 1. Characteristics of Danish cancer patients and reported QDs by gender, age and diagnoses in 2010 (n = 2177)**

<table>
<thead>
<tr>
<th>Cancer diagnosis</th>
<th>Female</th>
<th>Male</th>
<th>Total</th>
<th>P value*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n (%)</td>
<td>n (%)</td>
<td>n (%)</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1037 (100)</td>
<td>1140 (100)</td>
<td>2177 (100)</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>–</td>
<td>1140 (100)</td>
<td>1140 (52.4)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>1037 (100)</td>
<td>–</td>
<td>1037 (47.6)</td>
<td></td>
</tr>
<tr>
<td>Age group (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18–44</td>
<td>107 (10.3)</td>
<td>43 (3.8)</td>
<td>150 (6.9)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>45–54</td>
<td>186 (17.9)</td>
<td>88 (7.7)</td>
<td>274 (12.6)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>55–64</td>
<td>252 (24.3)</td>
<td>287 (25.2)</td>
<td>539 (24.8)</td>
<td>0.637</td>
</tr>
<tr>
<td>65–74</td>
<td>274 (26.4)</td>
<td>452 (39.6)</td>
<td>726 (33.3)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>75–</td>
<td>218 (21.0)</td>
<td>270 (23.7)</td>
<td>488 (22.4)</td>
<td>0.137</td>
</tr>
<tr>
<td>Cancer diagnosis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colon</td>
<td>118 (11.4)</td>
<td>113 (9.9)</td>
<td>231 (10.6)</td>
<td>0.267</td>
</tr>
<tr>
<td>Rectal</td>
<td>49 (4.7)</td>
<td>81 (7.1)</td>
<td>130 (6.0)</td>
<td>0.019</td>
</tr>
<tr>
<td>Lung</td>
<td>75 (7.2)</td>
<td>102 (8.9)</td>
<td>177 (8.1)</td>
<td>0.144</td>
</tr>
<tr>
<td>Malignant melanoma</td>
<td>78 (7.5)</td>
<td>58 (5.1)</td>
<td>136 (6.2)</td>
<td>0.019</td>
</tr>
<tr>
<td>Breast</td>
<td>375 (36.2)</td>
<td>3 (0.3)</td>
<td>378 (17.4)</td>
<td>n/a</td>
</tr>
<tr>
<td>Uterus</td>
<td>73 (7.0)</td>
<td>–</td>
<td>73 (3.4)</td>
<td>n/a</td>
</tr>
<tr>
<td>Ovarian</td>
<td>30 (2.9)</td>
<td>–</td>
<td>30 (1.4)</td>
<td>n/a</td>
</tr>
<tr>
<td>Prostate</td>
<td>–</td>
<td>399 (35.0)</td>
<td>399 (18.3)</td>
<td>n/a</td>
</tr>
<tr>
<td>Bladder</td>
<td>11 (1.1)</td>
<td>59 (5.2)</td>
<td>70 (3.2)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Others</td>
<td>228 (22.0)</td>
<td>325 (28.5)</td>
<td>553 (25.4)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>GP-reported QD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>308 (29.7)</td>
<td>312 (27.4)</td>
<td>620 (28.5)</td>
<td>0.228</td>
</tr>
<tr>
<td>No</td>
<td>699 (67.4)</td>
<td>796 (69.8)</td>
<td>1495 (68.7)</td>
<td>0.904</td>
</tr>
<tr>
<td>Unknown</td>
<td>30 (2.9)</td>
<td>32 (2.8)</td>
<td>62 (2.8)</td>
<td>0.224</td>
</tr>
<tr>
<td>Patient-reported QD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>315 (30.4)</td>
<td>317 (27.8)</td>
<td>632 (29.0)</td>
<td>0.187</td>
</tr>
<tr>
<td>No</td>
<td>347 (33.5)</td>
<td>424 (37.2)</td>
<td>771 (35.4)</td>
<td>0.572</td>
</tr>
<tr>
<td>Unknown</td>
<td>375 (36.2)</td>
<td>399 (35.0)</td>
<td>774 (35.6)</td>
<td>0.069</td>
</tr>
</tbody>
</table>

*Mann–Whitney test for difference between gender.
selection bias as the population was identified using a previously developed accurate algorithm (18). Furthermore, the case-mix of the cohort has been shown to resemble the case-mix of the Danish Cancer Register (18); this indicates that the inference to the source population seems valid.

The high response rates reduce the risk of selection bias. However, patients who were not included due to non-response may differ from included patients in terms of both registered QDs and time to diagnosis. Yet, we believe this bias to be stronger for patients with longer diagnostic interval, which implies that our associations are likely to be underestimated.

A particular challenge in this study was recall bias by both GP and patient; they may both mistakenly have ascribed certain circumstances to a specific diagnostic pathway or neglect influential factors if no impact was found on the diagnosis. Yet, we believe this bias to be stronger for patients with longer diagnostic interval, which implies that our associations are likely to be underestimated.

A particular challenge in this study was recall bias by both GP and patient; they may both mistakenly have ascribed certain circumstances to a specific diagnostic pathway or neglect influential factors if no impact was found on the diagnosis. Yet, we believe this bias to be stronger for patients with longer diagnostic interval, which implies that our associations are likely to be underestimated.

Comparison with existing literature
This study’s finding that both GPs and patients reported a QD in 30% of all cases is slightly lower than the percentages found in studies based on GPs only (19,20). However, these variations may be explained by design differences between the studies. Although previous studies have shown differences between patients and GPs (12,13), we found that GPs and patients in our study reported the same amount of QDs. The reason for this finding is unknown and needs further investigation.

The GPs and patients agreed only slightly on whether reported QDs were found in the same pathway. The degree of agreement in our study is lower than found in other studies, although both good agreement (11–13) and poor agreement (14,15) has been reported in other studies. As all other studies have been conducted in hospital settings, QD assessment in general practice might be more influenced by the respondent’s perspective than it does in secondary care.

One very interesting finding is why the GPs and patient agree on the amount of QDs, while the pairwise agreement was very low. The most plausible explanation is, that patients’ and GPs’ understanding of what constitutes a QD are fundamental different. The GP’s assessment is primarily based on a profound knowledge and experience with the health care. The patient’s assessment is, on the other hand, formed by the particular experience, the expectations and perhaps also what information the patient was given. Yet, more
in depth investigations are needed to understand this finding and the implications.

So far, this study is the first to estimate the association between diagnostic interval and agreement between GP-reported and patient-reported QDs. Our finding that a longer diagnostic interval was found when a QD was reported corresponds well with our previous findings of a 41-day longer diagnostic interval when only GPs were asked to assess QD presence (7). The diagnostic interval was 54 days longer in this study when both patient and GP agreed on QD presence than when no QD was present, which indicates that combining patient-reported QDs with GP-reported QDs adds a small value to the strength of impact of QDs on the diagnostic interval.

Still, if the diagnostic interval is regarded as a clinical outcome, the finding of GP-reported QDs had a stronger impact on the diagnostic interval corresponds well with a study by Basch et al., who found that doctor-reported QDs is a better predictor of unfavourable clinical outcomes, while patient-reported QDs are a better predictor of daily health status (16). The reason for this could be that patients and doctors have different perspectives on what constitutes a QD, and such possible variation needs to be examined by more qualitative methods.

Conclusions

Although GPs and patients reported the same amount of QDs in the diagnostic pathway of cancer, they only slightly agreed on whether it was the same pathway. This difference could indicate that GPs and patients have different perspectives on QDs, which is also supported by the fact that many patients reported that they did not know if any QD had occurred. Furthermore, if the GP reported a QD (either alone or together with the patient), this was a strong indicator of a delayed cancer diagnosis, which was not the case when only the patient reported a QD. In conclusion, if the health care system intends to address quality aspects related to delayed cancer diagnosis, the GP seems to be the most important source of information while patient-based aspects may add some information to this.

Declaration

Funding: Novo Nordisk Foundation; Danish Cancer Society.

Ethical approval: the study was approved by the Danish Data Protection Agency (record number: 2009-41-3471). The study did not require approval from the Committee on Health Research Ethics of the Central Denmark Region as no biomedical intervention was performed.

Conflict of interest: the authors of this manuscript declare to have no competing interests.

Acknowledgements

We would like to thank the cancer patients in Denmark for their time to provide us with highly valuable data for research. We also would like to thank the participating GPs in Denmark for their involvement in the cancer research in Denmark.

References


