Evidence of advanced stage colorectal cancer with longer diagnostic intervals: a pooled analysis of seven primary care cohorts comprising 11,720 patients in five countries

ML Tørring[1*], P Murchie[2], W Hamilton[3], P Vedsted[4], M Esteva[5], M Lautrup[6], M Winget[7], and G Rubin[8]

1) Marie Louise Tørring, *Associate Professor*, Department of Anthropology, School of Culture and Society, Aarhus University, Moesgaard Allé 20, DK-8720 Højbjerg, Denmark. mlt@cas.au.dk

2) Peter Murchie, *Clinical Senior Lecturer*, Centre of Academic Primary Care, Division of Applied Health Sciences, University of Aberdeen, Polwarth Building, Foresterhill, Aberdeen, AB25 2ZD Scotland, UK. p.murchie@abdn.ac.uk

3) William Hamilton, *Professor of Primary Care Diagnostics*, University of Exeter, College House, St Luke's Campus, Magdalen Road, Exeter, EX1 2LU England, UK. W.Hamilton@exeter.ac.uk

4) Peter Vedsted, *Professor and Research Director*, Research Centre for Cancer Diagnosis in Primary Care, Research Unit for General Practice, Aarhus University, Bartholins Allé 20, DK-8000 Aarhus C, Denmark. p.vedsted@ph.au.dk

5) Magdalena Esteva, MD, MSc, PhD, Primary Care Research Unit, Primary Care Majorca Department, Balearic Islands Health Research Institute (IdISBa), Reina Esclaramunda 9, 07003 Palma Mallorca, Spain. mesteva@ibsalut.caib.es

6) Marianne Djernes Lautrup, *Postdoctoral Researcher and Specialist Consultant*, Department of Organ and Plastic Surgery, Breast Centre, Vejle Hospital, Kabbeltoft 25, DK-7100 Vejle, Denmark. marianne.korsgaard@dadlnet.dk

7) Marcy Winget, PhD, *Clinical Associate Professor*, Division of General Medical Disciplines, School of Medicine, Stanford University, 1265 Welch Rd., MSOB #X214, CA 94305 Stanford, California, USA. mwinget@stanford.edu

8) Greg Rubin, *Professor of General Practice and Primary Care*, Durham University, School of Medicine, Pharmacy and Health, Wolfson Research Institute, Queen's Campus, University Boulevard, Stockton on Tees TS17 6BH England, UK. g.p.rubin@durham.ac.uk

*Corresponding author: Marie Louise Tørring. E-mail address: mlt@cas.au.dk; tel: +45 8716 2089.

Running head: Advanced stage colorectal cancer with longer diagnostic intervals

List of abbreviations:

CRC	Colorectal Cancer	PCP	Primary Care Physician
CI	Confidence Interval	IQI	Inter Quartile Interval
NA	Not Available		

Words in text: 3,190; Abstract: 273; Tables: 3; Figures: 3; and References: 42

ABSTRACT

BACKGROUND: The benefits from expedited diagnosis of symptomatic cancer are uncertain. We aimed to analyse the relationship between stage of colorectal cancer (CRC) and the primary and specialist care components of the diagnostic interval.

METHODS: We identified seven independent datasets from population-based studies in Scotland, England, Canada, Denmark and Spain during 1997-2010 with a total of 11,720 newly-diagnosed CRC patients, who had initially presented with symptoms to a primary care physician. Data were extracted from patient records, registries, audits and questionnaires, respectively. Datasets were required to hold information on dates in the diagnostic interval (defined as the time from the first presentation of symptoms in primary care until the date of diagnosis), symptoms at first presentation in primary care, route of referral, gender, age and histologically confirmed stage. We carried out reanalysis of all individual datasets and, using the same method, analyzed a pooled individual patient dataset.

RESULTS: The association between intervals and stage was similar in the individual and combined dataset. There was a statistically significant convex (∩-shaped) association between primary care interval and diagnosis of advanced (i.e. distant or regional) rather than localised CRC (P=0.002), with odds beginning to increase from the first day on and peaking at 90 days. For specialist care we saw an opposite and statistically significant concave (∪-shaped) association, with a trough at 60 days, between the interval and diagnosis of advanced CRC (P<0.001).

CONCLUSION: This study provides evidence that longer diagnostic intervals are associated with more advanced CRC. Furthermore, the study cannot define a specific 'safe' waiting time as the length of the primary care interval appears to have negative impact from day one.

Keywords: Delayed Diagnosis, Waiting Lists, Tumour Staging, Colorectal Cancer, Primary Health Care, Bias.

INTRODUCTION

The diagnostic interval for cancer is defined as the time from first presentation of symptoms in primary care to diagnosis (Weller *et al*, 2012). There is a general public perception – largely shared by health professionals – that prolonged diagnostic intervals lead to poorer cancer outcome (Richards, 2009). Existing epidemiological evidence on the relationship between diagnostic interval and outcomes in colorectal cancer (CRC) is contradictory, most of it suggesting no statistically significant association (Neal *et al*, 2015, Ramos *et al*, 2008, Ramos *et al*, 2007). Many of these studies are undermined by negative bias due to confounding by indication and lead-time, but the lack of clear evidence has nonetheless led some to question the current use of healthcare and research resources to reduce cancer delays (Rupassara *et al*, 2006, Murchie *et al*, 2014).

Most CRC patients present initially to a primary care physician (PCP) and are subsequently referred to specialist cancer centres for assessment and investigation (Banks *et al*, 2014, Hamilton and Sharp, 2004, Hamilton, 2010, Hansen *et al*, 2011). National guidelines such as the English *NICE Guidance*, the Scottish *SIGN* Guidelines and the Danish *Cancer Patient Pathways*, recommend urgent referral of patients with alarm signs and symptoms, including unexplained weight loss, change in bowel habit and/or rectal bleeding (NICE, 2005, SIGN, 2003, Probst *et al*, 2012). However, over half of people subsequently diagnosed with colorectal cancer have no alarm signs or symptoms at initial presentation to a PCP (Hamilton *et al*, 2005). These patients may experience longer diagnostic intervals, perhaps as a result of not being referred until alarm symptoms develop, or because they are ineligible for an urgent appointment (Jensen *et al*, 2014a, Neal *et al*, 2014).

We propose a temporal and relational way of understanding the association between time to diagnosis and CRC outcomes. In short, the temporal idea is that cancers become gradually easier to detect with time due to the exponential growth of tumours. The relational idea is that doctors exert their ability to identify cancer by virtue of the role they perform within the relational network to which they belong. It is our hypothesis that symptomatic cancer patients are being sorted and diagnosed at different pace based on their gradually changing clinical indications, and on the respective diagnostic tools available in primary and specialist care. This must be taken into account when analysing the prognostic influence of delays.

We identified seven CRC cohort studies to create a dataset sufficiently large to investigate the association between time to diagnosis and CRC outcomes. To call attention to the inevitable problem of confounding by indication and explain continued lack of evidence, we analysed stage of CRC as a flexible function of the length of time under PCP care and specialist care, respectively.

METHODS

We identified seven studies with datasets from Scotland, England, Canada, Denmark and Spain from population-based studies of incident CRC patients diagnosed between 1997 and 2010. Datasets were required to hold information on dates in the diagnostic interval (Weller *et al*, 2012), symptoms at first presentation in primary care, route of referral, gender, age and stage. All but one (ALBERTA) have published their results (Murchie *et al*, 2014, Robertson *et al*, 2004, Hamilton *et al*, 2005, Stapley *et al*, 2006, Korsgaard *et al*, 2006b, Korsgaard *et al*, 2006a, Jensen *et al*, 2014b, Esteva *et al*, 2007, Esteva *et al*, 2013, Rubin *et al*, 2011, Lyratzopoulos *et al*, 2013). The individual studies are summarized in Table 1 and more extensively described in the online supplemental material I.

Study population

All cohorts described in this study were from jurisdictions with publicly funded, universal health care systems, in which PCPs act as gatekeepers from primary to specialist care. Newly-diagnosed CRC patients were identified in discharge registries, hospital or primary care records. From each dataset we included all patients aged 20 years and over who had attended primary care before the cancer diagnosis. Screen-detected cases were excluded, as were those diagnosed as a result of emergency presentation with no prior contact with primary care. To render the datasets uniform, we only included patients with recorded CRC symptoms or CRC related visits in the year before diagnosis (Stapley *et al*, 2006).

Defining exposures, outcome and covariates

We defined and calculated three exposure variables as illustrated on Figure 1: 1) The primary care interval (time from first presentation to referral to a cancer specialist centre); 2) the secondary care interval (time from referral to diagnosis); and 3) the total diagnostic interval (time from first presentation to diagnosis) based on information on date of first presentation of symptoms in primary care and date of referral to a cancer specialist centre, and date of diagnosis, as defined in the Aarhus Statement (Weller *et al*, 2012). All datasets, except the NACDPC, recorded the date of diagnosis, defined as date of first histological confirmation of the malignancy in accordance with the hierarchy produced by the European Network of Cancer Registries (supplementary material I).

The primary outcome of the study was stage of CRC as defined by the local registry or provincial cancer registry in the case of the ALBERTA cohort (Table 1). All datasets included histologically confirmed stage, except the NACDPC dataset which used a simplified staging equating to SEER stages 0-3 that was determined by the PCP using information contained in cancer specialist letters. We used T, N and M and not Dukes' stage where both were available. Stage data were re-categorized to localised, regional, distant, or unknown; and then further simplified to a binary variable of advanced (i.e., distant or regional) vs. localised CRC (Supplementary table 1 in supplementary material II shows the algorithm we used for colorectal cancer staging according to classification system).

All datasets included data on gender, age at diagnosis and presenting symptoms. 'Alarm symptoms at first presentation' (yes/no), were defined based on whether patients' reported symptom(s) merited urgent referral according to the UK's NICE guidelines (Neal *et al*, 2014, NICE, 2005) (Supplementary table 2 in supplementary material III shows the pre-specified lists of colorectal cancer symptoms recorded in each dataset).

Statistical analyses

Each dataset was analysed separately, then combined for pooled individual patient data analysis. Time intervals are presented as medians with interquartile intervals (IQI). To test for confounding factors related to clinical triage, we estimated the difference in median care intervals between patients with and without 'alarm symptoms' and 'emergency admission' or not using quantile regression analysis (Supplementary table 3 and supplementary figure 4-5, online supplementary material IV).

To avoid assuming a linear or piecewise constant association between care intervals and stage, we treated primary and secondary care intervals as continuous variables, using restricted cubic splines with three knots and 30 days as the reference point (Durrleman and Simon, 1989). We estimated the odds ratio of being diagnosed with advanced vs. localised CRC as a function of the length of each care interval using logistic regression. We adjusted for age (20-64/65-74/≥75 years), gender and alarm symptoms. With the combined data, we allowed for between-dataset variability by adjusting for cohort. We combined interval data with no attention to distributions or weights. Eight sensitivity analyses tested the robustness of the model (Supplementary table 5, online supplementary material IV).

We tested each model against a model with no care interval term using the Wald test. A two-sided p-value of 0.05 or less was defined as significant. Analyses were done using Stata® v. 14 (StataCorp LP, College Station, TX, USA).

RESULTS

In total, the seven datasets included 15,023 incident CRC patients. Of these, 1,584 (11%) did not consult a PCP before their diagnosis of CRC, and 1,719 had incomplete data; these were excluded. Of the remaining 11,720 patients, 11,187 (95%) had information on primary care interval; 9,163 (78%) on secondary care interval and 9,696 (83%) on the total diagnostic interval. Stage information was available for 92% (Table 1 and 2).

The median age for all patients combined was 70 years; 56% were males, 59% presented with alarm symptoms, 20% were emergency hospital admissions, 61% had colon and 39% rectal cancer; 44% had localised CRC. Clinical features were remarkably similar for each of the seven cohorts, except for the proportion of patients presenting with alarm symptoms which varied from 49% (DECIRRE) to 78% (CAPER); and the proportion of emergency admissions which varied from 10% (CRCDK) to 43% (DECCIRE) (Table 3).

Total diagnostic interval and its primary care and secondary care components

The overall median primary care interval was 5 days (interquartile interval (IQI): 0-39), ranging from 2 days (IQI: 0-21) in CAP to 14 days (IQI: 0-64) in CRUX. Thirty nine percent of patients were referred immediately after presentation (i.e. primary care interval of zero days) varying from 14% in CRUX to 46% in the CAP and ALBERTA cohorts (~~online Figure 5~~ Supplementary figure 1, online supplementary material IV). The overall median secondary care interval was 20 days (IQI: 7-46) and varied from 14 days (IQI: 7-29) in CRCDK to 38 days (IQI: 17-82) in DECCIRE (Table 3 and Supplementary figure 2, online supplementary material IV~~online Figure 6~~). The overall median diagnostic interval was 46 days (IQI: 18–105) ranging from 35 days in CAP to 97 days in CAPER (Table 3 and Supplementary figure 3, online supplementary material IV~~online Figure 7~~).

The adjusted care intervals were significantly shorter for patients with alarm symptoms and for patients with emergency admissions. Overall, patients waited an additional 6 (95% CI: 4-7) days from presentation to referral and an additional 9 (95% CI: 7-10) days from referral to diagnosis at the 50th percentile if they had no alarm symptoms. An emergency admission shortened the secondary care interval by 18 (95% CI: 17-19) days at the 50th percentile. Alarm symptoms consistently shortened the length of the primary care interval in each cohort, whereas emergency admissions more convincingly affected the length of the secondary care interval (Supplementary table 3 and Supplementary figure 4-5, online supplementary material IV).

Primary and secondary care diagnostic intervals and stage of CRC

For the combined cohort, we observed a significant trend for a concave, ∩-shaped association with increasing and subsequently decreasing odds of advanced CRC with longer primary care intervals (P=0.004). The pointwise estimates showed that the adjusted odds of being diagnosed with an advanced stage tumour increased from the first day, and were around 8% (95% CI: 2-12%) higher for patients who waited 90 days compared to 30 days from first presentation to referral (Figure 2, blue curve). For the secondary care interval we saw the reverse effect: a significant ∪-shaped association with decreasing and subsequently increasing

odds of advanced CRC with longer secondary care intervals (*P*<0.001) (Figure 2, red curve). Crude estimates were similar to the adjusted curves on Figure 2 (not shown).

The cohort-specific associations were consistent with the overall trend (Figure 3 and 4), but the primary care interval model was only significant for CRUX (*P*=0.03); and the secondary care interval model was only significant for ALBERTA (*P*<0.001) and CAP (*P*<0.001) data. For the total diagnostic interval we found decreasing odds of advanced CRC with longer intervals (P<0.001), but not achieving significance in individual cohorts (Supplementary figure 6, online supplementary material IV).

The findings were similar after including patients with unknown tumour stage (915 patients) or restricting the analysis to patients with/without alarm symptoms or with/without emergency admission. Excluding the 41% of patients with zero days of primary care interval (zero-inflation) decreased the primary care interval trend. These sensitivity analyses are shown in Supplementary table 5, online supplementary material IV.

DISCUSSION

In this unique pooled individual patient data analysis of seven cohorts with 11,720 incident CRC patients attending primary care with symptoms before diagnosis, the odds of advanced CRC increased with longer primary care intervals up to ~90 days, after which the odds decreased. Conversely, the odds of advanced CRC decreased with longer secondary care intervals up to ~60 days, after which the odds slowly increased. Both of these associations were statistically significant and the trends were consistent across datasets. No clear trend was noted across individual cohorts for the total diagnostic intervals.

Strengths and limitations

The key strength of this study lies in the large number of cases, from seven datasets in five countries with comparable systems of universal health care. By excluding screen-detected patients and those without PCP involvement, the results are relevant to all healthcare systems featuring primary care gate-keeping. However, the proportions of patients presenting as an emergency are not routinely recorded outside the UK, so we cannot know if varying sizes of this proportion have affected our results. The data were sufficiently detailed to enable care intervals to be analysed as a continuous variable, enabling us to utilise recent methodological advances in a larger sample (Tørring *et al*, 2011, Tørring *et al*, 2012). A fundamental strength of the study is that lead time bias cannot explain results, as we used stage as our outcome measure, rather than survival time.

A number of limitations exist due to the cross-sectional study design, which does not permit direct inference of causality. Dates for interval calculations may have been recorded systematically differently across different datasets. Although the ascertainment of dates accorded with best practice in all cases (Weller *et al*, 2012), no gold standard data source exists for interval measures. Studies show good agreement between patient and PCP recorded dates of diagnosis (Adelstein *et al*, 2008); but dates of first presentation are usually more accurately recorded for alarm symptoms than for non-specific symptoms (Lynch *et al*, 2008). Furthermore, registry-based interval measures may be affected by the data collection methods, with some symptoms unrecorded, while others are potentially unrelated to the cancer (Stapley *et al*, 2006, Tate *et al*, 2009). Registry-based studies are highly dependent on underlying algorithms and cut-points; they benefit from high inclusion rates, but potentially have non-differential misclassification of date of first presentation. To ensure that registry-based recorded symptoms were likely to be related to the cancer, we only included recorded CRC symptoms or CRC-related visits in the year before diagnosis. Even so, registry-based measures were longer than PCP-reported time points (Table 3, and Supplementary figure 1-3 in supplementary material IV).

Missing information bedevils all studies on staging, and may have biased results if the quality of staging was associated with diagnostic timeliness. Although we used benchmarked registries and approaches to produce comparable stage information (Benitez-Majano *et al*, 2016, Ostenfeld *et al*, 2012, Tucker *et al*, 1999,

Walters *et al*, 2013), it is conceivable that there was some bias to exclusion of more advanced cases in some datasets, reflecting those dying during admission or treatment. The main effect of such information bias would be increased variation and fewer cases with short intervals and advanced stages, and thus towards no association between time and stage. Hence, our estimates are perhaps underestimates of the real association.

Despite differences in data sources and construction, the clinical features of patients were remarkably similar for each of the seven individual datasets, and they produced strikingly similar results suggesting that selection and information biases were not major methodological limitations.

Unmeasured confounding by factors such as socioeconomic status, co-morbidity, or tumour grade/aggressiveness, which were not universally available, may have influenced the results. We partly mitigated this by adjusting for alarm symptoms and emergency admissions, reducing confounding by indication.

Given the observed trends, it is difficult to predict the direction of bias. However, most of the potential biases from selection, information, confounding, and confounding by indication inherited from the different study designs are likely to have caused negative bias towards and even beyond the null hypothesis (i.e. odds ratio = 1).

When interpreting the results from the combined datasets, the weight of information contributed by each study is relevant. In a sensitivity analysis excluding ALBERTA (50-61% of subjects) we found almost identical results (Supplementary table 5, online supplementary material IV).

Comparison with findings from other studies

The study builds on a recent systematic review of the association between time to diagnosis and cancer outcome for all types of cancers, which called for higher quality and larger studies that addressed basic issues of bias (Neal *et al*, 2015). Few studies have considered the possibility that the associations may vary for different components of the diagnostic pathway (Afzelius *et al*, 1994, Crawford *et al*, 2002). They support our interpretation that the basis for assignment of delay (the sorting of patients) change during the diagnostic pathway and that interval-specific models are necessary to achieve valid comparisons of delay and mortality. Our findings also confirm a non-monotonic relationship between delay and stage, and thus consolidate important points made by Maguire and others, but in a much larger cohort enabling more valid models (Maguire *et al*, 1994, Murchie *et al*, 2014, Tørring *et al*, 2011, Tørring *et al*, 2012).

Underlying mechanisms

It is a widely held assumption that the waiting time paradox can be explained by the effect of high-risk precursors such as phenotype, biological virulence or tumour aggressiveness, which are thought to act as unmeasured confounders that mask the effect of the exposure (Afzelius et al, 1994, Crawford et al, 2002, Neal et al, 2015, Symonds, 2002). However, some studies have shown that a significant proportion of cancers

present with symptoms that are vague or non-specific, with the underlying problem more difficult to detect and act upon in time (Jensen et al, 2014a, Korsgaard et al, 2006b). It is reasonable to assume that symptomatic cancer patients are sorted and diagnosed at different pace based on their gradually changing clinical indications, and on the respective diagnostic tools available in primary and specialist care.

We believe the finding of a two-sided 'waiting time paradox' (lower odds of advanced CRC for very long primary care intervals and higher odds of advanced CRC for very short secondary care intervals) reflects confounding by indication, a bias stemming from the inherent difference in the prognosis of patients given different medical priority in primary care versus specialist care.

In primary care, PCPs will expedite patients using a fast-track cancer patient pathway or emergency admission if the patient's ill-health is obvious. At the same time, PCPs may be more reluctant to refer those with low-risk symptoms, leading to use of normal waiting list referral or watchful-waiting. Since many patients in primary care do not fall squarely into either of these categories, delays should be less contingent upon prognosis and thus more randomly distributed in the large group of intermediate patients. We, therefore, propose that the observed increasing odds of advanced CRC with longer primary care intervals up to ~90 days reflects the actual effect of primary care delays.

In contrast, when a patient is first seen in specialist care, primary care triage has already taken place – making the probability of CRC higher. Furthermore, the specialist's greater clinical experience of patients with CRC and ready access to hospital-based investigations ensures that patients with advanced disease attending specialist care are diagnosed and treated very quickly; those with less advanced disease are managed less urgently. Hence, negative bias (where the observed effect is lower than the true value) probably explains the decreasing odds of advanced CRC with time and may explain why many CRC studies to date have failed to reject the null-hypothesis of no association between delays and outcome (Neal *et al*, 2015, Ramos *et al*, 2008, Ramos *et al*, 2007). As with previous studies, we found no clear association between the total diagnostic interval and stage, presumably due to the conflicting selection-effects of primary care and secondary care intervals. As a final, tentative point we propose that the observation of increasing odds of advanced CRC with secondary care intervals longer than ~60 days could reflect the effect of false negative tests or unnecessary delays in investigation and/or treatment.

Clinical implications

We cannot define a specific 'safe' waiting time as the length of the primary care interval appears to have a negative impact from day one. It follows that patients with CRC, without alarm symptoms at presentation, are most at risk of a prolonged diagnostic interval.

Cancer diagnostic delays cannot be completely eradicated, so resources must be used proportionately to the objective of finding cancer sooner. Reducing the primary care interval by lowering the threshold for

urgent referral and enabling easier direct access to investigations by PCPs, may provide the greatest benefit, as other recent studies indicate (Maclean *et al*, 2015, Moller *et al*, 2015).

The study displays the immense complexity and difficulty of diagnosing cancer. Further research on similar combined datasets from longitudinal studies, using the same novel analytical approach, should now be conducted to confirm the relative impact of primary and secondary care diagnostic intervals on outcomes in patients with other cancers.

Conclusion

This study provides evidence that longer primary care and secondary care intervals are associated with more advanced CRC. The finding of similar trends when using different sources of information, for different time periods and in different health care systems (of Scotland, England, Canada, Denmark and Spain) strengthens the belief that the results can be generalised to other health care systems around the world.

ADDITIONAL INFORMATION

Acknowledgements

The authors are all members of the Cancer in Primary Care Research International (CaPRI) Network. We are thankful to the CRUX, CAPER, ALBERTA, CRC-DK, DECCIRE, CAP and NACDPC research teams for contributing data as well as detailed knowledge about the way it was collected. We are equally thankful to the participating general practices, hospitals, service agencies, and patients for their time and involvement.

In addition, we would like to express our gratitude to associate professor Morten Frydenberg from the Department of Biostatistics at Aarhus University for statistical support and making programmes available for generating cubic splines with specific reference values (centercsplines.ado) and estimates with standard errors for linear combinations (calcest.ado) in Stata (applied in Figure 2-4 and in Supplementary figure 6). To bioinformatics Martin Krzywinski at the Genome Science Centre in Vancouver for his advice on improving the data visualization of Figure 2. To Jochim Terhaar sive Droste from the Department of Gastroenterology and Hepatology, VU University Medical Centre, Amsterdam for his initial collaboration and willingness to share data from the Northern Holland cohort.

The workgroup sincerely thanks Dr. Fiona Walter and the Early Diagnosis Group at Cambridge's Primary Care Unit for providing a perfect space for thinking, working and writing together in the autumn of 2014.

Funding

The Ca-Pri Colorectal Cancer Collaboration project received no formally targeted donations, but, the primary author MLT coordinated the work as part of her postdoctoral studies at the Research Centre for Cancer Diagnosis in Primary Care (CaP), funded by the Novo Nordisk Foundation and the Danish Cancer Society. The **CRUX** study was funded by NHS Grampian Research Endowment Award 11/26, and a grant from The Colorectal Study Fund (a NHS Grampian Endowment fund). The **CAPER** study was funded by the Department of Health, UK. The **CRC-DK** study was supported by grants from Western Danish Research Forum, Danish Medical Research Council, Dagmar Marshall's Fund and the Danish Cancer Society. The **ALBERTA** was funded by the National Cancer Institute of Canada, Alberta Cancer Foundation, and the Canadian Institute of Health Research. The **CAP** study was supported by the Novo Nordisk Foundation, the Danish Cancer Society, the Health Foundation, the Tryg Foundation, and the Central Denmark Region's "Praksisforskningsfond". The **DECCIRE** study was financed with grants from the Ministry of Health, Carlos III Institute, and also received support from the Health Promotion and Preventive Activities-Primary Health Care Network, sustained by the Ministry of Health in Spain. The **NACDPC** study was financed by the Department of Health, England. The sponsors were not involved in any part of the studies.

No competing interests

We have read and understood the journal's policy on declaration of interests and declare that we have no competing interests.

Contributors

GR and MLT initiated the Ca-PRI Colorectal Cancer Collaboration Workgroup and came up with the idea for the study. MLT, PM and GR evolved the study design in subsequent meetings and discussions. All authors collaborated on drafting an analysis plan and prepared their individual datasets in accordance with the agreed protocol. The author and data custodian of the CRCDK data, M. Lautrup is identical with M. Korsgaard as cited in the reference list. MLT co-ordinated communication with and data acquisition from involved research groups, performed the statistical analyses, and wrote the paper together with PM. She (MLT) is the guarantor. GR, PV and WH provided critical revision of the intellectual content of the manuscript. All authors had full access to their own dataset and take full responsibility for the accuracy of the data description and analysis. All authors contributed to the writing of and approved the final manuscript.

Transparency declaration

The guarantor, MLT, affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that all discrepancies from the study as planned have been explained.

Ethical approval

MLT was granted permission to analyse data according to the laws in each country. The **CRUX** study was approved by the Privacy Advisory Committee of ISD Scotland. According to the North of Scotland Research Ethics Committee formal ethical approval was not required. Research and Development approval was granted from NHS Grampian. In the **CAPER** study ethical approval was obtained from North and East Devon local research ethics committee. All patients in the **CRC-DK** study signed an informed consent form and the study was approved by the Danish Data Protection Agency and the Ethical Committees of Aarhus, Ringkoebing and Ribe counties. The principle investigator of the **ALBERTA** dataset obtained necessary regulatory approvals to send data to MLT. The **CAP** study was approved by the Danish Data Protection Agency. The Danish Health and Medicines Authority gave permission to obtain information from the PCPs' medical records. Health research ethics was not required. The **DECCIRE** study was approved by the Primary Health Care Committee of each health district and by the Ethical and Clinical Research Committee of each participating region in Spain. The **NACDPC** audit was approved as an audit by the Research and Audit committees of the Primary Care Trusts of participating practices. Permission to use anonymised data for research purposes was granted by the NACDPC oversight group.

REFERENCES

Adelstein BA, Irwig L, Macaskill P, Katelaris PH, Jones DB, Bokey L (2008) A self administered reliable questionnaire to assess lower bowel symptoms. *BMC Gastroenterol* **8**: 8-230X-8-8, doi:10.1186/1471-230X-8-8 [doi]

Afzelius P, Zedeler K, Sommer H, Mouridsen HT, Blichert-Toft M (1994) Patient's and doctor's delay in primary breast cancer. Prognostic implications. *Acta Oncol* **33**: 345-351

Banks J, Walter FM, Hall N, Mills K, Hamilton W, Turner KM (2014) Decision making and referral from primary care for possible lung and colorectal cancer: a qualitative study of patients' experiences. *Br J Gen Pract* **64**: e775-82, doi:10.3399/bjgp14X682849 [doi]

Benitez-Majano S, Fowler H, Maringe C, Di Girolamo C, Rachet B (2016) Deriving stage at diagnosis from multiple population-based sources: colorectal and lung cancer in England. *Br J Cancer* **115**: 391-400, doi:10.1038/bjc.2016.177 [doi]

Crawford SC, Davis JA, Siddiqui NA, de Caestecker L, Gillis CR, Hole D, Penney G (2002) The waiting time paradox: population based retrospective study of treatment delay and survival of women with endometrial cancer in Scotland. *BMJ* **325**: 196

Durrleman S and Simon R (1989) Flexible regression models with cubic splines. *Stat Med* **8**: 551-561

Esteva M, Leiva A, Ramos M, Pita-Fernandez S, Gonzalez-Lujan L, Casamitjana M, Sanchez MA, Pertega-Diaz S, Ruiz A, Gonzalez-Santamaria P, Martin-Rabadan M, Costa-Alcaraz AM, Espi A, Macia F, Segura JM, Lafita S, Arnal-Monreal F, Amengual I, Bosca-Watts MM, Hospital A, Manzano H, Magallon R, DECCIRE GROUP (2013) Factors related with symptom duration until diagnosis and treatment of symptomatic colorectal cancer. *BMC Cancer* **13**: 87, doi:10.1186/1471-2407-13-87; 10.1186/1471-2407-13-87

Esteva M, Ramos M, Cabeza E, Llobera J, Ruiz A, Pita S, Segura JM, Cortes JM, Gonzalez-Lujan L, DECCIRE research group (2007) Factors influencing delay in the diagnosis of colorectal cancer: a study protocol. *BMC Cancer* **7**: 86, doi:1471-2407-7-86 [pii]

Hamilton W (2010) Cancer diagnosis in primary care. *Br J Gen Pract* **60**: 121-128, doi:10.3399/bjgp10X483175

Hamilton W, Round A, Sharp D, Peters TJ (2005) Clinical features of colorectal cancer before diagnosis: a population-based case-control study. *Br J Cancer* **93**: 399-405, doi:10.1038/sj.bjc.6602714

Hamilton W and Sharp D (2004) Diagnosis of colorectal cancer in primary care: the evidence base for guidelines. *Fam Pract* **21**: 99-106

Hansen RP, Vedsted P, Sokolowski I, Sondergaard J, Olesen F (2011) Time intervals from first symptom to treatment of cancer: a cohort study of 2,212 newly diagnosed cancer patients. *BMC Health Serv Res* **11**: 284, doi:10.1186/1472-6963-11-284

Jensen H, Torring ML, Olesen F, Overgaard J, Vedsted P (2014a) Cancer suspicion in general practice, urgent referral and time to diagnosis: a population-based GP survey and registry study. *BMC Cancer* **14**: 636-2407-14-636, doi:10.1186/1471-2407-14-636 [doi]

Jensen H, Torring ML, Larsen M,Bach, Vedsted P (2014b) Existing data sources for clinical epidemiology: Danish Cancer in Primary Care cohort. *Clin Epidemiol* **6**: 237-246, doi:10.2147/CLEP.S62855 [doi]

Korsgaard M, Pedersen L, Sorensen HT, Laurberg S (2006a) Delay of treatment is associated with advanced stage of rectal cancer but not of colon cancer. *Cancer Detect Prev* **30**: 341-346

Korsgaard M, Pedersen L, Sorensen HT, Laurberg S (2006b) Reported symptoms, diagnostic delay and stage of colorectal cancer: a population-based study in Denmark. *Colorectal Dis* **8**: 688-695

Lynch BM, Youlden D, Fritschi L, Newman B, Pakenham KI, Leggett B, Owen N, Aitken JF (2008) Self-reported information on the diagnosis of colorectal cancer was reliable but not necessarily valid. *J Clin Epidemiol* **61**: 498-504, doi:10.1016/j.jclinepi.2007.05.018 [doi]

Lyratzopoulos G, Abel GA, McPhail S, Neal RD, Rubin GP (2013) Gender inequalities in the promptness of diagnosis of bladder and renal cancer after symptomatic presentation: evidence from secondary analysis of an English primary care audit survey. *BMJ Open* **3**: 10.1136/bmjopen-2013-002861. Print 2013, doi:10.1136/bmjopen-2013-002861; 10.1136/bmjopen-2013-002861

Maclean R, Jeffreys M, Ives A, Jones T, Verne J, Ben-Shlomo Y (2015) Primary care characteristics and stage of cancer at diagnosis using data from the national cancer registration service, quality outcomes framework and general practice information. *BMC Cancer* **15**: 500-015-1497-1, doi:10.1186/s12885-015-1497-1 [doi]

Maguire A, Porta M, Malats N, Gallen M, Pinol JL, Fernandez E (1994) Cancer survival and the duration of symptoms. An analysis of possible forms of the risk function. **30**: 785-792

Moller H, Gildea C, Meechan D, Rubin G, Round T, Vedsted P (2015) Use of the English urgent referral pathway for suspected cancer and mortality in patients with cancer: cohort study. *BMJ* **351**: h5102, doi:10.1136/bmj.h5102 [doi]

Murchie P, Raja EA, Brewster DH, Campbell NC, Ritchie LD, Robertson R, Samuel L, Gray N, Lee AJ (2014) Time from first presentation in primary care to treatment of symptomatic colorectal cancer: effect on disease stage and survival. *Br J Cancer* **111**: 461-469, doi:10.1038/bjc.2014.352 [doi]

Neal RD, Din NU, Hamilton W, Ukoumunne OC, Carter B, Stapley S, Rubin G (2014) Comparison of cancer diagnostic intervals before and after implementation of NICE guidelines: analysis of data from the UK General Practice Research Database. *Br J Cancer* **110**: 584-592, doi:10.1038/bjc.2013.791 [doi]

Neal RD, Tharmanathan P, France B, Din NU, Cotton S, Fallon-Ferguson J, Hamilton W, Hendry A, Hendry M, Lewis R, Macleod U, Mitchell ED, Pickett M, Rai T, Shaw K, Stuart N, Torring ML, Wilkinson C, Williams B, Williams N, Emery J (2015) Is increased time to diagnosis and treatment in symptomatic cancer associated with poorer outcomes? Systematic review. *Br J Cancer* **112 Suppl**: S92-S107, doi:10.1038/bjc.2015.48 [doi]

NICE (2005) Referral guidelines for suspected cancer.

Ostenfeld EB, Froslev T, Friis S, Gandrup P, Madsen MR, Sogaard M (2012) Completeness of colon and rectal cancer staging in the Danish Cancer Registry, 2004-2009. *Clin Epidemiol* **4 Suppl 2**: 33-38

Probst HB, Hussain ZB, Andersen O (2012) Cancer patient pathways in Denmark as a joint effort between bureaucrats, health professionals and politicians--a national Danish project. *Health Policy* **105**: 65-70, doi:10.1016/j.healthpol.2011.11.001

Ramos M, Esteva M, Cabeza E, Campillo C, Llobera J, Aguilo A (2007) Relationship of diagnostic and therapeutic delay with survival in colorectal cancer: a review. *Eur J Cancer* **43**: 2467-2478, doi:10.1016/j.ejca.2007.08.023

Ramos M, Esteva M, Cabeza E, Llobera J, Ruiz A (2008) Lack of association between diagnostic and therapeutic delay and stage of colorectal cancer. *Eur J Cancer* **44**: 510-521, doi:10.1016/j.ejca.2008.01.011

Richards MA (2009) The National Awareness and Early Diagnosis Initiative in England: assembling the evidence. *Br J Cancer* **101 Suppl 2**: S1-4, doi:10.1038/sj.bjc.6605382; 10.1038/sj.bjc.6605382

Robertson R, Campbell NC, Smith S, Donnan PT, Sullivan F, Duffy R, Ritchie LD, Millar D, Cassidy J, Munro A (2004) Factors influencing time from presentation to treatment of colorectal and breast cancer in urban and rural areas. *Br J Cancer* **90**: 1479-1485, doi:10.1038/sj.bjc.6601753 [doi]

Rubin G, McPhail S, Elliott K (2011) National Audit of Cancer Diagnosis in Primary Care.

Rupassara KS, Ponnusamy S, Withanage N, Milewski PJ (2006) A paradox explained? Patients with delayed diagnosis of symptomatic colorectal cancer have good prognosis. *Colorectal Dis* **8**: 423-429, doi:10.1111/j.1463-1318.2006.00958.x

SIGN SIGN (2003) Management of CRC: a national clinical guideline. **67**

Stapley S, Peters TJ, Sharp D, Hamilton W (2006) The mortality of colorectal cancer in relation to the initial symptom at presentation to primary care and to the duration of symptoms: a cohort study using medical records. *Br J Cancer* **95**: 1321-1325, doi:10.1038/sj.bjc.6603439

Symonds RP (2002) Cancer biology may be more important than diagnostic delay... Crawford SC, Davis JA, Siddiqui NA et al. The waiting time paradox: population based retrospective study of treatment delay and survival of women with endometrial cancer in Scotland. BMJ 2002;325:196 (27 July). *BMJ* **325**: 774-774

Tate AR, Martin AG, Murray-Thomas T, Anderson SR, Cassell JA (2009) Determining the date of diagnosis--is it a simple matter? The impact of different approaches to dating diagnosis on estimates of delayed care for ovarian cancer in UK primary care. *BMC Med Res Methodol* **9**: 42-2288-9-42, doi:10.1186/1471-2288-9-42 [doi]

Tørring ML, Frydenberg M, Hansen RP, Olesen F, Hamilton W, Vedsted P (2011) Time to diagnosis and mortality in colorectal cancer: a cohort study in primary care. *Br J Cancer* **104**: 934-940

Tørring ML, Frydenberg M, Hansen RP, Olesen F, Vedsted P (2012) Diagnostic interval and mortality in colorectal cancer: U-shaped association demonstrated for three different datasets. *J Clin Epidemiol* **65**: 669-678

Tucker T,C, Howe H,L, Weir H,K (1999) Certification for Population-Based Cancer Registries. *J Registry Manag* **26**: 24-27

Walters S, Maringe C, Butler J, Brierley JD, Rachet B, Coleman MP (2013) Comparability of stage data in cancer registries in six countries: lessons from the International Cancer Benchmarking Partnership. *Int J Cancer* **132**: 676-685, doi:10.1002/ijc.27651 [doi]

Weller D, Vedsted P, Rubin G, Walter FM, Emery J, Scott S, Campbell C, Andersen RS, Hamilton W, Olesen F, Rose P, Nafees S, van Rijswijk E, Hiom S, Muth C, Beyer M, Neal RD (2012) The Aarhus statement: improving design and reporting of studies on early cancer diagnosis. *Br J Cancer* **106**: 1262-1267, doi:10.1038/bjc.2012.68; 10.1038/bjc.2012.68

FIGURE LEGENDS

Figure 1: Definition of exposure variables

We calculated three exposure variables based on information on date of first presentation of symptoms in primary care (B); date of referral to a cancer specialist centre (C); and date of diagnosis (D). "The primary care interval" is defined as B-C = time from first presentation to referral to a cancer specialist centre. "The secondary care interval" is defined as C-D = time from referral to diagnosis). "The total diagnostic interval" is defined as B-D = time from first presentation to diagnosis.

Figure 2: The risk of being diagnosed with advanced colorectal cancer as a function of time to diagnosis.

Estimated odds ratios of being diagnosed with advanced (distant or regional) versus localised colorectal cancer as a function of the length of the primary care interval (blue) and the secondary care interval (red) analysed for all cohorts combined (patients with unknown tumour stage excluded). We adjusted for age, gender, alarm symptoms and cohort. The area around the fitted curves indicates 95% confidence limits. The spikes below the curves show the distribution of the primary care interval (blue) and secondary care interval (red) on a squared scale. The grey horizontal lines indicate the chosen reference point of 30 days (see logistic regression details in Supplementary table 4, supplementary material IV). Crude estimates are not shown.

Figure 3: The risk of being diagnosed with advanced colorectal cancer as a function of time from presentation to referral.

Estimated odds ratios of being diagnosed with advanced (distant or regional) versus localised colorectal cancer as a function of the length of the primary care interval (time from first presentation of symptoms to referral); analysed for six cohorts (in total 10,333 patients). We excluded patients with unknown tumour stage excluded and adjusted for age, gender and alarm symptoms at first presentation. The grey dashed curves with 95% confidence limits are fitted on the combined datasets with grey spikes showing the distribution of the care intervals on a squared scale. The grey horizontal lines indicate the chosen reference point of 30 days. Crude estimates are not shown.

Figure 4: The risk of being diagnosed with advanced colorectal cancer as a function of time from referral to diagnosis

Estimated odds ratios of being diagnosed with advanced (distant or regional) versus localised colorectal cancer as a function of the length of the secondary care interval (time from referral to diagnosis) analysed for five cohorts (in total 8,415 patients). We excluded patients with unknown tumour stage and adjusted for age, gender and alarm symptoms at first presentation. The grey dashed curves with 95% confidence limits are fitted on the combined datasets with grey spikes showing the distribution of the care intervals on a squared scale. The grey horizontal lines indicate the chosen reference point of 30 days. Crude estimates are not shown.

TABLE LEGENDS

Table 1: Study characteristics of seven colorectal cancer cohort datasets

Study characteristics of seven colorectal cancer cohort dataset. Abbreviations: CRC, colorectal cancer; PCP, primary care physician; ICD-10, International Classification of Diseases 10th revision; ICD-9 International Classification of Diseases 9th revision; ICD-O-3, WHO International Classification of Disease for Oncology; ICPC, International Classification of Primary Care; TNM, Tumor, Node, Metastasis; SEER, Surveillance, Epidemiology and End Results (cancer reporting standard of the National Cancer Institute).

Table 2: Patient flow for each colorectal cancer cohort dataset and all data combined

Patient flow for each colorectal cancer cohort dataset and all data combined (far right). Abbreviations: CRC, colorectal cancer; CRUX, Comparing Rural and Urban Cancer Care; CAPER, Cancer Prediction in Exeter; CRCDK, Colorectal Cancer in Denmark; CAP, Cancer in Primary Care; DECCIRE, Delay Cancer Colon i Recto; NACDPD, National Audit of Cancer Diagnosis in Primary Care.

Table 3: Clinical features for patients attending general practice before diagnosis displayed for each colorectal cancer cohort dataset and all data combined

Clinical features for patients attending general practice before diagnosis displayed for each colorectal cancer cohort dataset and all data combined. Abbreviations: NA= Not Available; * In the DECCIRE dataset, the primary and secondary care intervals were only recorded in the Baleares, Galicia, Valencia and Catalunya regions (n=250). ** Emergency admission was only recorded in the 1st subcohort of the CAP dataset (n=272).

ONLINE SUPPLEMENTARY MATERIAL (titles and legends)

Supplementary material I: Information on data collection

Supplementary material II: Information on staging

Supplementary table 1: Algorithm for colorectal cancer staging according to classification system
Algorithm for colorectal cancer staging according to classification system. Abbreviation: TNM, tumor, node, metastasis. SEER, Surveillance, Epidemiology and End Results of the National Cancer Institute (applied from Ostenfeld et al. 2012).

Supplementary material III: Information on symptoms

Supplementary table 2: Pre-specified lists used to record symptoms of colorectal cancer at first presentation in the year before diagnosis

Supplementary material IV: Descriptive statistics and sensitivity analyses

Supplementary figure 1: Distribution of primary care interval (days from first presentation of symptoms in primary care and until referral), total N=11,187.

Supplementary figure 2: Distribution of secondary care intervals (days from referral to diagnosis), total N= 9,163.

Supplementary figure 3: Distribution of total diagnostic interval (days from first presentation of symptoms to diagnosis), total N= 9,696.

Supplementary figure 4: Distribution of primary care and secondary care intervals according to alarm symptom at first presentation
Distribution of primary care interval (blue); and secondary care intervals (red) according to alarm symptom at first presentation for the combined colorectal cancer data.

Supplementary figure 5: Distribution of primary care and secondary care intervals according to emergency admission
Distribution of primary care interval (blue); and secondary care intervals (red) according to emergency admission for the combined colorectal cancer data. The CAP dataset only recorded emergency admission for the 1st subcohort (n=272).

Supplementary figure 6: The risk of being diagnosed with advanced colorectal cancer as a function of the length of the total diagnostic interval
Estimated odds ratios of being diagnosed with advanced (distant or regional) versus localised colorectal cancer as a function of the length of the total diagnostic interval (combined primary and secondary care intervals) analysed for six cohorts and in total (N= 8,907). We excluded patients with unknown tumour stage

and adjusted for age, gender and alarm symptoms at first presentation. The grey dashed curves with 95% confidence limits are fitted on the combined datasets with grey spikes showing the distribution of the care intervals on a squared scale. The grey horizontal lines indicate the chosen reference point of 30 days. Crude estimates are not shown.

Supplementary table 3: Estimated difference in median care intervals

Estimated difference in median care intervals between patients with and without alarm symptoms of cancer; and between patients with and without emergency admission to hospital - displayed for each colorectal cancer cohort dataset and all data combined. **Abbreviations:** Med= Median; IQI= inter-quartile intervals; NA = Not Available. **Method:** We employed quantile regression analyses (Hao et al. 2007) to estimate the difference in median using the 'qcount' procedure (Miranda 2006) on the smoothed quantiles (Machado et al. 2005), as we considered the outcome to be count data (discrete). We adjusted for gender and age (centred at the median age of 70). Confidence intervals were calculated using standard errors (SEs) estimated from 1,000 repetitions of bootstrapping. Point estimates marked in bold are statistically significant at minimum level of $p < 0.05$. * Could not be estimated because all CRCDK patients with emergency admission to hospital were recorded to have zero days of primary care interval. **Emergency admission was only recorded for the 1st subcohort of the CAP dataset (n=272 patients). **References:** Hao L, Naiman DQ. (2007) *Quantile regression*. Thousand Oaks, Calif.: Sage Publications. Miranda A. (2006): QCOUNT: *Stata program to fit quantile regression models for count data*. Machado JAF, Silva JMCS (2005): Quantiles for Counts. *Journal of the American Statistical Association* 100(472):1226-1237.

Supplementary table 4: Logistic regression details for the two care interval models based on combined datasets displayed on Figure 2.

Supplementary table 5: Eight sensitivity analyses testing the robustness of the basic model

We performed eight sensitivity analyses testing the robustness of the basic model presented on Figure 2. Each figure below show estimated odds ratios of being diagnosed with advanced colorectal cancer as a function of the length of the primary care interval (blue) and the secondary care interval (red) analysed for all cohorts combined. The models are adjusted for age, gender, alarm symptoms and cohort. The area around the fitted curves indicates 95% confidence limits. The spikes below the curves show the distribution of the primary care interval (blue) and secondary care interval (red) on a squared scale. The grey horizontal lines indicate the chosen reference point of 30 days. Crude estimates are not shown.

A) We investigate the implications of missing tumour stage in by recoding unknown tumour stage as advanced CRC and hence including 915 more patients.
B) We estimated the odds of distant vs. regional or localised CRC to test a measure which may approximate better to the relative risk.

C) We restricted the analysis to patients with alarm symptoms.

D) We restricted the analysis to patients with alarm no alarm symptoms.

E) We restricted the analysis to patients with emergency admission.

F) We restricted the analysis to patients with no emergency admission.

G) We excluded patients with zero days from presentation to diagnosis.

H) We excluded patients from the ALBERTA dataset which made up 50% of the combined cohort.