

Endoscopy No. 9 / 2016

ENDOS-2015-13583.R1 / Plumb

Original Article

Terminal digit preference biases polyp size measurements at endoscopy, computed tomographic colonography, and histopathology

Andrew A. Plumb<sup>1</sup>, Claire Nickerson<sup>2</sup>, Katherine Wooldrage<sup>3</sup>, Paul Bassett<sup>4</sup>, Stuart A. Taylor<sup>1</sup>, Douglas Altman<sup>5</sup>, Wendy Atkin<sup>3</sup>, Steve Halligan<sup>1</sup>

<sup>1</sup> Centre for Medical Imaging, Division of Medicine, University College London, London, United Kingdom

<sup>2</sup> NHS Cancer Screening Programmes, Fulwood House, Sheffield, United Kingdom

<sup>3</sup> Cancer Screening and Prevention Research Group, Department of Surgery and Cancer, Imperial College London, London, United Kingdom

<sup>4</sup> Research Support Centre, University College London, London, United Kingdom

<sup>5</sup> Centre for Statistics in Medicine, University of Oxford, Oxford, United Kingdom

Short title: Polyp size measurement bias at endoscopy, CTC, and histopathology

Corresponding author

Steve Halligan, MD, PhD, FRCP, FRCR

Centre for Medical Imaging

3rd Floor East

250 Euston Road

London NW1 2PG

United Kingdom

Email: [s.halligan@ucl.ac.uk](mailto:s.halligan@ucl.ac.uk)

**Background and study aims:** Terminal digit preference bias for “pleasing” numbers has been described in many areas of medicine. The aim of this study was to determine whether endoscopists, radiologists, and pathologists exhibit such bias when measuring colorectal polyp diameters.

**Methods:** Colorectal polyp diameters measured at endoscopy, computed tomographic colonography (CTC), and histopathology were collated from a colorectal cancer screening program and two parallel multicenter randomized trials. Smoothing models were fitted to the data to estimate the expected number of polyps at 1-mm increments, assuming no systematic measurement bias. The difference between the expected and observed numbers of polyps was calculated for each terminal digit using statistical modeling. The impact of measurement bias on per-patient detection rates of polyps  $\geq 10$  mm was estimated for each modality.

**Results:** A total of 92 124 individual polyps were measured by endoscopy (91 670 screening and 454 from trials), 2 385 polyps were measured by CTC (1664 screening, 721 trials), and 79 272 were measured by histopathology (78 783 screening, 489 trials). Clustering of polyp diameter measurements at 5-mm intervals was demonstrated for all modalities, both in the screening program and the trials. The statistical models estimated that per-patient detection rates of polyps  $\geq 10$  mm were over-inflated by 2.4% for endoscopy, 3.1% for CTC, and 3.3% for histopathology in the screening program, with similar trends in the randomized trials.

**Conclusion:** Endoscopists, radiologists, and pathologists all exhibit terminal digit preference when measuring colorectal polyps. This will bias trial data, referral rates for further testing, adenoma surveillance regimens, and comparisons between tests.

## **Introduction**

Timely endoscopic polypectomy prevents colorectal cancer; meta-analysis [1] of randomized trials [2–5] confirms that both incidence and mortality are reduced by screening. Following an index colonoscopy, whether performed in a screening or symptomatic setting, patients may be categorized according to their subsequent risk of colorectal neoplasia, which depends on the number and maximal diameter of adenomas measured at the index colonoscopy [6,7]. Any bias in the measurement of polyp diameter will influence risk categorization and hence the requirement for surveillance colonoscopy.

Measuring polyp diameter is also important in other disciplines. Computed tomographic colonography (CTC) is a sensitive technique for the detection of colorectal polyps [8] and cancers [9]. CTC can estimate the size, number, and morphology of polyps, all factors that are related to the risk of advanced histology [10], thereby guiding referral for subsequent colonoscopy. Such referral largely depends on polyp diameter and number [11]. However, phantom [12] and clinical studies [13] have shown that CTC may not assess diameter accurately, and any measurement bias will affect referral rates to colonoscopy. Histopathologists also measure colorectal polyps, most commonly providing a reference measurement of the adenomatous portion of a polypectomy specimen [14], thereby also impacting on surveillance schedules if biased.

As the importance of specific diameter thresholds are well known to endoscopists, radiologists, and pathologists, it is possible that this knowledge might bias estimates of maximal polyp diameter. For example, if 9-mm polyps are up-rated to 10 mm consistently, advanced adenoma rates will be overestimated, as this diameter threshold partly defines an adenoma as “advanced” [15]. Because detection rates of

adenomas  $\geq 10$  mm are often used to compare screening modalities [16,17], if one modality is subject to bias more frequently than the others, an artificial difference in detection rates of adenomas  $\geq 10$  mm will be reported whereas, in reality, this may be due to error. If such bias exists, it will likely be manifest as “terminal digit preference.” This describes over-representation of certain numbers, usually digits ending with 5 or 0, and has been reported in studies of hypertension [18], renal cancers [19], and emergency department waiting times [20].

The aim of the current study was to determine whether terminal digit preference occurred in measurements of maximal polyp diameter by three groups of observers: (1) colonoscopists measuring polyps in vivo; (2) radiologists reporting CTC; and (3) histopathologists measuring polyps ex vivo. Both symptomatic and screening patient populations were examined.

## **Materials and methods**

A waiver to publish nonidentifiable data was obtained from the University College London / University College London Hospitals Joint Research Office.

### *Data sources*

Participants were included from centrally held data from the English National Health Service Bowel Cancer Screening Programme (“BCSP cases”), and two randomized controlled trials of symptomatic patients [21,22] (“RCT cases”). The BCSP is a national screening program of adults aged 60–74 years, using guaiac fecal occult blood test (gFOBT), followed by colonoscopy for those testing positive [23]. CTC is used when colonoscopy is incomplete or judged unsuitable (e.g. due to co-morbidity) as per international consensus recommendations [24,25]. The RCTs were parallel, multicenter, randomized trials of barium enema vs. CTC and colonoscopy vs. CTC for

diagnosis of colorectal cancer and polyps  $\geq 10$  mm in symptomatic patients. The primary outcomes of both trials have been published [21,22].

#### *Endoscopy, CTC, and pathology*

Endoscopists working within the BCSP must have completed >1000 colonoscopies (including >150 in the previous year), passed a written multiple-choice question examination, and been assessed during two peer-observed colonoscopies [26]. Polyp diameter is usually judged in comparison to open biopsy forceps or an endoscopic snare, measured to the nearest millimeter. Sedoanalgesia is administered in approximately 90% of cases, according to endoscopist preference and patient requirements.

The CTC technique within the BCSP requires multidetector-row computed tomography (CT) at a slice thickness of 1–3 mm, dual patient positioning, gas insufflations, and a colonography workstation [27]. Oral contrast fecal tagging and administration of antispasmodics are recommended (and now mandatory).

Pathologists are requested to record the absolute diameter of nonadenomatous polyps and the maximal diameter of the adenomatous portion for adenomas, using a dome magnifier or microscope micrometer scale, to the nearest millimeter. Large lesions that cannot be accommodated are measured using a ruler and the formalin-fixed specimen to the nearest millimeter.

The RCTs were conducted at 21 hospitals and included 217 accredited endoscopists. Polyp diameters were measured according to routine local practice, entailing judgement of size in comparison to biopsy forceps. CTC examinations required full bowel purgation, multidetector-row CT at a minimum collimation of 2.5 mm, dual patient positioning, colonic gas insufflations, and a colonography workstation. CTC

interpretation was by 41 consultant radiologists with experience of CTC, including a 2-day training course. Polyp measurement was discussed during training, and radiologists were asked to measure the long axis of all suspected polyps using electronic calipers, using multiplanar reformatting and a suitable CT window. Maximal polyp diameters for both CTC and colonoscopy were recorded on a case report form, to the nearest millimeter.

#### *Data selection*

Polyp diameters for BCSP cases were extracted from the program database for all screenees undergoing colonoscopy as their first-line test following positive gFOBT from January 2011 to December 2012. For CTC, we selected all patients who underwent CTC as their first-line test to investigate positive gFOBT from April 2006 to March 2013, and who had a polyp detected.

The RCT database was searched by a statistician, who identified and extracted all polyp measurements (at endoscopy, CTC, or histopathology) for all patients included in either trial. If the same polyp was visualized at multiple endoscopic examinations (e.g. diagnostic endoscopy followed by a repeat examination for polypectomy), all measurements were included, on the basis that each individual measurement might be subject to terminal digit preference. Data were collated in Microsoft Excel 2011 for Mac (Microsoft Corp., Redmond, Washington, USA) and analyzed using R version 3.0.1 (R Foundation for Statistical Computing, Vienna, Austria).

#### *Statistical analysis*

Patients may have more than one polyp, and measurements of polyp diameter may be correlated within individuals (i.e. clustered). To account for this, we conducted a cluster bootstrap analysis, generating 2000 replicates by sampling with replacement

from the highest level (patient) and without replacement from the lower level (polyp), as this is preferred to sampling with replacement at all levels [28]. The 95% confidence intervals were calculated using the percentile method. Subsequently, to estimate over-representation of a particular terminal digit, we required assumptions regarding the true distribution of polyp diameters; this is unknown, and nonlinear. Small polyps are more frequent than large polyps, so we expect more polyp diameters with terminal digits of 1 or 2 than 8 or 9. We used two previously published methods of estimating distributions subject to digit preference [29,30]. Camarda et al. [29] developed a composite link model to estimate the probability of misreported data and reassign these to a neighboring category, thereby providing an estimate of the underlying distribution. The Wang et al. method [30] smooths the observed distribution of polyp sizes by using a kernel density estimator. Both methods were implemented in R version 3.0.1 using code provided by each method's lead author (Camarda: <https://sites.google.com/site/carlogiovannicamarda/r-stuff/digit-preference-model>; Wang: bda package [31]). We were unable to apply these methods to each individual bootstrap replicate owing to computational demands; therefore, these analyses were conducted on the original dataset.

Having estimated the underlying distribution of polyp sizes without digit preference, we compared this to the observed number of polyps at 1-mm diameter increments, pooled across each terminal digit (e.g. 2 mm, 12 mm, 22 mm – terminal digit of 2). The difference between the observed number of polyps at each terminal digit and that predicted by the models was taken as an estimate of the number of “re-allocated” polyps (i.e. moved to a different size measurement). This was then converted to a percentage to allow comparison between endoscopy, CTC, and histopathology.



Because of the importance of the 10-mm diameter threshold, we calculated the effect that bias might have on detection rates of polyps  $\geq 10$  mm on a per-patient basis. The analysis was restricted to polyps nearing 10 mm, because digit preference bias at larger sizes rarely alters clinical management. We specifically focused on polyps within 3 mm of the 10-mm threshold (i.e. 7 mm to 13 mm) as we judged that larger rounding errors purely due to digit preference were implausible. For example, if 1000 polyps were recorded as measuring 10 mm, but no polyps measured 7 mm, 8 mm, 9 mm, 11 mm, 12 mm, or 13 mm, we assumed that some of these 1000 polyps would, in reality, have measured less than 10 mm and some would have measured more. We used the Camarda et al. and Wang et al. models to predict the correct distribution of diameters of these 1000 polyps, and compared them to the observed number. Statistical significance testing was not performed because the very large number of polyps included meant that even small deviations from the model were highly significant. We did not attempt to test which modality was most “accurate,” or agreement between the three methods of polyp measurement, as our interest was to document and quantify digit preference bias, rather than polyp measurement accuracy against a reference standard. Furthermore, we were unable to reliably match all polyps between all three modalities, rendering such a comparison impossible. Finally, we conducted an exploratory analysis, treating adenomas and nonadenomas separately, on the basis that the former are of greater biological significance and therefore any bias carries greater clinical importance.

## Results

### *Screening program cases*

#### Endoscopic measurements

Polyp diameters were available for 91 670 polyps in 36 850 individuals, with a mean diameter of 6.6 mm (range 1–200 mm). In total, 67 453 polyps (73.6%) were sessile, 21 890 (23.9%) were pedunculated, and 2327 (2.5%) were flat. Diameter distribution is shown in Fig. 1a, with clusters evident at 5-mm intervals. Positive deviation from the fitted model of polyp diameters was substantially greater for the terminal digit of 0 (i.e. multiples of 10 mm) than for all other sizes (Fig. 1b [Camarda model] and Fig. 1c [Wang model]). A similar, but less marked effect was present for polyp diameters with a terminal digit of 5 mm. Approximately 10% of all re-allocated polyp diameters occurred for a terminal digit of zero (estimate by Camarda method 10.1%; Wang method 9.5%), and approximately 7% of re-allocation occurred for a terminal digit of 5 (Camarda estimate 6.9%, Wang method 7.3%). In contrast, there was a relative deficiency of polyps measured as having a diameter with a terminal digit of 1 (Camarda estimate –6.5%, Wang estimate –7.9%) or of 9 (Camarda estimate –4.4%, Wang estimate –5.7%). There was a tendency for the effect to be of greater magnitude for adenomatous vs. nonadenomatous polyps (Fig. e2a,b, available online).

#### CTC measurements

Measurements were available for 1664 polyps in 892 individuals. Mean diameter was 9.5 mm (range 1–91 mm). In total, 1108 polyps (66.6%) were sessile, 494 (29.7%) were pedunculated, and 62 (3.7%) were flat. Diameter distribution is shown in Fig. 3a, with clusters at 5-mm intervals, similarly to the endoscopy data. Deviation from the fitted model of polyp diameters was substantially greater for the terminal digit 0

(i.e. multiples of 10 mm) than for all other sizes. There was a small excess of polyps with a terminal digit measurement of 5, and a relative deficiency of polyps with a maximum diameter ending in either 1 or 9 (Fig. 3b,c). Approximately 8% of all re-allocated polyp diameters occurred for a terminal digit of 0 (Camarda estimate 8.0%, Wang estimate 8.3%), and approximately 5% of re-allocation occurred for a terminal digit of 5 (Camarda estimate 5.1%, Wang estimate 5.3%).

### Histopathological measurements

Diameters were available for 78 783 polyps in 33 891 individuals (fewer than endoscopy because of incomplete data entry and polyp nonretrieval). Mean polyp diameter was 6.4 mm (range 1–190 mm). Once again, diameters clustered at 5-mm intervals (Fig. 4a), with an excess of polyps with diameters ending in 0 (approximately 4.0% of re-allocated polyps by both Camarda and Wang methods) or 5 (Camarda estimate 2.9%, Wang estimate 3.5%), and relative under-representation of polyps with a diameter measurement ending in 1 or 9 mm (Fig. 4b,c). As for endoscopy, the effect was more marked for adenomatous polyps (Fig. e5a,b, available online).

### Estimated impact on per-patient detection rates of polyps $\geq 10$ mm

Of the 36 850 screenees with polyps diagnosed endoscopically, 15 952 (43.3%) had at least one polyp measured as  $\geq 10$  mm. Assuming the fitted models represent the true distribution of polyps, the Camarda method suggests that only 15 223 (41.3%) patients would truly have polyps  $\geq 10$  mm. The Wang method estimated this figure to be 14 924 patients (40.5%). The average difference between the observed number of polyps  $\geq 10$  mm and the number derived by the two methods was  $-2.4\%$  (i.e. we estimate that 2.4% of all patients undergoing endoscopy were erroneously diagnosed

with polyps  $\geq 10$  mm as a result of digit preference). For CTC, of the 892 screenees with polyps diagnosed, 483 (54.1%) had one or more polyps  $\geq 10$  mm. The Camarda model suggests that this figure would be 463 (51.9%) without digit preference bias, and the Wang method estimates the true figure to be 447 (50.1%); the average difference was -3.1%. Finally, for histological measurement, 14 282 of 33 891 (42.1%) screenees had polyps  $\geq 10$  mm. This figure would fall to 13 653 (40.3%) under the assumption of a smooth distribution of polyp diameters using the Camarda method, and to 12 660 (37.4%) using that of Wang – a mean difference of -3.3% (Table 1).

#### *RCT cases*

##### Endoscopic measurements

Data were available for 454 polyps, of which 329 had diameters measured. Mean diameter was 10.2 mm (range 1–115mm). Similarly to the screening data, clustering at 5-mm intervals was evident (Fig. 6a), particularly for terminal digits of 0 (Fig. 6b), with approximately 24% (both Camarda and Wang methods) of re-allocation occurring for polyps with a maximum diameter of a multiple of 10 mm.

##### CTC measurements

A total of 721 polyps were documented, of which 680 had diameters recorded: 41 polyps had either no recorded diameter or a size category only ( $<5$  mm, 5–9 mm, or  $\geq 10$  mm). Mean diameter was 13.6 mm (range 2–110 mm), and again showed clustering at 5-mm intervals (Fig. 7a). Approximately 18.3% (Camarda) to 22.4% (Wang) of re-allocation occurred for polyps with a maximal diameter of a multiple of 10 mm (Fig. 7b,c).

## Histopathological measurements

A total of 489 polyps were documented, 244 of which had diameters recorded (mean 18.8 mm, range 1–230 mm). As for other modalities, the number of polyps at multiples of 5 mm was greater than expected (Fig. 8a). Re-allocated polyps were considerably more common for terminal digits of 5 (Camarda estimate 17.0%, Wang estimate 19.7%) or 0 (Camarda estimate 19.4%, Wang estimate 20.9%) (Fig. 8b,c).

## Estimated impact on per-patient detection rates of polyps $\geq 10$ mm

Endoscopic data were available for 169 patients, of whom 81 (47.9%) had polyps  $\geq 10$  mm. No patient had either a 9 mm or an 11 mm polyp recorded as their largest lesion; 24 patients had polyps measuring precisely 10 mm. The two models predicted that 73 patients (43.2% Camarda method) or 72 patients (42.6% Wang method) would have polyps  $\geq 10$  mm, corresponding to a mean difference of -5.0% from the observed percentage. A total of 333 patients had their largest polyp diameter recorded to 1-mm precision at CTC, of whom 192 (57.7%) had their largest polyp measured as  $\geq 10$  mm, whereas the Camarda model predicted that this figure would be 194 (58.3%) without digit preference. The Wang model gave a very similar estimate of 191 patients (57.4%); the mean difference was 0.26%. For pathology, 162 patients had pathological diameters recorded; 102 (63.0%) harbored polyps of  $\geq 10$  mm. This figure dropped to 98 patients (60.5%) by the Camarda method and to 96 patients (59.3%) when using the Wang estimate (Table 1) – an average difference of -3.1% from the observed value.

## Discussion

This study demonstrated that measurement of polyps by endoscopists, radiologists, and histopathologists, in both screening and symptomatic practice, is subject to

terminal digit preference bias. The effect is sizeable and apparent across all polyp diameters. Our data suggest that terminal digit preference may increase the proportion of patients with polyps  $\geq 10$  mm measured at CTC by up to 3%, increasing the referral rate for colonoscopy inappropriately. Furthermore, the same phenomenon occurs at both colonoscopy and histopathology, with an estimated 2%–5% of patients erroneously diagnosed with polyps  $\geq 10$  mm, and therefore potentially subjected to unnecessarily intensive colonoscopic surveillance.

Terminal digit preference has been described in many areas of medicine and epidemiology. “Age heaping” in national census data has been recognized for decades [32], and demonstrates clustering at digits ending in 0 or 5 in the West and multiples of 12 in China (corresponding to preferred years of the Chinese animal calendar) [33]. Medical practitioners seem equally affected by this bias, with studies from disparate areas such as blood pressure monitoring [18], neonatal birthweight [34], tumor size measurement [35], and emergency department waiting times [20] all reporting the phenomenon.

The effect remains when small differences in measurement make a substantial difference to patient outcome, and may partly arise because of this clinical significance. For example, Hayes [19] described digit preference in the measurement of renal cancers, for which maximal diameter determines tumor stage (and hence treatment). Similarly, while biases in polyp diameter may extend to only a millimeter or so, they can affect subsequent management (i.e. referral for colonoscopy after CTC [11], or the timing and frequency of surveillance colonoscopy [6,7]).

Whether polyps are most accurately measured by colonoscopy or CTC is not clear [36]. An in vivo study comparing CTC and several colonoscopic measurement techniques found that 2D CTC was the most reproducible between observers,

although tended to record larger diameters than colonoscopy [13]. The technique used at endoscopy influences size measurement considerably, with comparison against open biopsy forceps being the least accurate in one study [37]. Nor is pathology a universal gold standard; even the measurement of polyp phantoms ex vivo (to mimic a polypectomy specimen in the pathology laboratory) is subject to considerable measurement error and poor reproducibility [38].

Irrespective of which technique is most accurate, current criteria for both referral for colonoscopy after CTC [11,24,25] and surveillance colonoscopy after polypectomy [6,7] are based on studies reporting the risk of histologically advanced neoplasia within polyps of various sizes (measured endoscopically and pathologically). As we have demonstrated that digit preference bias occurs for both colonoscopists and histopathologists, it is highly likely that this historical literature also incorporates digit preference. This might have the effect of erroneously placing many 9-mm polyps into the  $\geq 10$ -mm category, which would (counter-intuitively) lower the proportion of lesions with advanced histology in both groups. As 9-mm polyps are, on average, the highest risk of the subcentimeter polyp group, moving them into the  $\geq 10$ -mm category will tend to reduce the overall proportion of advanced histology in the remaining pool of subcentimeter polyps. Furthermore, as these same 9-mm polyps will, on average, be less likely to harbor advanced histology than most lesions  $\geq 10$  mm, their addition to the large polyp group will also reduce the proportion of advanced neoplasia here. In contrast, 11-mm polyps being “rounded down” to 10 mm would have no effect, as they would remain within the same size category. This so-called “Will Rogers phenomenon” (the apparently counter-intuitive reduction in risk for both groups) has been well-described for cancer stage migration [39]. The degree to which this affects

rates of histologically advanced neoplasia in the historical literature is not clear, but our data suggest that such an effect is possible.

The current study has some limitations. Polyp diameters recorded onto a central database or trial case report forms may not reflect the measurements given on the clinical endoscopy, radiology, or pathology report. It is possible that clinicians are more precise in their routine practice than when recording data for central analysis. Additionally, the technique of polyp measurement was not stipulated in detail, either for endoscopy or CTC, although arguably this reflects real-world practice most closely. Furthermore, the behavior of practitioners in both the BCSP and the RCTs, although broadly representative, may not generalize more widely. We were unable to retrieve data regarding individual practitioners, meaning we were unable to explore the possibility that the effect is underpinned by a subgroup of individuals who are particularly prone to such bias. We measured the effect of digit preference on the detection rate of polyps 10 mm, but this does not directly equate to a need for surveillance colonoscopy – this may be required in any case (e.g. because of multiplicity of subcentimeter adenomas [6,7], a situation that applied to 24% of patients in our cohort). Finally, statistical methods for estimating terminal digit preference are relatively under-developed, particularly when the true underlying distribution of the variable of interest is unknown. Accordingly, the techniques we applied here may not accurately model the genuine distribution of polyp diameters, potentially under- or over-estimating the magnitude of digit preference. The degree of uncertainty (i.e. the degree of under- or over-estimation) is unknown, due to the relative immaturity of the relevant statistical literature. We applied the methods to datasets ranging from 329 to >90 000 polyps and the methods may not be equally precise across these extremes. They also do not take into account the statistical



clustering of multiple polyps within a single individual. This does not alter our primary conclusion that the digit preference bias phenomenon undoubtedly occurs, which is clear from simple inspection of the histograms of polyp diameter.

In summary, we have shown that terminal digit preference occurs when endoscopists, radiologists, and pathologists are asked to measure the maximal diameter of colorectal polyps, for both symptomatic and screening patients. The magnitude of the effect was similar between the three specialities in terms of diagnostic yield at the  $\geq 10$ -mm threshold. This bias may influence the number of patients entering polyp surveillance regimens and being referred for endoscopic excision, and will also affect comparisons of detection rates between tests in research studies. All relevant practitioners should be aware of this phenomenon when estimating the maximal diameter of colorectal polyps in order to reduce biases introduced by mis-classification.

### **Acknowledgment**

This article presents independent research funded in part by the UK National Institute for Health Research (NIHR) under its Programme Grants for Applied Research funding scheme (RP-PG-0407-10338). A proportion of this work was undertaken at University College London, which received a proportion of funding from the NIHR Biomedical Research Centre funding Scheme. Funding was also received from NIHR Health Technology Assessment Programme (HTA 02/02/01), Royal College of Radiologists Kodak Fund, and the Wellcome Trust Institutional Strategic Support Fund.

The authors acknowledge the assistance of Dr. Manuel Rodriguez-Justo, Consultant Pathologist, for providing details of histopathological polyp measurement in the BCSP.

*Competing interests:* None.

## **References**

- 1 Brenner H, Stock C, Hoffmeister M. Effect of screening sigmoidoscopy and screening colonoscopy on colorectal cancer incidence and mortality: systematic review and meta-analysis of randomised controlled trials and observational studies. *BMJ* 2014; 348: g2467
- 2 Atkin WS, Edwards R, Kralj-Hans I et al. Once-only flexible sigmoidoscopy screening in prevention of colorectal cancer: a multicentre randomised controlled trial. *Lancet* 2010; 375: 1624–1633
- 3 Segnan N, Armaroli P, Bonelli L et al. Once-only sigmoidoscopy in colorectal cancer screening: follow-up findings of the Italian Randomized Controlled Trial-SCORE. *J Natl Cancer Inst* 2011; 103: 1310–1322
- 4 Hoff G, Grotmol T, Skovlund E et al. Risk of colorectal cancer seven years after flexible sigmoidoscopy screening: randomised controlled trial. *BMJ* 2009; 338: b1846
- 5 Schoen RE, Pinsky PF, Weissfeld JL et al. Colorectal-cancer incidence and mortality with screening flexible sigmoidoscopy. *N Engl J Med* 2012; 366: 2345–2357
- 6 Cairns SR, Scholefield JH, Steele RJ et al. Guidelines for colorectal cancer screening and surveillance in moderate and high risk groups (update from 2002). *Gut* 2010; 59: 666–689

- 7 Hassan C, Quintero E, Dumonceau JM et al. Post-polypectomy colonoscopy surveillance: European Society of Gastrointestinal Endoscopy (ESGE) Guideline. *Endoscopy* 2013; 45: 842–851
- 8 de Haan MC, van Gelder RE, Graser A et al. Diagnostic value of CT-colonography as compared to colonoscopy in an asymptomatic screening population: a meta-analysis. *Eur Radiol* 2011; 21: 1747–1763
- 9 Pickhardt PJ, Hassan C, Halligan S et al. Colorectal cancer: CT colonography and colonoscopy for detection – systematic review and meta-analysis. *Radiology* 2011; 259: 393–405
- 10 Kolligs FT, Crispin A, Graser A et al. Risk factors for advanced neoplasia within subcentimetric polyps: implications for diagnostic imaging. *Gut* 2013; 62: 863–870
- 11 Zalis ME, Barish MA, Choi JR et al. CT colonography reporting and data system: a consensus proposal. *Radiology* 2005; 236: 3–9
- 12 Punwani S, Halligan S, Irving P et al. Measurement of colonic polyps by radiologists and endoscopists: who is most accurate? *Eur Radiol* 2008; 18: 874–881
- 13 de Vries AH, Bipat S, Dekker E et al. Polyp measurement based on CT colonography and colonoscopy: variability and systematic differences. *Eur Radiol* 2010; 20: 1404–1413
- 14 Quirke P, Risio M, Lambert R et al. European guidelines for quality assurance in colorectal cancer screening and diagnosis. First Edition – Quality assurance in

pathology in colorectal cancer screening and diagnosis. *Endoscopy* 2012; 44 Suppl 3: SE116–130

- 15 Winawer SJ, Zauber AG. The advanced adenoma as the primary target of screening. *Gastrointest Endosc Clin N Am* 2002; 12: 1–9
- 16 Stoop EM, de Haan MC, de Wijkerslooth TR et al. Participation and yield of colonoscopy versus non-cathartic CT colonography in population-based screening for colorectal cancer: a randomised controlled trial. *Lancet Oncol* 2012; 13: 55–64
- 17 Kim DH, Pickhardt PJ, Leung WK et al. CT colonography versus colonoscopy for the detection of advanced neoplasia. *N Engl J Med* 2007; 357: 1403–1412
- 18 Nietert PJ, Wessell AM, Feifer C et al. Effect of terminal digit preference on blood pressure measurement and treatment in primary care. *Am J Hypertens* 2006; 19: 147–152
- 19 Hayes SJ. Terminal digit preference occurs in pathology reporting irrespective of patient management implication. *J Clin Pathol* 2008; 61: 1071–1072
- 20 Locker TE, Mason SM. Digit preference bias in the recording of emergency department times. *Eur J Emerg Med* 2006; 13: 99–101
- 21 Atkin W, Dadswell E, Wooldrage K et al. Computed tomographic colonography versus colonoscopy for investigation of patients with symptoms suggestive of colorectal cancer (SIGGAR): a multicentre randomised trial. *Lancet* 2013; 381: 1194–1202

- 22 Halligan S, Wooldrage K, Dadswell E et al. Computed tomographic colonography versus barium enema for diagnosis of colorectal cancer or large polyps in symptomatic patients (SIGGAR): a multicentre randomised trial. *Lancet* 2013; 381: 1185–1193
- 23 Logan RFA, Patnick J, Nickerson C et al. Outcomes of the Bowel Cancer Screening Programme (BCSP) in England after the first 1 million tests. *Gut* 2012; 61: 1439–1446
- 24 Spada C, Stoker J, Alarcon O et al. Clinical indications for computed tomographic colonography: European Society of Gastrointestinal Endoscopy (ESGE) and European Society of Gastrointestinal and Abdominal Radiology (ESGAR) Guideline. *Endoscopy* 2014; 46: 897–915
- 25 Spada C, Stoker J, Alarcon O et al. Clinical indications for computed tomographic colonography: European Society of Gastrointestinal Endoscopy (ESGE) and European Society of Gastrointestinal and Abdominal Radiology (ESGAR) Guideline. *Eur Radiol* 2015; 25: 331–345
- 26 Joint Advisory Group on GI Endoscopy. Accreditation of screening colonoscopists. BCSP Guidelines. Royal College of Physicians; 2013. Available from:  
[http://www.saas.nhs.uk/documents/Accreditation%20of%20screening%20colonoscopists%20guidelines%20final\\_June\\_2013.pdf](http://www.saas.nhs.uk/documents/Accreditation%20of%20screening%20colonoscopists%20guidelines%20final_June_2013.pdf)
- 27 Burling D, Patnick J. Guidelines for the use of imaging in the NHS Bowel Cancer Screening Programme. In: *NHS Cancer Screening Programmes*; 2010

- 28 Ren S, Lai H, Tong W et al. Nonparametric bootstrapping for hierarchical data. *J Applied Stat* 2010; 37: 1487–1498
- 29 Camarda CG, Eilers PHC, Gampe J. Modelling general patterns of digit preference. *Stat Modelling* 2008; 8: 385–401
- 30 Wang B, Wertenlecker W. Density estimation for data with rounding errors. *Comput Stat Data Anal* 2013; 65: 4–12
- 31 Wang B. bda: Density estimation for binned/weighted data. R package version 3.2.0-3. 2014. Available from: <https://cran.r-project.org/web/packages/bda/index.html>
- 32 Myers RJ. Errors and bias in the reporting of ages in census data. *Transactions of the Actuarial Society of America* 1940; 41: 395–415
- 33 Jowett AJ, Li Y. Age-heaping: contrasting patterns from China. *GeoJournal* 1992; 28: 427–442
- 34 Emmerson AJ, Roberts SA. Rounding of birth weights in a neonatal intensive care unit over 20 years: an analysis of a large cohort study. *BMJ Open* 2013; 3: e003650
- 35 Hayes SJ. Does terminal digit preference occur in pathology? *J Clin Pathol* 2008; 61: 975–976
- 36 Summers RM. Polyp size measurement at CT colonography: what do we know and what do we need to know? *Radiology* 2010; 255: 707–720

- 37 Gopalswamy N, Shenoy VN, Choudhry U et al. Is in vivo measurement of size of polyps during colonoscopy accurate? *Gastrointest Endosc* 1997; 46: 497–502
- 38 Rubio CA, Grimelius L, Lindholm J et al. Reliability of the reported size of removed colorectal polyps. *Anticancer Res* 2006; 26: 4895–4899
- 39 Feinstein AR, Sosin DM, Wells CK. The Will Rogers phenomenon. Stage migration and new diagnostic techniques as a source of misleading statistics for survival in cancer. *N Engl J Med* 1985; 312: 1604–1608

**Fig. 1** Endoscopist measurement of polyps in the screening program (BCSP).

**a** Histogram of observed endoscopically measured polyp diameters in the BCSP (with 95% confidence intervals). **b,c** Number of re-allocated polyps grouped by terminal digit, expressed as a percentage of all re-allocated polyps, estimated by the Camarda method (**b**) and the Wang method (**c**).

**Fig. e2** Endoscopist measurement of adenomas and nonadenomas in the screening program, showing clusters for terminal digits of 0 and 5.

**a** Adenomas. **b** Nonadenomas. Bars represent 95% confidence intervals.

**Fig. 3** Computed tomographic colonography (CTC) measurement of polyps in the BCSP. **a** Histogram of observed CTC-measured polyp diameters in the BCSP (with 95% confidence intervals). **b,c** Number of re-allocated polyps grouped by terminal digit, expressed as a percentage of all re-allocated polyps, estimated by the Camarda method (**b**) and the Wang method (**c**).

**Fig. 4** Histopathological measurement of polyps in the screening program (ScrP). **a** Histogram of observed histologically measured polyp diameters in the ScrP (with 95% confidence intervals). **b,c** Number of re-allocated polyps grouped by terminal digit, expressed as a percentage of all re-allocated polyps, estimated by the Camarda method (**b**) and the Wang method (**c**).



**Fig. e5** Histopathological measurement of adenomas and nonadenomas in the screening program, showing clusters for terminal digits of 0 and 5.

**a** Adenomas. **b** Nonadenomas. Bars represent 95% confidence intervals.

**Fig. 6** Endoscopic measurements of polyps in randomized controlled trials (RCTs). **a** Histogram of observed endoscopically measured polyp diameters in the RCTs (with 95% confidence intervals). **b,c** Number of re-allocated polyps grouped by terminal digit, expressed as a percentage of all re-allocated polyps, estimated by the Camarda method (**b**) and the Wang method (**c**).

**Fig. 7** Computed tomographic colonography (CTC) measurement of polyps in randomized controlled trials (RCTs). **a** Histogram of observed CTC-measured polyp diameters in RCTs (with 95% confidence intervals). **b,c** Number of re-allocated polyps grouped by terminal digit, expressed as a percentage of all re-allocated polyps, estimated by the Camarda method (**b**) and the Wang method (**c**).

**Fig. 8** Histopathological measurement of polyps in randomized controlled trials (RCTs). **a** Histogram of observed histologically measured polyp diameters in RCTs (with 95% confidence intervals). **b,c** Number of re-allocated polyps grouped by terminal digit, expressed as a percentage of all re-allocated polyps, estimated by the Camarda method (**b**) and the Wang method (**c**).



**Table 1** Observed and estimated (i.e. corrected for digit preference) number (percentage) of patients with a polyp  $\geq 10$  mm as their largest measured lesion, split by diagnostic modality, for both screening (BCSP) and trial (RCT) cases. The mean difference is the difference between the number of patients who were recorded to have at least one polyp  $\geq 10$  mm and the arithmetic mean of the estimate by the two smoothing methods.

	BCSP cases			RCT cases		
	Endoscopy (n = 36 850)	CTC (n = 892)	Histopathology (n = 33 891)	Endoscopy (n = 169)	CTC (n = 333)	Histopathology (n = 162)
Observed, n (%)	15 952 (43.3)	483 (54.1)	14 282 (42.1)	81 (47.9)	192 (57.7)	102 (63.0)
Camarda estimate, n (%)	15 223 (41.3)	463 (51.9)	13 653 (40.3)	73 (43.2)	194 (58.3)	98 (60.5)
Wang estimate, n (%)	14 924 (40.5)	447 (50.1)	12 660 (37.4)	72 (42.6)	191 (57.4)	96 (59.3)
Mean difference, n (%)	-878.5 (-2.4)	-28 (-3.1)	-1125.5 (-3.3)	-8.5 (-5.0)	0.5 (0.26)	-5 (-3.1)

CTC, computed tomographic colonography.