European Radiology (2019) 29:5264–5271
https://doi.org/10.1007/s00330-019-06141-8

COMPUTED TOMOGRAPHY

CT diagnostic reference levels: are they appropriately computed?

Thibault Vanaudenhove[1] · Alain Van Muylem[2] · Nigel Howarth[3] · Pierre Alain Gevenois[4] · Denis Tack[5]

Received: 26 October 2018 / Revised: 18 February 2019 / Accepted: 8 March 2019 / Published online: 8 April 2019

Abstract

Objectives To estimate the variability of CT diagnostic reference levels (DRLs) according to the methods used for computing collected data.

Methods Dose-length products (DLP) were collected by our national nuclear control agency from the 250 devices installed in 140 medical centers in the country. In 2015, the number of head, thorax, abdomen, and lumbar spine examinations collected in these centers ranged from approximately 20,000 to 42,000. The impact on DRLs of the number of devices considered, as well as the differences in descriptive statistics (mean vs. median DLP) or methods of pooling DLP data (all devices vs. all patients), was investigated. Variability in DRLs was investigated using a bootstrapping method as a function of the numbers of devices and examinations per device.

Results As expected, DRLs derived from means were higher than those from medians, with substantial differences between device- and patient-related DRLs. Depending on the numbers of devices and DLP data per device, the variability ranged from 10 to 40% but was stabilized at a level of 10–20% if the number of devices was higher than 50 to 60, regardless of the number of DLP data per device.

Conclusion Number of devices and of DLP data per device, descriptive statistics, and pooling data influence DRLs. As differences in methods of computing survey data can artificially influence DRLs, harmonization among national authorities should be recommended.

Key Points

- *Due to CT dose variability, that of DRLs is at least of 10%.*
- *DRLs derived from medians are lower than from means and differ from those obtained by pooling all patient data.*
- *Fifty to 60 devices should be sufficient for estimating national DRLs, regardless of the number of data collected per device.*

Keywords Radiation protection · Tomography · Surveys and questionnaires

✉ Denis Tack
denis.tack@skynet.be

[1] Federal Agency for Nuclear Control (FANC), Rue Ravenstein 36, 1000 Brussels, Belgium

[2] Department of Chest Medicine, Hôpital Erasme, Route de Lennik 808, 1070 Brussels, Belgium

[3] Department of Radiology, Clinique des Grangettes, 7 Chemin des Grangettes, 1224 Chêne-Bougeries, Switzerland

[4] Department of Radiology, Hôpital Erasme, Route de Lennik 808, 1070 Brussels, Belgium

[5] Department of Radiology, Epicura, Clinique Louis Caty, 136 Rue Louis Caty, 7331 Baudour, Belgium

Abbreviations

D-P75	Device-related diagnostic reference level
DLP	Dose-length product
DRL	Diagnostic reference level
EU	European Union
P-P75	Patient-related diagnostic reference level
P75	75th percentile of dose distribution

Introduction

Establishing diagnostic reference levels (DRLs) on radiation dose delivered by computed tomography (CT) is currently

mandatory in the European Union (EU) member states [1, 2]. The 75th percentiles (P75) of dose distributions are considered as the upper limit of good medical practice and can be used to define DRLs [1–9]. The EU recommends a method based on that used in the UK for establishing these distributions but its application in each EU state is at the discretion of national authorities [8, 9]. This method is based on collecting dose values in ten consecutive so-called standard weight patients (1.70 m and 70 kg) for any given examination. This method has been adapted in the EU states in four different manners. First, as recruiting "standard weight patients" can be difficult, some national authorities disregard the weight and size criteria but request sample sizes larger than ten and up to 40 consecutive patients [6, 10–15]. Second, the method recommended originally considers means of dose values and subsequently device-related DRLs [8–11] but some national authorities pool the patient values—with subsequent patient-related DRLs—in order to overcome huge variabilities between samples and to estimate the global radiation dose delivered to the population [12, 16]. Third, whereas the method recommended originally suggests collecting data from all devices installed in the country, some national authorities accept a voluntary contribution from each center, limiting the collection to 20 to 30% of installed devices [6–12]. Last, the method recommended originally considers means of dose values delivered by each device. As these values are not normally distributed, and as recently suggested by the International Commission on Radiological Protection, medians could be more appropriate than means [3].

Comparisons of DRLs between countries using different methods are thus based on the hypothesis that national adaptations do not influence DRL values [17, 18]. The purpose of this study was to verify this hypothesis by comparing P75 of dose values delivered during four frequently performed CT examinations and collected by methods applied in various countries in the EU (i.e., means vs. medians, patient-related DRLs vs. device-related DRLs, and data collection from all devices vs. from a sample of devices).

Materials and methods

According to the EU legislation (i.e., the Directive 95/46/EC regarding the protection of data of individuals), a purely observational study with complete anonymization of the data at the source, which removes any possibility of identifying the individual patients, is not subject to ethical review [19]. We analyzed CT dose indicators of head, thorax, abdomen, and lumbar spine examinations, anonymized at the source, and collected by the agency in charge of the nuclear control in Belgium (FANC) during the year 2015. For each given examination, the collected dose indicators were the volume computed tomography dose index ($CTDI_{vol}$) and the dose-length product (DLP), along with patient's gender and age.

About 250 CT devices are installed in our country in 180 medical centers. These centers are obliged by law to complete yearly surveys on CT dose and to provide anonymized dose indicator values for given standard examinations delivered to a minimum of 20 consecutive patients or, if less than 20, for all examinations performed within a 3-month period.

As the whole body or effective dose delivered by each examination is the most important metric for regulatory authorities, we focused this study on DLP. For each examination, we considered the sum of DLPs of all acquisitions including all phases of multiphasic examinations, as well as stationary acquisitions for bolus-tracking techniques.

In this article, the expression "device-related DRL" refers to P75 of mean or median DLP distributions among all devices for a given examination. This percentile is abbreviated D-P75. Similarly, the expression "patient-related DRL" refers to P75 of the DLP distribution among all patients undergoing a given examination in the country. This percentile is abbreviated P-P75.

Statistical analyses

For a given examination, a database of M devices named and numbered CT_i (with i ranging from 1 to M) was constituted. Each device provided a corresponding N_i DLP values.

Combined influence of number of devices and sample size per device on D-P75

In this procedure, we considered as eligible all devices with $N_i \geq 100$.

Step 1 For a given eligible device CT_i, a random DLP sample of size n_j was drawn with replacement from its DLP distribution. From this sample, a median (or mean) μ_{ij} was computed. By repeating this procedure 2000 times (yielding $\mu_{ij}(1)$, $\mu_{ij}(2)$, …,$\mu_{ij}(2000)$), the sampling distribution of μ_{ij} was derived and stored. This procedure is illustrated in panel A of Fig. 1. The procedure was executed for each eligible device and for four sample sizes ($j = 1...4$, $n1 = 10$, $n2 = 20$, $n3 = 50$, $n4 = 100$). At this stage, M × 4 median (or mean) sampling distributions MSDij were established for the M eligible devices ($i = 1...M$) and the four considered sample sizes ($j = 1...4$).

Step 2 Out of the M eligible devices, mk devices were drawn at random with replacement. For each of the m_k-drawn devices and a given sample size n_j, one μ_{ij} value was drawn at random from MSDij (i being the number of the considered device). The 75th percentile of this sample of m_k values was computed giving $D\text{-}P75_{jk}$ (for a sample size nj and a number of devices m_k). This procedure is illustrated in panel B of Fig. 1. We considered four numbers of devices ($k = 1...4$, $m1 = 5$, $m2 = 10$, $m3 = 25$,

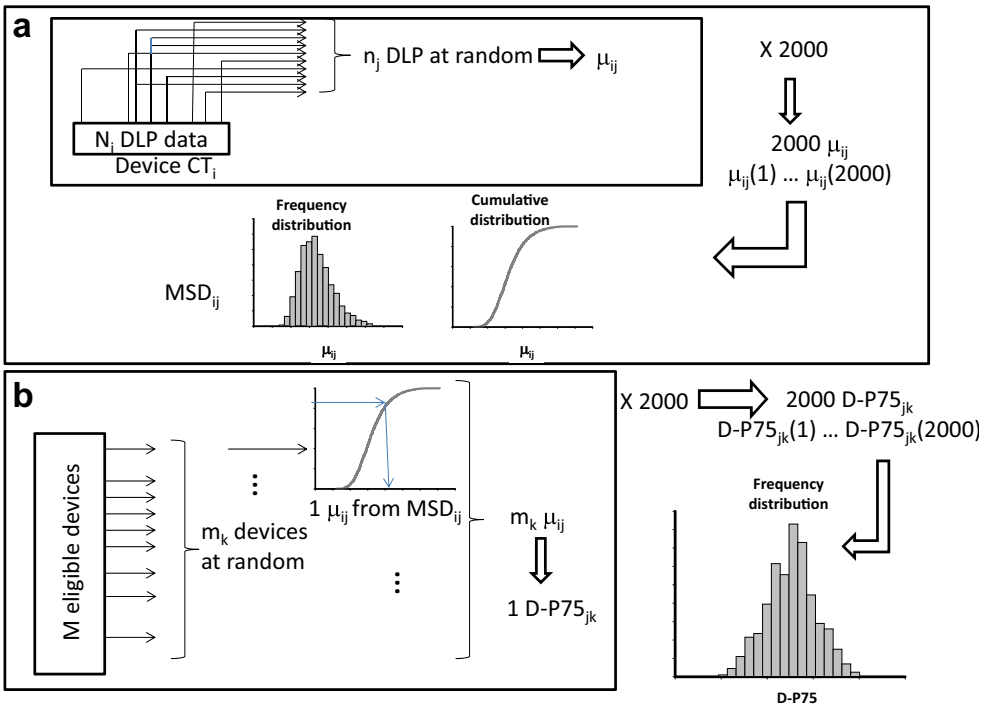

Fig. 1 Statistical method. Panel A illustrates the first step of the statistical analysis and panel B illustrates the second one

$m4 = 50$). By repeating this procedure 2000 times (yielding D-$P75_{jk}(1)$, D-$P75_{jk}(2)$,..., D-$P75_{jk}(2000)$), the sampling distribution of D-$P75_{jk}$ was derived. From this distribution, the median and the 95% confidence interval (percentile 97.5–percentile 2.5) of D-P75 were computed.

Influence of the parameter used (median, mean, or all DLP values)

This procedure compared the P75 distributions, obtained in three different ways of using all available DRLs from a given set of m devices: (1) the median of the available DRL (m medians); (2) the mean of the available DRLs (m means); or (3) all DRLs of the *m* given devices pooled together. In this procedure, we took into account all CT scanners with $N_i \geq 20$. In cases 1 and 2, we considered the 75th percentile of m medians and m means, respectively (D-P75medians and D-P75means). In case 3, we considered the 75th percentile of all DRLs from the *m* devices (P-P75).

Cases 1 and 2

Step 1 This step was analogous to step 1 above, except that for a given eligible device CTi, the random DLP sample drawn with replacement had the size N_i (considering all DLP values of this device) (bootstrap procedure). At this stage, M median and M mean sampling distributions were established for the eligible devices (MSD_i, $i = 1...M$).

Step 2 This step was analogous to step 2 above with m devices drawn at random with replacement from the M eligible devices, each providing one median and one mean value randomly drawn from MSDi. The 75th percentile of this sample of m values was computed, giving D-P75. We considered four values of m (5, 10, 25, and 50). By repeating this procedure 2000 times (yielding D-P75(1), D-P75(2),..., D-P75(2000)), the sampling distribution of D-P75 was derived for each value of m.

Case 3

The first step to determine the sampling distribution may be skipped in this case as each device was represented by all available DRLs. All DRLs of m devices drawn at random with replacement from the M eligible devices were pooled and the 75th percentile of this sample was computed, giving P-P75. We considered four values of m: 5, 10, 25, and 50. By repeating this procedure 2000 times (yielding P-P75(1), P-P75(2),…, P-P75(2000)), the sampling distribution of P-P75 was derived for each value of m.

In the three cases, the median and the 95% confidence interval (percentile 97.5–percentile 2.5) were computed from D-P75 (from medians and means) and P-P75 sampling distributions.

Results

The numbers of CT devices and patients' data collected in 2015 are displayed in Table 1. The D-$P75_{means}$, D-

Table 1 Number of CT devices and DLP data per device

Body region		Head	Thorax	Abdomen	Lumbar spine
Number of patients at national level		27,232	19,683	42,350	22,822
Number of devices at national level		219	215	245	216
Number of devices with number of DLP data	> 5	217	214	244	216
	≥ 25	95	85	116	85
	≥ 40	80	73	96	74
	≥ 100	62	55	80	58
	≥ 200	45	32	59	39
	≥ 500	11	5	29	8
	≥ 1000	3	0	7	2

$P75_{medians}$, and P-P75 are illustrated in Table 2 for the four examinations selected.

The influence of both the number of CT devices included in the survey and of the number of DLP data per device on DRLs is shown in Fig. 2, by the variability of the D-P75 of DLP medians or means, for the head, abdomen, thorax, and lumbar spine examinations. From these figures, the variability of DRL's was stable at around 20%, providing the number of included devices exceeded 50 to 60.

The influence of the number of CT devices—considering 10, 50, or 200 devices—included in the survey on DRLs is shown in Fig. 3 for means, medians, and for all data pooled, for the brain, abdomen, thorax, and lumbar spine examinations.

Discussion

This study shows that (1) differences are observed between device-related DRLs (D-P75) (either based on means or medians) and patient-related DRLs (P-P75), (2) even with large samples of devices, intrinsic variability of D-DRLs (based either on means or medians) approximates 10% to 20% and, (3) this variability is only moderately dependent on the number of DLP data per device and on the number of devices included, provided this number is higher than 50 to 60. These results deserve further discussion at the level of national authorities, international authorities, and radiology departments.

Table 2 DRLs derived from device-related means, medians, and from all DLP data pooled

Body region		Head	Thorax	Abdomen	Lumbar spine
D-P75	Mean	1027	329	712	698
	Median	967	297	617	625
P-P75	All patient's data	978	292	716	645

Dose-length product (DLP) is expressed in mGy cm

At the level of national authorities, as device-related DRLs based on means and medians differ by at least 10%—even in surveys including 100% of available CT devices—any change in the method used to report DRLs could induce an artificial increase (if changing medians into means) or decrease (if changing means into medians) of the DRLs. Recently, the International Commission on Radiation Protection (ICRP) proposed a change in the method for reporting doses at the national level by using medians instead of means [3]. If national authorities follow this proposal, they should recalculate DRLs derived from previous surveys in order to prevent erroneous dose biases, when comparing surveys performed with different methods (i.e., medians vs. means). Regarding the number of devices included in the surveys, the ICRP recommended to perform initial surveys with 20–30 devices and to increase the number of devices in subsequent surveys. Our data suggest that increasing the number of devices included from 50 to 70 would decrease the DRL variability from 40 to 20%, or even 10% if increased to at least 200 devices. Such a higher number of devices—reasonably achievable in large countries at least—could be considered as representative of all devices in the country, regardless of the sample size of examinations for each device. More importantly, the number of devices included, if above a minimum of 70, influences only weakly the DRL variability, regardless of the metric used (i.e., means, medians, or all data pooled) and regardless of the number of DLP data for each device.

At the level of international authorities, caution is essential when comparing DRLs between countries. Indeed, methods for calculating DRLs (means vs. medians vs. all data pooled, as well as the number of devices included) differ between countries and dose data strongly depend on patients' height and weight, as modern scanners adapt the delivered dose to the patient's absorption, which varies among populations [15].

At the level of radiology departments, our data confirm that small sample sizes induce huge variabilities, limiting

comparisons between doses delivered by one particular device in one particular department and national DRLs. On the one hand, these variabilities could lead to inappropriate dose optimization and impair image quality or, on the contrary, fail to perform optimization [16]. In addition, comparisons between doses delivered by one particular device in one particular department and P-P75 could induce further errors, as there are differences between D-P75 and P-P75. Finally, comparisons between the dose delivered to a particular patient and the P-P75 should be discouraged, as the most important reasons for delivering high individual doses are the patient's weight and diameter, and multiphasic or repeated acquisitions [17].

Following these comments, recommendations aiming at improving dose surveys can be proposed. First, as recommended by the ICRP, medians should be preferred to means, with recalculation of DRLs from older surveys. Second, to establish reliable DRLs with an error minimized to 10% of their actual values, at least 200 devices with at least 20 data per device should be included, with a limitation in countries where too few devices are available. Third, regional differences in patient's body habitus should be taken into account. Fourth, large samples of patient's dose complete reports should be collected to compare medians in a given department to national DRLs, ideally through dose-tracking software. Fifth, as long as uncertainty persists over the actual doses delivered to patients, national authorities should consider with caution whether any penalty should be inflicted on a department where a device might be delivering doses higher than DRLs.

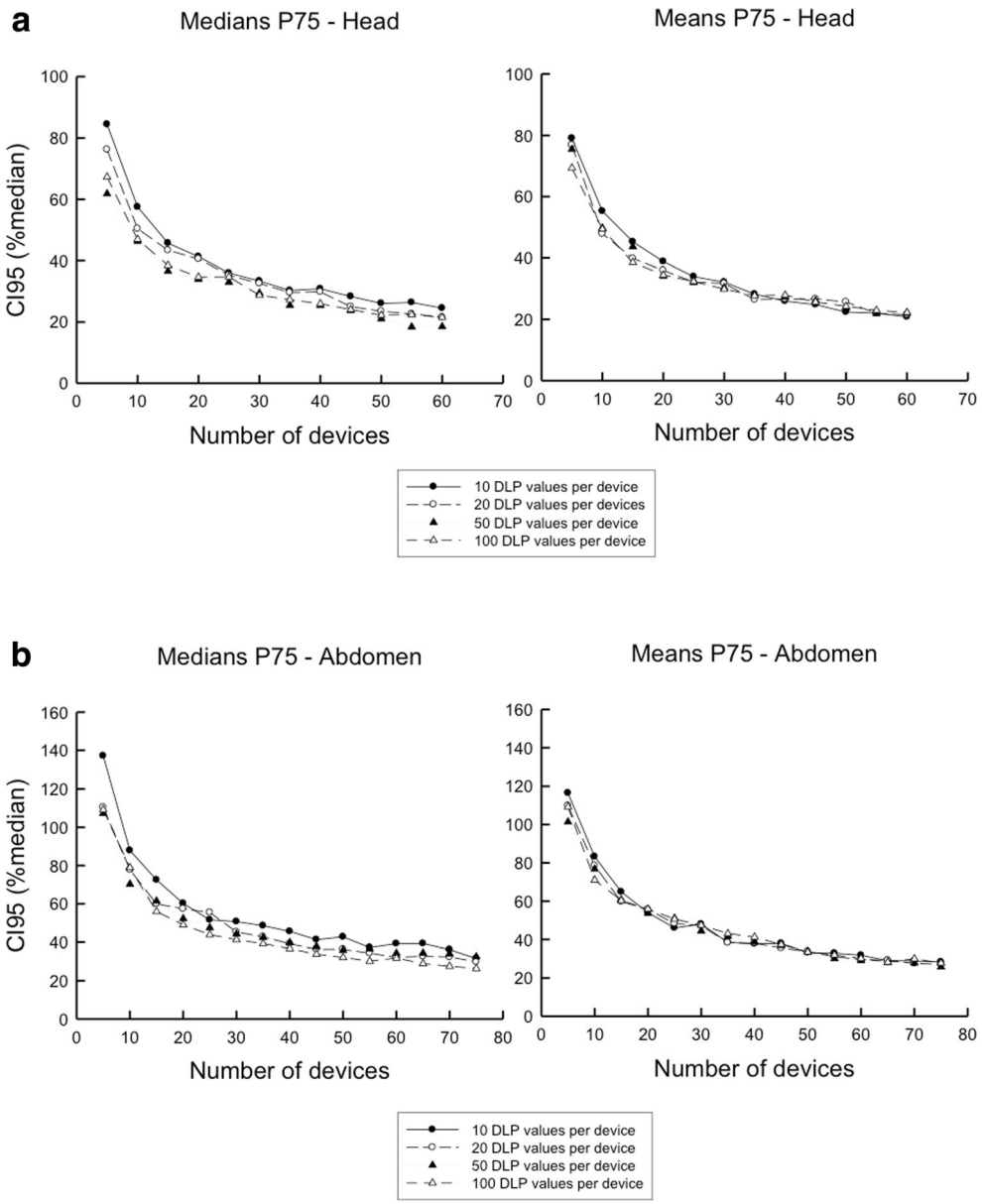

Fig. 2 Variability (95% confidence interval in percentage of median) of the 75th percentile of device-related DLP medians (left panel) and means (right panel) as a function of the number of devices included. Closed circles, open circles, closed triangles, and open triangles correspond to 10, 20, 50, and 100 DLP values per device, respectively. (**a**), head. (**b**), abdomen. (**c**), thorax. (**d**), lumbar spine

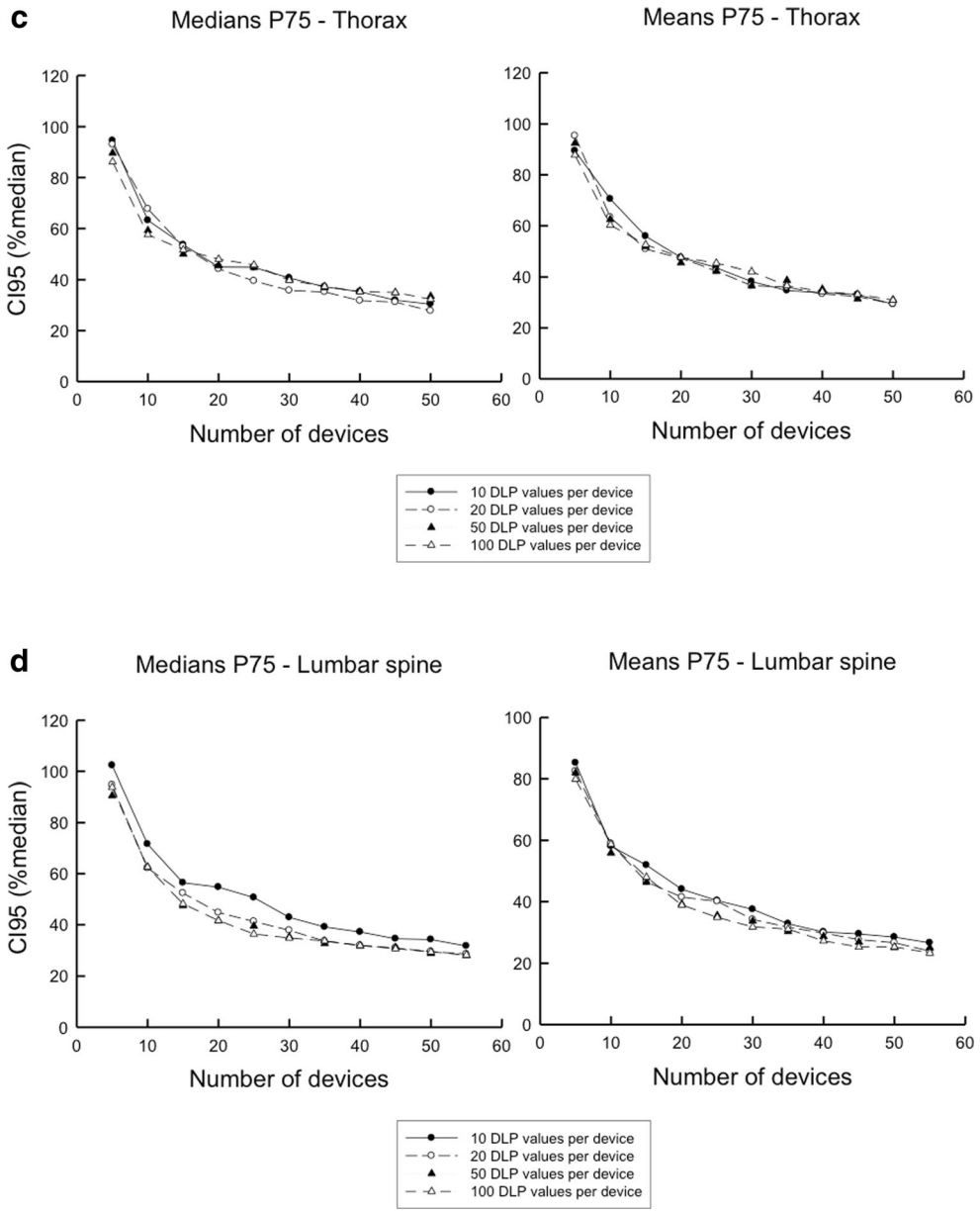

Fig. 2 (continued)

This study has a limitation. As the sample size varied between devices, the bootstrap procedures were based on the 60 to 80 devices providing the largest samples, instead of using all the devices available. However, as the drop in the variability of DRLs illustrated in Fig. 2 tends to stabilize when more than 40 devices are included, such bias would be minimal. In addition, we were unable to investigate factors influencing DLP and DRL variabilities such as patient size, number of acquisitions per examinations—particularly in abdominal studies, scanner type, as well as the variability in operator-defined factors, especially scanned volume, knowledge and use of dose reduction software, and diagnostic quality of images which may not always be taken for granted. Finally, it could be interesting to link DLP values to exam codes in order to reduce the variability of both DLP and their corresponding DRLs. This would permit homogenization of the clinical indications and the number of acquisition phases per examination. However, the more categories are created, the lower will be the number of data collected per category within a 3-month period (as allowed by law for data collection), and the higher the local DLP variability.

In conclusion, DRLs can differ, depending on whether they are based on medians, means, or all data pooled. DRLs can vary by 10% to 40% depending on the number of devices included in the surveys. Comparisons between local dose data and DRLs can be over- or underestimated and should therefore be considered with caution. Harmonization of the computing method should be recommended between the authorities of the EU states.

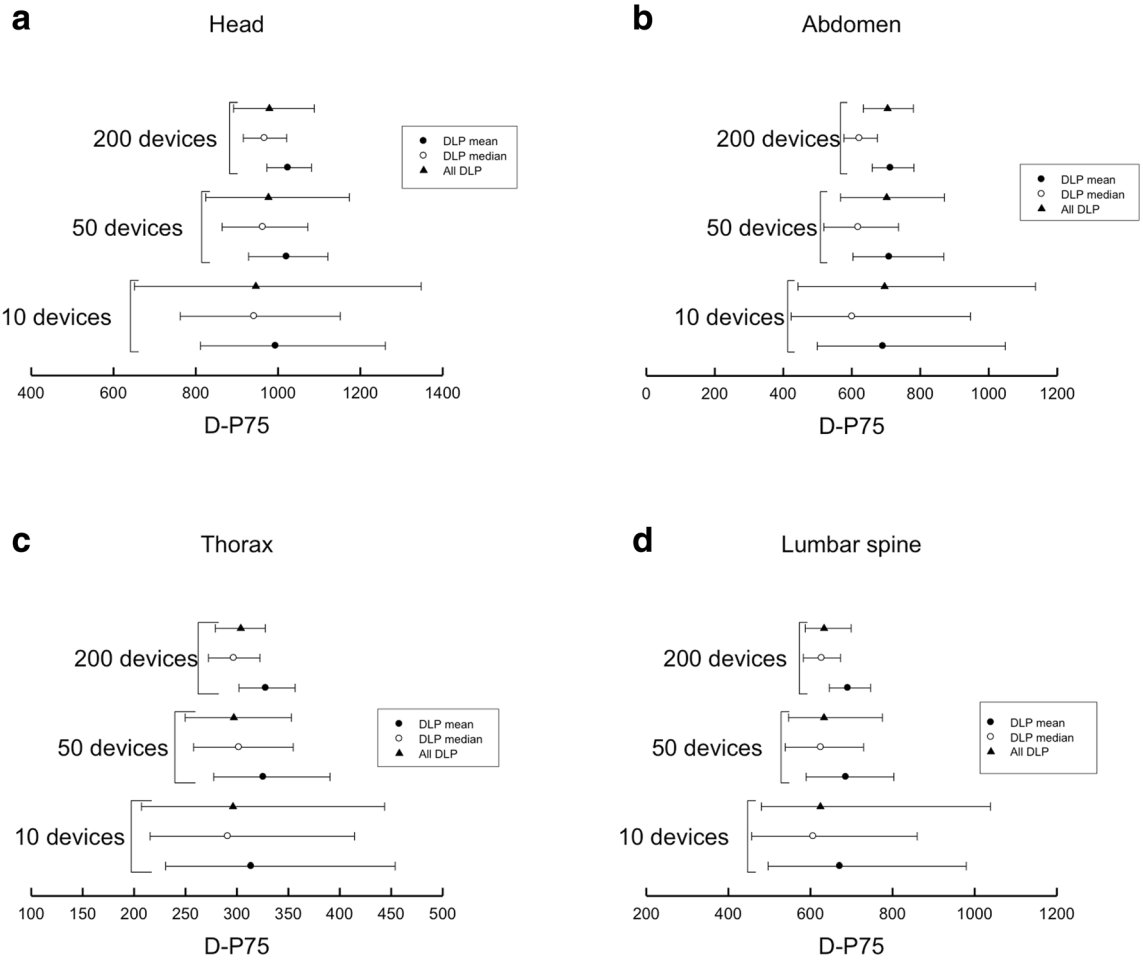

Fig. 3 Median and 95% confidence interval of the 75th percentile (D-P75, computed considering one DLP mean value) (closed circles) and one DLP median value per device (open circles); median and 95% confidence interval of all DLP values of the centers included (closed triangles). These medians and intervals were computed considering 10, 50, or 200 devices. (**a**) Head. (**b**) Abdomen. (**c**) Thorax. (**d**) Lumbar spine

Funding The authors state that this work has not received any funding.

Compliance with ethical standards

Guarantor The scientific guarantor of this publication is Denis Tack.

Conflict of interest The authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

Statistics and biometry One of the authors has significant statistical expertise.

Informed consent Written informed consent was not required for this study because according to the EU legislation (i.e., the Directive 95/46/EC regarding the protection of data of individuals), a purely observational retrospective study with complete anonymization of the data at the source, which removes any possibility of identifying the individual patients, is not subject to ethical review or to written informed consent.

Ethical approval Institutional Review Board approval was not required because for the same reasons as the informed consent. (See above.)

Methodology
- retrospective
- observational
- anonymized dose registry data analysis at national level.

References

1. The council of the European Union (1997) Council Directive 97/43/Euratom of 30 June 1997 on health protection of individuals against the dangers of ionizing radiation in relation to medical exposure, and repealing Directive. 84/466/Euratom Available via: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31997L0043:EN:HTML. Accessed 21 Jun 2016
2. The council of the European Union (2013) Council Directive 2013/59/Euratom of 5 December 2013 on laying down basic safety standards for protection against the dangers arising from exposure to ionising radiation, and repealing Directives 89/618/Euratom, 90/641/Euratom, 96/29/Euratom, 97/43/Euratom and 2003/122/Euratom. Available via http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2014:013:0001:0073:EN:PDF. Accessed 21 Jun 2015

3. Authors on behalf of ICRP, Vañó E, Miller DL, et al (2017) ICRP Publication 135: Diagnostic Reference Levels in Medical Imaging. Ann ICRP 46:1–144
4. Office Fédéral de Santé Publique Suisse (2010) Niveaux de référence diagnostiques (NRD) en tomodensitométrie. Available via http://www.bag.admin.ch/themen/strahlung/10463/10958/index.html?lang=fr&download=NHzLpZeg7t,lnp6I0NTU042l2Z6ln1ae2IZn4Z2qZpnO2Yuq2Z6gpJCJdXx6fWym162epYbg2c_JjKbNoKSn6A%2D%2D. Accessed 23 Apr 2016
5. Agence Fédérale de Contrôle Nucléaire AFCN (2015) Niveaux de référence pour les examens CT. Available via https://afcn.fgov.be/fr/professionnels/professions-medicales/applications-radiologiques/nrd. Accessed 23 Jun 2016
6. Stamm G (2012) Collective radiation dose from MDCT. Critical review of survey studies. In: Tack D, Kalra MK, Gevenois PA (eds) Radiation dose from multidetector CT. Springer, Heidelberg, pp 209–229
7. Tack D, Jahnen A, Kohler S et al (2014) Multidetector CT radiation dose optimisation in adults: short- and long-term effects of a clinical audit. Eur Radiol 24:169–175
8. European commission (1998) Radiation protection 102: implementation of the medical exposure directive (97/43/Euratom). Proceedings of the international workshop held in Madrid on 27 April 1997. Available via http://ec.europa.eu/energy/sites/ener/files/documents/102_en_0.pdf. Accessed 21 Jun 2015
9. European commission (1999) Radiation protection 109: guidance on diagnostic reference levels for medical exposures. Available via https://ec.europa.eu/energy/sites/ener/files/documents/109_en.pdf. Accessed 21 Jun 2015
10. Shrimpton PC, Hillier MC, Lewis MA, Dunn M (2006) National survey of doses from CT in the UK: 2003. Br J Radiol 79:968–980
11. Jessen KA, Shrimpton PC, Geleijns J, Panzer W, Tosi G (1999) Dosimetry for optimisation of patient protection in computed tomography. Appl Radiat Isot 50:165–172
12. IRSN (2010) Doses délivrées aux patients en scanographie et en radiologie conventionnelle. Résultats d'une enquête multicentrique en secteur public. Rapport DRPH/SER N°2010-12. Available via http://www.irsn.fr/FR/expertise/rapports_expertise/Documents/radioprotection/IRSN-Rapport-dosimetrie-patient-2010-12.pdf. Accessed 21 Dec 2017
13. AFCN/FANC (2011) La dosimétrie des patients - Arrêté de l'Agence fédérale de contrôle nucléaire du 28 septembre 2011 concernant la dosimétrie des patients. Available via http://www.jurion.fanc.fgov.be/jurdb-consult/consultatieLink?wettekstId=15119&appLang=fr&wettekstLang=fr. Accessed 21 Dec 2017
14. NRPB (1992) Institute of Physical Sciences in Medicine, Institute of Physical Sciences in Medicine (IPSM), National Radiological Protection Board (NRPB) and College of Radiographers (CoR). National protocol for patient dose measurements in diagnostic radiology. Available via https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/337175/National_Protocol_for_Patient_Dose_Measurements_in_Diagnostic_Radiology_for_website.pdf. Accessed 21 Dec 2017
15. European Commission (2008) European guidance on estimating population doses from medical X-ray procedures. Radiation Protection 154. Luxembourg. Available via http://ddmed.eu/_media/background_of_ddm1:rp154.pdf. Accessed 21 Dec 2017
16. European Commission. Radiation protection. 2014. Diagnostic reference levels in thirty six European countries. Part 2/2. RP180. Available via https://ec.europa.eu/energy/sites/ener/files/documents/RP180%20part2.pdf. Accessed 5 Oct 2018
17. Taylor S, Van Muylem A, Howarth N, Gevenois PA, Tack D (2017) CT dose survey in adults: what sample size for what precision? Eur Radiol 27(1):365–373
18. Smith-Bindman R, Wang Y, Yellen-Nelson TR, et al (2017) Predictors of CT radiation dose and their effect on patient care: a comprehensive analysis using automated data. Radiology 282(1): 182–193
19. European Commission (1995) Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement ofsuch data. Available via http://eur-lex.europa.eu/LexUriServ/ LexUriServ.do?uri=CELEX:31995L0046:en:HTML. Accessed 12 Oct 2018