IPEM Topical Report

Honey, Ian; Rose, Amy; Baker, Chris; Charnock, Paul; Fazakerley, Jason; Iball, Gareth Richard

Published in:
Physics in Medicine and Biology

DOI:
10.1088/1361-6560/aaf112

Publication date:
2018

Document Version
Peer reviewed version

Physics in Medicine & Biology

IPEM Topical Report: An evidence and risk assessment based analysis of the efficacy of quality assurance tests on fluoroscopy units – part I; dosimetry and safety

To cite this article before publication: Mark Worrall *et al* 2019 *Phys. Med. Biol.* in press https://doi.org/10.1088/1361-6560/ab3bd7

IPEM Topical Report: An evidence and risk assessment based analysis of the efficacy of quality assurance tests on fluoroscopy units – part I; dosimetry and safety

Authors Mark Worrall[1], Dan Shaw[2], Chris Baker[3], Paul Charnock[4], Jason Fazakerley[4], Ian Honey[5], Gareth Iball[6], Manthos Koutalonis[7], Mandy Price[8], Caroline Renaud[9], Amy Rose[8] and Tim Wood[10]

[1]NHS Tayside, Dundee, United Kingdom

[2]The Christie NHS Foundation Trust, Manchester, United Kingdom

[3]Royal Free London NHS Foundation Trust, London, United Kingdom

[4]Integrated Radiological Services, Liverpool, United Kingdom

[5]Guy's and St Thomas' NHS Foundation Trust, London, United Kingdom

[6]Leeds Teaching Hospitals NHS Trust, Leeds, United Kingdom

[7]Colchester Hospital University NHS Foundation Trust, Colchester, United Kingdom

[8]Barts Health NHS Trust, London, United Kingdom

[9]Imperial College Healthcare NHS Trust, London, United Kingdom

[10]Hull and East Yorkshire Hospitals NHS Trust, Hull, United Kingdom

Abstract

This work aims to assess the efficacy of x-ray quality assurance tests undertaken on fluoroscopy units in the UK. Information was gathered on the results of dosimetry and safety tests recommended by the reports of the Institute of Physics and Engineering in Medicine, and those additionally undertaken by medical physics departments. The assessment of efficacy considers the frequency with which a test result breaches the remedial level or other relevant threshold where applicable. The third quartile of those results exceeding the remedial level or threshold is used to estimate the severity of such a breach in terms of potential impact on patient dose and image quality. A risk assessment approach is then used to recommend to what degree, if any, the test should be included in an on-going test regimen.

Data was analysed from 468 testing sessions to 336 unique fluoroscopy units throughout the UK. Across all tests, the rate with which the remedial level was exceeded varied from 0 – 29.5%, with severity ranging from little or none to major degradation to image quality or significant increase on population dose. Where possible, the data has also been used to produce representative ranges for the results of dosimetric tests. These could be useful as an up to date comparator for those sites considering the purchase of or commissioning new equipment.

Overall the results indicate a wide range for the efficacy of those tests undertaken at present; this can be used to review local test protocols and to inform future changes to national guidance in the UK. The results also highlight some tests where measurement technique varies significantly throughout the UK, making any valid comparison difficult. This may indicate a need for further guidance on how best to undertake these tests.

1. Introduction

This is the second paper published by the Institute of Physics and Engineering in Medicine's (IPEM) Evidence Based Quality Assurance (QA) Working Party. The primary aim is to assess the effectiveness of various QA tests undertaken on x-ray imaging units in terms of patient and staff safety, with due consideration to rates of failure and the consequence of a failure. Secondary aims are to provide data on typical ranges of results to guide commissioning of newly installed units and to support future decisions on appropriate tolerances. A detailed discussion of the aims of the working party is given in the first paper, 'IPEM Topical Report: An evidence and risk assessment based analysis of the efficacy of tube and generator quality assurance tests on general x-ray units' by Honey et al [1].

This paper reviews the results submitted for dosimetry and safety tests on fluoroscopy units, as recommended by IPEM report 91 [2] and IPEM report 32 parts I [3], II [4] and VI [5], by working party members and colleagues from medical physics departments across the UK. Data for additional tests not recommended in the IPEM reports [2-5] was also requested and are reviewed. The analysis of results related to image quality tests forms part of a separate study.

2. Materials and methods

2.1 Data acquisition

A template spreadsheet was designed listing the fluoroscopy and fluorography tests contained within IPEM reports 91 [2] and 32 parts I [3], II [4] and VI [5], and any additional tests that were considered by the working party to have become commonplace since the publication of these reports.

The first part of the pro-forma asked the respondents to indicate how often their department undertook each of the listed tests and to provide details, including frequency, of additional tests performed. This was designed to be quick and simple to complete to encourage a high submission rate. Collection of the details of testing protocols for departments across the UK is considered useful to aid in identifying where UK experts consider deviations from the existing IPEM reports are required. Details of the methodology for each test were requested to ensure that all of the results for a single test could be compared across all testing sessions and all fluoroscopy units, with due regard for, and necessary corrections made to, the technique used to produce the result.

The second part of the pro-forma was more detailed, requesting summary test results for as many testing sessions to as many different fluoroscopy units as possible, to allow for analysis of intra and inter unit variability. For each fluoroscopy unit, the pro-forma requested information about the unit including manufacturer and model, the detector type (image intensifier (II) or flat panel detector (FPD)), the equipment category (mobile c-arm, mini c-arm or fixed fluoroscopy unit), year of installation and the size of its normal and magnification fields (diameter or diagonal length).

The pro-forma asked which procedures the unit was primarily used for (options were diagnostic cardiac and percutaneous coronary intervention (PCI), electrophysiological and device implantation, therapeutic cardiac, general cardiac, vascular, neurology, orthopaedics, theatre, barium procedures, iodine contrast procedures, urology, lithotripsy, general). For each unique testing session the pro-forma requested the date of the testing, reason for testing (commissioning, routine, fault

investigation or other) and the testing equipment used by the medical physics department for the dose measurements. The tests pertaining to dosimetry and safety included in the pro-forma and the requested test result(s) are summarised in table 1.

Table 1: Dosimetry and safety tests and the corresponding results requested in the pro-forma

Fluoroscopy		
Test	Reference	Requested result(s)
Dose rate at entrance surface of phantom under automatic dose rate control (ADRC)	IPEM 91, FLU06	Largest % change from baseline AND measured dose rates (for varying programmes and field sizes)
Dose rate to the input face of the image receptor under ADRC	IPEM 91, FLU07	Largest % change from baseline AND measured dose rates (for varying programmes and field sizes)
Dose rate to the input face of the image receptor under ADRC for a collimated beam	IPEM 32(II), 1.2.2	Measured dose rate
Dose rate at entrance surface of phantom under ADRC for a collimated beam	Additional	Measured dose rate
kV accuracy	IPEM 91, RAD12	% or absolute kV deviation from stated
Accuracy of KAP meter calibration	BIR report 18	% difference between displayed and measured
Accuracy of Reference air KERMA (rate) calibration	BS EN 60601-2-54: 2009	% difference between displayed and measured
Tube leakage	IPEM report 107	Highest measured leakage
Tube filtration	ICRP publication 33	Measured filtration
Fluoroscopy timer accuracy	BS EN 60601-2-54: 2009	Pass or fail
Fluoroscopy automatic cut-out (10 minutes continuous)	BS EN 60601-2-54: 2009	Pass or fail
Fluorography		
Test	Reference	Requested result(s)
Dose per image at the input face of the image receptor under automatic exposure control (AEC)	IPEM 91, FLG04	Largest % change from baseline AND measured dose per image (for varying programmes and field sizes)
Collimated dose per image at the input face of the image receptor under AEC	Additional	Measured dose per image (for varying programmes and field sizes)
Dose per image at the entrance surface of a phantom under AEC	Additional	Largest % change from baseline AND measured dose per image (for varying programmes and field sizes)
Collimated dose per image at the entrance surface of a phantom under AEC	Additional	Measured dose per image (for varying programmes and field sizes)

Note that although the term Automatic Brightness Control (ABC) is used in IPEM reports 91 [2] and 32 [4,5], this paper will use the term Automatic Dose Rate Control (ADRC) throughout.

The pro-forma was circulated to the medical physics community via a well-established UK based medical physics mailing list (with a membership of over 2600 at the time of circulation). Data was also submitted by members of the Evidence Based QA Working Party. The call for data was left open from September to December 2014. Contributors were asked where possible to send original test results rather than results measured after corrections had been made by an engineer. This is most likely to occur at equipment commissioning. For some units, data was received for multiple years testing of the same unit; this data has all been included in the analysis subject to a prior check to ensure that it does not skew the analysis.

Once the data had been acquired, the results were validated. Obvious outliers were identified (either very abnormal results or an above average failure rate in the data submitted by a single centre) which could be indicative of a mis-typed result or a test technique incompatible with the pro-forma or very different from that of other medical physics departments across the UK. Outliers were discussed with representatives from the sites that sent the data. If the validity of the test technique and recording of results was confirmed, the data was retained. Less than 1% of the data that was submitted was rejected.

2.2 Assessment of test efficacy using risk assessment methodology

The methodology used for assessing the test efficacy was carried out following the risk assessment methodology described in Honey et al [1]. The frequency of exceeding a remedial level was graded on a scale of 1 to 5 to give a 'likelihood rating' (LR) as described in table 2. A minimum data size of 50 was required before a LR was assigned. A severity rating (SR) was assigned based on consideration of the effects on image quality, the estimated percentage increase to the population effective dose for exposures affected by the fault and the maximum absolute increase to individual patient effective dose. Values were assigned based on the highest score from these 3 categories when applying the ranges in table 3. For each test, the third quartile of those results that exceeded remedial level was considered when assigning a SR.

The risk matrix (figure 1) indicates the possible outcome for each test. For a risk matrix category of green it is proposed that this test may no longer be needed during routine testing sessions. For tests in the yellow, orange or red categories, it is proposed that these tests are continued, but consideration is given to reducing the testing frequency of yellow category tests, maintaining the frequency for orange category tests, and increasing the frequency of red category tests (possibly by recommending their inclusion in radiographer QA programs). Alternatively, the results may indicate that current remedial levels are inappropriate.

Table 2: Definition of Likelihood Rating (LR)

Likelihood rating	Frequency (%) of exceeding a remedial level
1	<2
2	≥2 and <5
3	≥5 and <10
4	≥10 and <15
5	≥15%

Table 3: Definition of Severity Rating (SR)

Severity Rating	Effect On Image Quality	% increase on Population Dose	Max. Increase In Individual Patient Dose (mSv)
1	Little/none	<5	0.1
2	Small, unlikely to have much effect	≥5 and <10	0.1-0.5
3	Noticeable, but clinical IQ likely to still be acceptable	≥10 and <20	0.5-2
4	Significant degradation in clinical image quality	≥20 and <40	2-5
5	Major degradation in image quality likely to seriously affect diagnosis	≥40	>5 or externally reportable incident

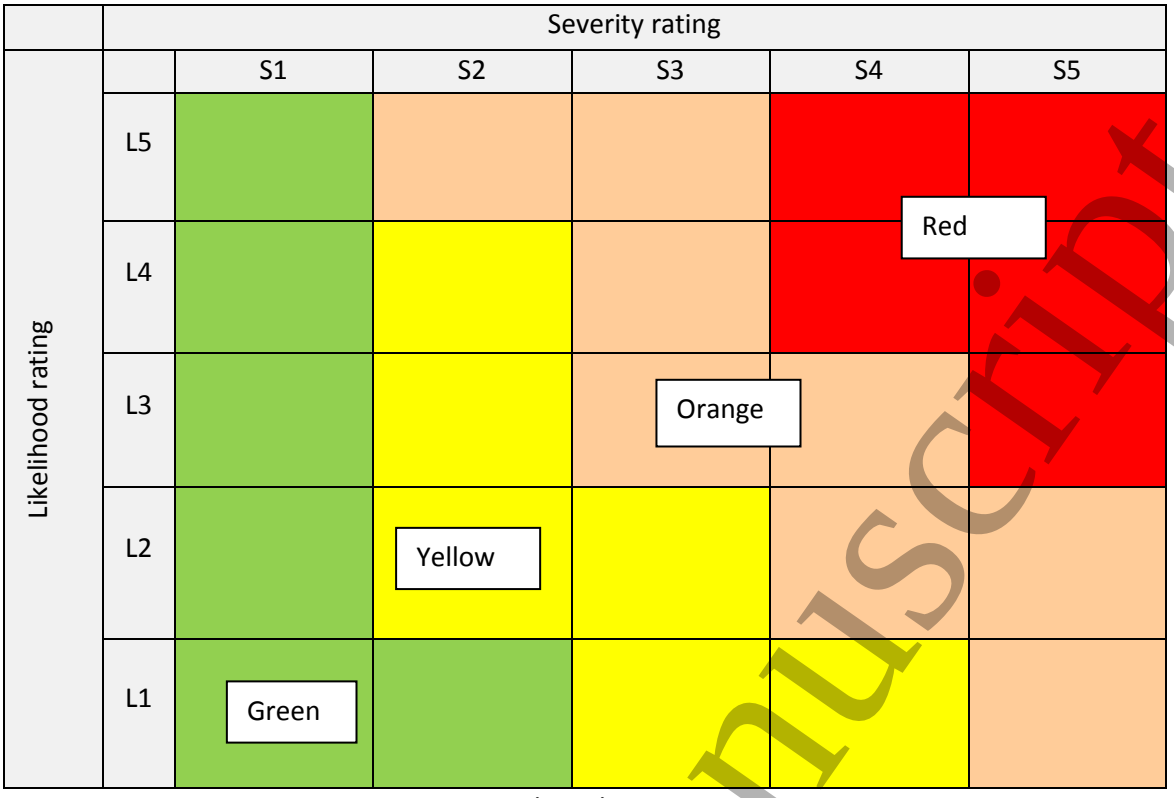

Figure 1: The risk rating matrix

2.3 Statistical analysis

For the statistical analysis a linear mixed effect model was employed [6]. It was postulated each result was potentially influenced by the 'testing reason' – the reason for undertaking the tests (e.g. commissioning, routine, fault investigation or other) where this is a relevant factor. Since many of the tests involve a comparison with baseline, the results from commissioning testing sessions would not be relevant as these results are used for establishing the baseline. Further influencing factors include the 'equipment age' (> 10 years old or < 10 years old), the 'equipment manufacturer', the 'medical physics department' undertaking the tests, the 'equipment type' (e.g. mobile fluoroscopy unit, fixed fluoroscopy unit etc.) and the 'detector type' (II or FPD). All were considered as fixed effects. Most of the analysis was undertaken with the data already separated by detector type (i.e. II or FPD). For some tests, additional analysis was undertaken with all of the data together, with detector type added to the list of fixed effects.

Equipment identification code was included as a random effect and is denoted by (1|ID), where ID is an identification number unique to the unit. This was considered necessary as a correlation between results for the same piece of equipment returned for testing performed over several years cannot be assumed to be independent. A model was then evaluated for the results of each individual test. The square brackets indicate factors not included for the analysis of every test. For example, for some tests commissioning results are not relevant and, in these situations, 'testing reason' was omitted. Likewise, 'detector type' was omitted when the test results were already stratified by II and FPD units. As per the notation of Winter [6], the model can be written as;

Result ~ [Testing Reason +] Equipment Age + Equipment Manufacturer + Medical Physics Department + Equipment Type [+ Detector Type] + (1|ID) (1)

The statistical analysis package 'R' (The R Foundation for Statistical computing, version 3.3.3) was used to perform this analysis. Null models were created by removing one of each of the fixed variables in turn and comparing against the model using an Analysis of Variance (ANOVA) test. This returned a p value which indicated whether removing a parameter from the model had a statistically significant effect on the result. For the purpose of this work p-values of <0.025 were considered statistically significant. For this analysis, the absolute value of all results was considered as it was known that some centres had returned results that included positive and negative values whilst others had returned only the absolute value. A logarithmic transform was applied to all data in order to ensure linearity, homoscedasticity and normality.

It is worth noting that this work does not intend to examine or comment upon differences between manufacturers. Such comparisons could not be undertaken fairly because sample sizes will not be equal and nothing is known about how well optimised any unit is. Nevertheless, it was felt worth including manufacturer as a fixed effect to determine if this is an important consideration for departments to consider locally.

3. Results and discussion

Data were received from 25 UK based medical physics departments. Of these, all 25 provided data pertaining to test frequencies and 15 contained test data for at least one testing session to one fluoroscopy unit. In total data were submitted for 468 testing sessions to 336 unique fluoroscopy units as indicated in table 4.

Table 4: The number of units by detector type and equipment type that sets of data were provided for

Detector type	Number of unique units
Image intensifier	213
Flat panel detector	121
Unspecified	2
Equipment type	
Fixed fluoroscopy units	135
Mobile c-arms	186
Mini c-arms	12
Unspecified	3

For the 468 testing sessions, there were data for at least one of the tests identified in table 1. The years of first installation range from 1997 to 2015 with a median of 2008. Test results for the reported data extended from February 2006 to January 2015.

Whilst the pro-forma asked which procedures each unit was primarily used for, it is noteworthy that for 68% of units this question was either not completed or it was indicated that the equipment was

for general use. It was indicated by some large regional departments that this information was often not forthcoming, and it was also commented that often equipment was used for more than one primary purpose, particularly in smaller centres that have limited equipment. The multi-functional nature of modern equipment has been made possible by being able to run several different software packages on one unit making it suitable for many different applications.

3.1 Fluoroscopy results

3.1.1 Frequency of testing

The results pertaining to the frequency with which each test is undertaken by the responding medical physics departments are summarised in table 5.

Table 5: Frequency with which each fluoroscopy test is undertaken by responding medical physics departments

Fluoroscopy Tests	Test identifier	Never	Commissioning only	Annual	Biennial
Dose rate at entrance surface of phantom under ADRC	IPEM 91, FLU06	-	1	23	1
Dose rate to the input face of the image receptor under ADRC	IPEM91, FLU07	-	-	24	1
Dose rate to the input face of the image receptor under ADRC for a collimated beam	IPEM 32(II), 1.2.2	15	-	4	-
Dose rate at entrance surface of phantom under ADRC for a collimated beam	Additional	15	-	3	-
kV accuracy	IPEM 91, RAD12	1	2	19	1
Accuracy of KAP meter calibration	BIR report 18	-	1	23	1
Accuracy of Reference air KERMA (rate) calibration	BS EN 60601-2-54: 2009	6	3	11	-
Tube leakage	IPEM report 107	-	20	3	-
Tube filtration	ICRP publication 33	-	9	14	1
Fluoroscopy timer accuracy	BS EN 60601-2-54: 2009	15	4	1	-
Fluoroscopy automatic cut-out (10 mins continuous)	BS EN 60601-2-54: 2009	6	10	7	-

3.1.2 Entrance Surface Dose Rate (ESDR) and Detector Input Dose Rate (DIDR) under ADRC

IPEM Reports 91 [2] and 32 part II [4] and Martin et al [7] recommend measuring the ESDR on a 20cm thick water, or water equivalent, phantom located at the approximate position of the patient. 18.5cm of polymethacrylate (PMMA) is a good approximation to 20cm of water, though there is a dependence on beam quality [7]. The attenuating medium and thickness used by 13 of the 14 medical physics departments that provided data for this test are shown in table 6. The remaining centre used a 20cm Perspex attenuator at commissioning and a 3.6mm copper attenuator positioned at the detector for routine testing to reduce the equipment required for routine testing.

Table 6: The attenuating mediums and their thicknesses used by the responding medical physics departments

No. of medical physics departments	Thickness		
Medium	15cm	18.5cm	20cm
PMMA / Perspex	1	3	5
Water	0	0	4

For DIDR under ADRC, IPEM reports 91 [2] and 32 part II [4] recommend placing a copper attenuator at the output of the x-ray tube and making dose rate measurements at the image receptor.

The IPEM 91 [2] remedial level for both of these tests is baseline ±25% and the suspension level is baseline ±50% or greater. IPEM report 91 [2] has a further remedial level for ESDR of a phantom under ADRC as > 50mGy min^{-1} and a suspension level of > 100mGy min^{-1} for a standard sized patient at the largest available field size. This is only for fluoroscopy modes, not acquisition.

3.1.2.1 Absolute ESDR values

ESDR results exceeding the remedial and suspension levels of 50mGy min^{-1} and 100mGy min^{-1} respectively for the largest field size and for a standard sized patient (attenuators in the range of 15 to 20cm were considered as equivalent) were analysed for 400 testing sessions to 251 unique units. As most testing sessions contained data for more than one fluoroscopy program, a total of 1332 different program combinations were analysed. There were 2 instances of a dose rate exceeding the 50mGy min^{-1} remedial level and these were for two different programs on the same unit. There were no instances of the suspension level being exceeded.

3.1.2.2 Percentage change from baseline

The results for the analysis of percentage change from baseline for ESDR of a phantom under ADRC are given in table 7 and for DIDR under ADRC in table 8 (excluding commissioning as these measurements form the baseline data for subsequent analysis), where n is the number of submitted data sets in the analysis (i.e. from testing sessions fitting the described criteria)

Table 7: ESDR of phantom under ADRC

Analysis	n	Exceeding remedial level	Exceeding suspension level	Mean result of those exceeding remedial	3rd quartile result of those exceeding remedial	Maximum deviation from baseline
All testing sessions on all units	221	25.8%	8.6%	44.7%	54.5%	104.5%
Most recent testing session on all unique units	142	25.4%	8.5%	47%	56.2%	104.5%
All testing sessions on image intensifiers	128	28.9%	7%	42.3%	48.6%	104.5%
All testing sessions on flat panel detectors	93	21.5%	10.8%	49.3%	57%	90%

Table 8: DIDR of the image receptor under ADRC

Analysis	n	Exceeding remedial level	Exceeding suspension level	Mean result of those exceeding remedial	3rd quartile result of those exceeding remedial	Maximum deviation from baseline
All testing sessions on all units	288	29.5%	5.2%	44.5%	42.9%	233.3%
Most recent testing session on all unique units	203	29.1%	5.4%	46.0%	40.5%	233.3%
All testing sessions on image intensifiers	189	34.4%	5.8%	45.6%	43.5%	233.3%
All testing sessions on flat panel detectors	97	20.6%	4.1%	40.7%	36.9%	115.0%
All testing sessions on fixed fluoroscopy units	80	23.8%	2.5%	38.2%	37.5%	115.0%
All testing sessions on interventional units	53	24.5%	3.8%	36.1%	41.5%	72.0%
All testing sessions on mobile c-arm units	147	33.3%	7.5%	49.7%	47.5%	233.3%
All testing sessions on mini c-arms	6	66.7%	0.0%	37.1%	41.5%	43.8%

For all testing sessions to all units, the failure rates for both tests received a LR of 5. The SR was assigned based on the 3rd quartile result. The analysis considered the absolute value of all results. It was considered that an increase in the dose rate would lead to increased population dose, whereas a drop in dose rate could have a detrimental effect on image quality through increased noise and reduced low contrast spatial resolution. For both tests this resulted in a SR of 5. This gives both tests a risk matrix category of red indicating the importance of remaining part of routine testing and the continued inclusion of dose rate monitoring as part of radiographer QA.

For the ESDR of a phantom under ADRC, results were found to be statistically significant related to 'equipment manufacturer' ($p<0.01$) and 'medical physics department' ($p<0.01$). There was no statistical significance related to 'equipment age' ($p=0.78$), 'equipment type' ($p=0.51$) or 'detector type' ($p=0.84$).

For DIDR there was no statistical difference for 'equipment age' ($p=0.03$), 'equipment manufacturer' ($p=0.08$), 'medical physics department' ($p=0.56$), 'equipment type' ($p=0.56$) or 'detector type' ($p=0.70$).

'Medical physics department' was found to be a significant factor ($p<0.01$) for ESDR of a phantom under ADRC but not for DIDR ($p=0.56$). The data submitted indicates that test methods for DIDR are closely aligned across departments [2, 4] but less so for ESDR. Further guidance on ESDR testing, including a proposed standardisation of technique, would be welcome.

Whilst both tests demonstrate high fail rates, it is not thought to be completely indicative of performance fault condition. It is known that manufacturers' representatives make adjustments to fluoroscopy units as new software and processing is developed, or in response to user requests. Some units also allow users to make adjustments to programmes. Where the difference in dose rate is the result of an intentional and approved change to the unit it is not appropriate to compare the result to a previous baseline. For ESDR of a phantom under ADRC, statistical significance was found between manufacturers, which could indicate that the equipment fails this test more often for some manufacturers or that some manufacturers undertake these on-going adjustments more often than others. The significance found between medical physics departments could reflect the relative distribution of these manufacturers in their reported data.

The data submitted to the working party cannot distinguish between intentional and unintentional changes to dose rate. Following discussion with representatives from medical physics departments that submitted data it was noted that the Medical Physics Expert (MPE) is often not informed of intentional changes (a particular concern for larger regional medical physics departments where there is reliance on the equipment users to pass on information). The representatives of the medical physics departments that submitted data were of the opinion that whilst a breach of remedial level would trigger an investigation in which intentional changes would be identified, changes to dose rate would not be automatically requested because of the result. Little guidance is given in the literature regarding when it is appropriate to set new baselines, or indeed when a unit is no longer acceptable for clinical use and it is left as a task for the MPE in collaboration with the users to make these decisions.

A possible solution would be to only apply remedial levels in the form of percentage change from baseline to dose rates measured using a dedicated medical physics or QA program, or a user quality control mode as described in NEMA XR-27 [8]. Typically, these programs are set up by the manufacturer during commissioning and are stored such that they will be unaffected by future adjustments to clinical imaging programs. These programs are only used during QA testing and offer a true comparison against baseline.

It is still necessary to measure the dose rate for a representative sample of regularly used clinical programmes during QA testing to ensure that significant changes are detected and investigated, however it is not useful to apply remedial levels in the form of percentage change from the baseline. For clinical programmes, and for all programmes on units where an unchanging QA program is not an option, representative dose rate ranges might be more useful (discussed in section 3.1.5). The MPE should be involved in considering whether the measured dose rates are appropriate.

It is possible that intentional reductions in dose rate following improvements to image processing will not affect the image quality. The working party is unable to differentiate between results exceeding the remedial level that are the result of intentional changes to the unit or not. The SR of 5 is assigned in the knowledge that the 3rd quartile result of those exceeding the remedial level could have had little or no effect, but to account for the possibility that it did. In either case, the ESDR and DIDR needs to be routinely tested so changes can be found and investigated.

3.1.3 Variation with dose rate setting

Fluoroscopy equipment is provided with two or more [9] dose modes (typically low, normal and high) for each program to allow the operator to make adjustments in relation to dose rate and the required level of image quality. Martin et al [7] undertook an evidence based exercise in 1997 on 80 II units to evaluate the effect of varying dose settings on the ESDR of a phantom. The 2002 edition of the MDGN [10] used their results to recommend that for a 20 – 30cm field size and 20cm of water equivalent attenuator the low dose mode should result in an ESDR less than 50% that of the normal setting, and the high dose setting should result in an ESDR greater than 150% of the normal setting. The IEC standards require that for interventional equipment, low and normal dose modes should be supplied such that the low dose mode provides an ESDR that is 50% of the normal dose rate setting, measured at the patient entrance reference point (PERP, formerly the interventional reference point (IRP)) [9,11]. It should be noted that this variation does not directly translate to DIDR as the effects on ESDR can be partly achieved through changes to beam quality.

The data submitted allowed for a review of this relationship. The analysis was restricted to the most recent test results for each unique unit for ESDR with an attenuator thickness in the range of 18.5 – 20cm. The results for low to normal dose and normal to high dose for all units are shown in tables 9 and 10.

Table 9: The ratio of ESDR measured for Low and Normal dose settings for all units by field size

Field size (cm)	<20	20 – 30	>30
No. Unique units with field size	55	128	61
Minimum Low – Normal ratio	0.32	0.24	0.22
1st quartile Low – Normal ratio	0.43	0.46	0.46
Median Low – Normal ratio	0.51	0.61	0.52
Average Low – Normal ratio	0.57	0.61	0.59
3rd quartile Low – Normal ratio	0.67	0.70	0.78
Maximum Low – Normal ratio	1.07	1.39	0.99

Table 10: The ratio of ESDR measured for High and Normal dose settings for all units by field size

Field size (cm)	<20	20 – 30	>30
No. Unique units with field size	16	93	59
Minimum High - Normal ratio	1.01	0.45	0.58
1st quartile High – Normal ratio	2.13	1.53	1.35
Median High – Normal ratio	2.33	1.75	1.48
Average High – Normal ratio	2.09	1.76	1.59
3rd quartile High - Normal ratio	2.36	2.11	1.77
Maximum High - Normal ratio	2.54	2.5	3.39

These results suggest some variation from the ratios suggested by the IEC [9,11] and Martin et al [7]. The mean Low – Normal ratio is quite consistent across all field sizes at 0.59, above the < 0.5 suggested. The median ratio is close to 0.5 except for the 20-30cm field size. The mean High – Normal ratio did not exhibit the same consistency across field size but was greater than the 1.5 suggested by the IEC [9,11] and Martin et al [7] for all field sizes. The median ratio was also above 1.5 for all but the >30cm field size.

Comparisons with the ratios in the IEC standard must take into account the measurement conditions. The IEC measurements are specified at the PERP, whereas values derived in this work are from ESDR testing. The IEC standards are intended as a guide to equipment manufacturers for the design and construction of the equipment, whereas ratios derived from ESDR measurements are more clinically representative. These can be used by the MPE for optimisation.

Anomalous results such as the Low setting giving a higher ESDR than Normal (ratios > 1 in table 9) and the High setting giving a lower ESDR than Normal (ratios < 1 in table 10) indicate the usefulness of this test. The Low setting gave a higher ESDR than Normal in 1.6% (4) of the results analysed, and the Higher setting gave a lower ESDR than Normal in 4.8% (8) of results analysed. Discussion with representatives from the sites that reported these failures indicated the causes are related to incorrectly set up clinical programs at equipment commissioning or following software upgrade, or changes to the clinical programs made throughout the lifetime of the equipment. Another cause reported is a failure to reload user pre-sets properly following equipment maintenance. Some service agents choose to save user pre-sets prior to starting maintenance work and work with default settings to prevent accidental changes to user programs. These must be properly loaded back onto the unit at the end of the maintenance work and this process may occasionally fail.

3.1.4 Variation with field size

Data for all the units submitted indicated that they all had a range of field sizes. Historically for II units, it has been necessary for the dose rate to increase with an inversely proportional relationship to the area of the selected field size in order to maintain a constant light photon flux at the output phosphor. For FPD units this constraint does not apply. Instead, the emphasis is on maintaining a constant perceived noise level with the changing ratio of display size to irradiated field size (i.e. the effective pixel size). As such, it is still expected that dose rate will increase towards smaller field sizes.

The measured ESDR for all field sizes under continuous fluoroscopy with a normal dose mode and 18.5 – 20cm attenuator is shown in figure 2 for II units and figure 3 for fixed installations with FPD. Data for fixed installations with II is from 13 unique units, data for mobile units with II is from 137 unique units and data for fixed installations with FPD is from 15 unique units.

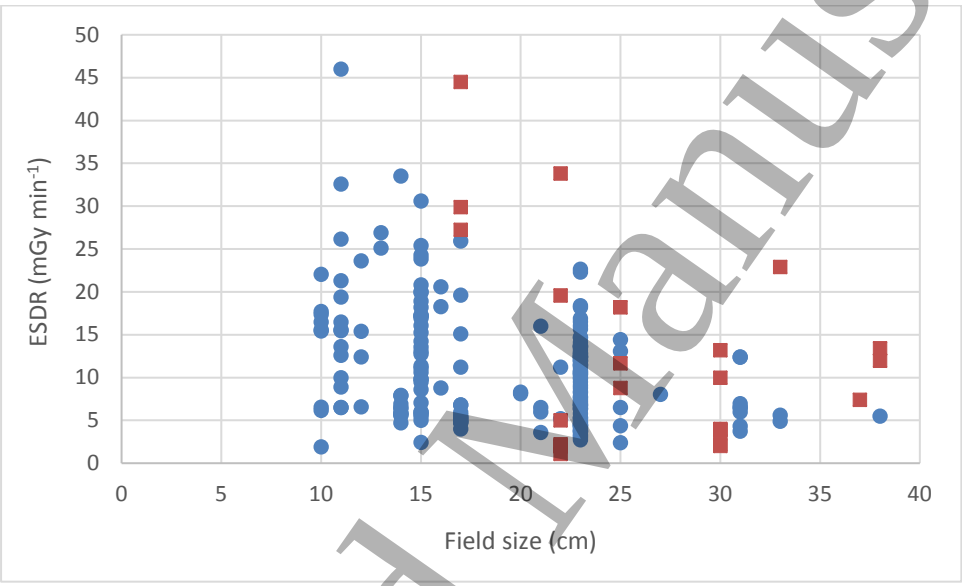

Figure 2: ESDR for different field sizes for mobile (blue circles) and fixed installation (red squares) II units measured under continuous fluoroscopy with a normal dose mode and 18.5 – 20cm of PMMA or Water attenuating medium

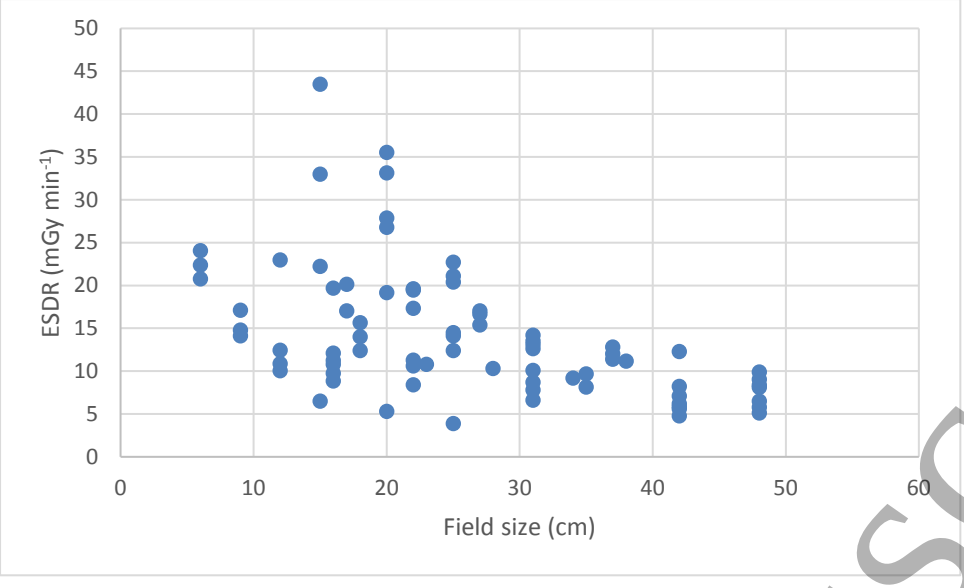

Figure 3: ESDR at different field sizes for fixed installation FPD units measured under continuous fluoroscopy with a normal dose mode and 18.5 – 20cm of PMMA or Water attenuating medium

Figures 2 and 3 demonstrate a large variation in the measured ESDR for the same field size for both fixed installations and mobile units and II and FPD units. Further analysis indicates a level of consistency for identical make and model where the dose rates were generally found to agree to within 25% for an equivalent attenuator and program. This is maintained across different medical physics departments. However, there is little agreement across models by the same manufacturer or between different manufacturers. This is shown in figure 4 for mobile II units and in figure 5 for fixed FPD units. For each manufacturer and equivalent model, the mean ESDR from at least 3 different units for continuous fluoroscopy, normal dose and using a normal/standard/general fluoroscopy program is shown.

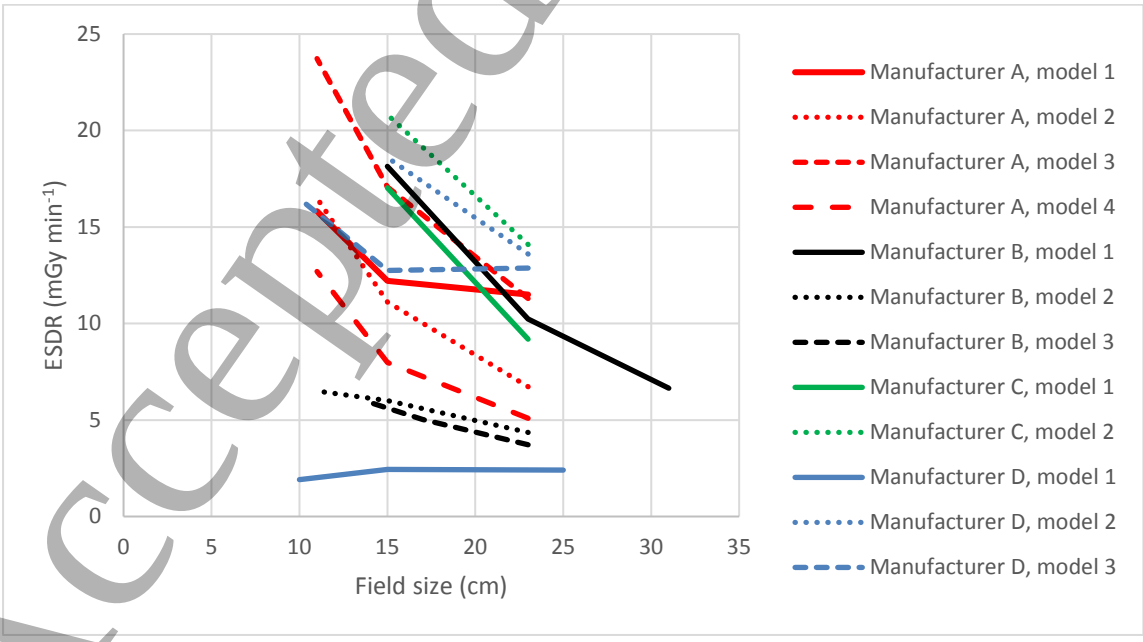

Figure 4: Mean ESDR for different manufacturer and model mobile II units where data received for least 3 different units of same manufacturer and model measured using a general fluoroscopy programme, normal dose mode, continuous fluoroscopy and 18.5 – 20cm of PMMA or Water attenuating medium

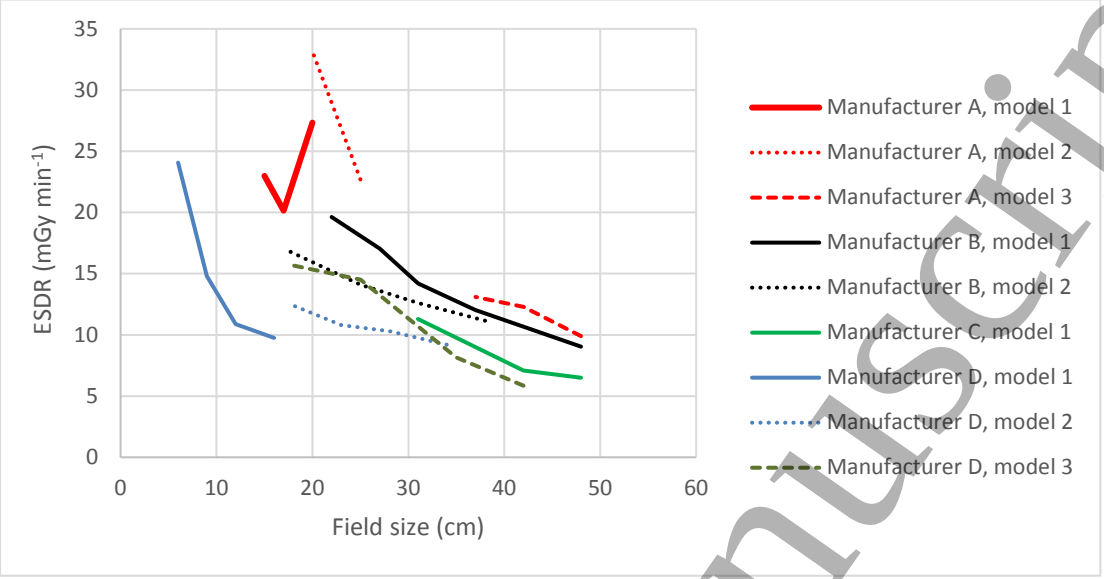

Figure 5: Mean ESDR for different manufacturer and model fixed FPD units where data received for least 3 different units of same manufacturer and model measured using a general fluoroscopy programme, normal dose mode, continuous fluoroscopy and 18.5 – 20cm of PMMA or Water attenuating medium

The clinical program selected was also observed to have an effect. The analysis can offer no useful range of normal values for variation in ESDR with field size. Local values should be compared against the data provided by the manufacturer; therefore it is important that this data is available. Comparison between equivalent units from the same manufacturer where the results were acquired using similar settings may also be a valid approach for establishing the expected ESDR for those units.

3.1.5 Representative dose rate ranges

Information was requested pertaining to the equipment's primary use. For 68% of submissions, the field was left blank or indicated general use. Sub-division of results by confirmed equipment use resulted in sample sizes too small ($n < 10$) to produce reliable representative ranges by clinical application. Sub-division by program name was considered, however there was too great a variability in program name with manufacturer and model to be confident that the comparison was appropriate.

3.1.5.1 Entrance surface dose rate

Representative ranges for ESDR to a phantom under ADRC are presented for fixed fluoroscopy and mobile fluoroscopy units in table 11. The representative ranges are for the largest field size, continuous fluoroscopy, 18.5 – 20cm of attenuator and all clinical programs.

Table 11: Representative ESDR for fixed II, mobile II and fixed FPD units for 18.5-20cm of attenuating medium

Analysis	Dose mode	Diagonal Field Size (cm)	n	Median (mGy min^{-1})	Range (mGy min^{-1})	Inter quartile range (mGy min^{-1})
Fixed Image Intensifiers	Normal, Continuous	25 - 40	14	6.0	2.0 – 41.2	4.1 - 12.0
Mobile image intensifiers	Normal, Continuous	22 - 33	130	8.7	0.8 – 23.7	3.8 - 12.1
Fixed Flat Panel Detectors	Normal, Continuous	20 - 48	53	11.1	1.8 – 34.2	5.3 – 17.9

3.1.5.2 Detector input dose rate

IPEM Report 32 part II [4] indicates that the air-KERMA rate under ADRC should be in the range of 0.2 to 1.0 µGy s^{-1} for 23 or 25cm input field size. It goes on to state that for the maximum field size available the dose rate should not exceed 1.0 µGy s^{-1} and should never exceed 2.0 µGy s^{-1}. Jones et al [12] indicate that for an II of 30 cm nominal diameter at 30 frames per second, the DIDR should be 260 nGy s^{-1} and suggest corrections for pulse rate and field size for both II and FPD. They also note that for FPD there is a restriction on the lowest dose rate of 20 to 50nGy per pulse due to levels of electronic noise (this may vary with different generations of digital detector) and that the dose floor can be lower in the case of II due to lower amounts of electronic noise present.

Analysis of the DIDR of the image receptor under ADRC was complicated by the different measurement techniques employed with sites using a range of solid state and ionisation chambers which may or may not include back scattered radiation, different thicknesses of copper, a wide range of dose rate modes, anatomical programmes, pulse rates and anti-scatter grid state. It was necessary to group some data together by varying the range of field sizes to maintain sample sizes > 10 in each category.

The range of DIDR for all the 3499 measurements submitted (all phantoms, field sizes, pulse rates and modes) was 0.009 to 89.6 µGy s^{-1}, with a mean value of 0.87 µGy $s^{-1}$and an inter-quartile range of 0.27 to 0.74 µGy s^{-1}. The analysis for continuous fluoroscopy (including 25 and 30pulses per second (PPS)) and a 'normal' dose rate setting as specified by the manufacturer (includes terminologies such as normal, 1/1 and standard), low dose rate setting as specified by the manufacturer (includes terminologies such as ½ dose, low dose and reduced dose) and high dose rate setting as specified by the manufacturer (includes terminologies such as high level fluoroscopy

(HLF) and 'mA Boost') is given in table 12 and figure 6. Insufficient data was submitted using comparable parameters for fixed II systems to present values.

Table 12: Representative DIDR for fixed II, mobile II and fixed FPD for 1.5-2mm Cu attenuating medium, using a Continuous dose rate setting (including 25 and 30pps), grid in where the number of test results received (n) was ≥10

Analysis	Dose mode	Diagonal Field Size (cm)	n	Median (µGy s^{-1})	Range (µGy s^{-1})	Inter quartile range (µGy s^{-1})
Mobile Image Intensifiers	Normal, Continuous	23	11	0.37	0.29-0.68	0.32-0.43
		15	10	0.81	0.43-1.02	0.60-0.90
	Low, Continuous	23	27	0.17	0.09-0.36	0.15-0.21
		15	12	0.18	0.16-0.50	0.17-0.31
	High, Continuous	23	11	0.33	0.27-0.98	0.32-0.63
Flat Panel Detectors	Normal, Continuous	17 - 48	26	1.00	0.15-1.72	0.82-1.20
	Low, Continuous	30-32	12	0.64	0.12-1.94	0.14-0.84
		25	17	0.95	0.24-1.74	0.86-0.99
		20	38	0.56	0.23-2.03	0.50-0.60
		17	26	0.64	0.38-0.77	0.64-0.66
		15	30	0.73	0.43-1.63	0.70-0.79
		12	28	0.79	0.35-0.96	0.77-0.83

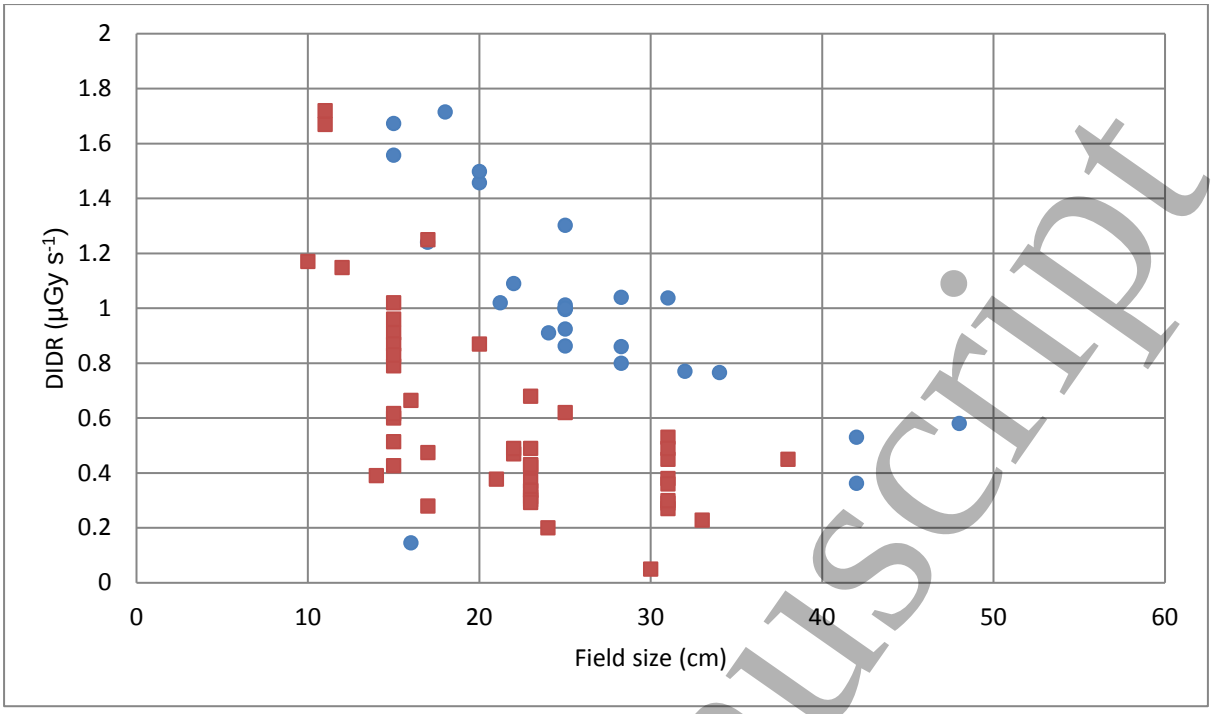

Figure 6: DIDR for different field sizes using 1.5mm of copper phantom at the x-ray tube, on continuous dose rate (25pps and 30pps accepted as continuous) on Normal (or Standard) fluoroscopy dose setting for FPD (blue circles) and II (red squares)

3.1.6 Collimated DIDR to the image receptor under ADRC

IPEM Report 32 part II [4] recommends that when collimating to an area of 7 x 7cm at the detector for the largest field size no excessive increase in DIDR should result. There is no remedial level suggested for this test, it is left as a task for the MPE to determine if an increase in dose is excessive. To maintain the approach used throughout for assigning a SR, a remedial level of ±25% of the measured DIDR for an uncollimated beam was used. This is in keeping with the remedial levels used for all tests involving DIDR.

Whilst 3 respondents indicated they perform this test routinely and at commissioning, testing data was only received from one site. The analysis is presented in table 13.

Table 13: Percentage increase in DIDR for a collimated beam relative to an uncollimated beam at the largest field size under ADRC

Analysis	n	Mean difference	3rd Quartile increase	Maximum increase
All units	134	4.6%	11.6%	35.5%
Image intensifier units	111	5.0%	11.6%	31.1%
Flat panel detector units	23	2.6%	10.3%	35.5%

The 3rd quartile result indicates a SR of 3. Owing to a lack of specified remedial limit in the guidance, and that data for this test was submitted from a single site it was considered that a LR was not appropriate.

3.1.7 Relationship between changes in ESDR of a phantom and DIDR at the image receptor under ADRC

Table 5 indicates that the majority of respondents test both ESDR and DIDR annually. There are many reasons to test both. ESDR data is useful for optimisation, and there are published [2,4,5] remedial and suspension levels for ESDR that sites can compare against. DIDR data can provide a useful indication of equipment function such as the performance of the ADRC and has a bearing on image quality tests. Both tests can also be undertaken to monitor changes in equipment performance over time through comparison to baseline measurements. The data were evaluated to determine if a significant change in one was reflected in the results of the other; if this were the case, it might only be necessary to routinely test one and a significant reduction in the testing could be possible.

For the most recent testing session to all units, the percentage variation from baseline measured for both the ESDR and DIDR tests was compared directly to see how closely they agreed. For 105 unique units, the results are shown in table 14.

Table 14: The percentage agreement between the measured deviation from baseline for ESDR and DIDR tests undertaken during the same instance of testing for all testing sessions to all units

	% agreement between the measured deviation from baseline of ESDR and DIDR						
	0 - 5%	5 – 10%	10 – 15%	15 – 20%	20 – 25%	25 – 50%	>50%
No. of unique units	21.9% (n=23)	21.0% (n=22)	8.6% (n=9)	6.7% (n=7)	9.5% (n=10)	24.8% (n=26)	7.6% (n=8)

Table 14 shows that 21.9% (23) of the total number of unique units had a percentile variation from baseline for DIDR and a percentile variation from baseline for ESDR that agreed with each other to within 5%, suggesting an apparent agreement in the change in dose rate from baseline. Conversely, table 14 also shows that 7.6% (8) of the total number of unique units had a percentile variation from baseline for DIDR and a percentile variation from baseline for ESDR that were > 50% different to each other, suggesting no agreement in the change in measured dose rate from baseline and that an acceptable result for one dosimetric quantity with respect to baseline is not indicative of an acceptable result for the other.

51.5% (54) of units demonstrate agreement between the measured deviations from baseline for the ESDR and DIDR tests within 15%. 48.5% of units demonstrated a deviation from baseline for the ESDR and DIDR tests that disagreed by more than 15%, with a deviation of more than 25% for just over 32% of units. This is likely due to differences in the way the dose is delivered to the detector due to different QA setup with different pulse rates, mA and kV curve and spectral filtration. This

means that potentially the same DIDR may be delivered in many ways. The recording of those parameters can aid with the detailed analysis of results.

The data indicates that the measurements of ESDR and DIDR are complementary to each other and comparison of the results can indicate which changes have occurred within the unit. It is important that both are retained as part of equipment QA.

3.1.8 kV accuracy

This test is described in the tube and generator sections of IPEM report 91[2] and report 32 I and II [3,4]. It is applicable to fluoroscopy units and can be measured with a suitable meter. The remedial level is the greater of nominal ±5% or ±5kV. The suspension level is the greater of nominal ±10% or ±10kV. The results are shown in table 15. Note that errors are given as absolute (kV) where the set kV for the test is < 100 and as a percentage where the set kV for the test is ≥ 100.

Table 15: The results for absolute and percentage deviations between measured and nominal kV

Analysis	n	Exceeding remedial level	Largest error	Mean result of those exceeding remedial	Third quartile result of those exceeding remedial	Exceeding suspension level
All testing sessions on all units	145	12.4%	10.7kV, 73%	7.0kV, 13.8%	7.3kV, 9.3%	2.1%
Most recent testing session on all unique units	76	11.8%	7.1kV, 73%	6.2kV, 20.6%	6.5kV, 9.8%	1.3%
All testing sessions on units with image intensifiers	69	15.9%	7.1kV, 9.8%	6.2kV, 6.5%	6.5kV, 6.8%	0%
All testing sessions on units with flat panel detectors	76	9.2%	10.7kV, 73%	9.3kV, 21%	10kV, 23.3%	4.0%
All testing sessions on fixed fluoroscopy units	50	2%	5.7%	-	-	0%
All testing sessions on interventional units	39	18%	10.7kV, 73%	9.3kV, 21.0%	10kV, 23.3%	7.7%

All testing sessions on mobile c-arm units	55	18.2%	7.1kV, 9.8%	6.12kV, 6.6%	6.5kV, 7.2%	0%
All commissioning testing sessions on all units	33	3%	5.4%	-	-	0%

Where no result is entered there is insufficient data to make a robust analysis.

There were no statistically significant results for kV below 100 (' test reason' (p=0.04), ' equipment age' (p=0.85), 'equipment manufacturer' (p=0.09), 'medical physics department' (p=0.63), 'equipment type' (p=0.88) and 'detector type' (p=0.43)), nor were there any statistically significant results for kV above 100 ('test reason' (p=0.96), 'equipment age' (p=0.63), 'equipment manufacturer' (p=0.81), 'medical physics department' (p=0.99), 'equipment type' (p=0.64) and 'detector type' (p=0.69)).

Based on the failure rate this test is given a LR of 4. A SR of 3 is applied based on the potential effect on image quality, especially with respect to image contrast. This SR is also relevant where the unit is capable of providing radiographic imaging such as in the case of mobile II and FPD units. This gives a risk matrix category of orange for this test indicating that there is enough evidence to suggest the test should remain or be adopted as part of routine testing.

As the remedial level for kV accuracy of fluoroscopy units is the same for radiographic units a comparison with the results from Honey et al [1] is made. For radiographic units, 6.4% exceeded the remedial level for tube potentials exceeding 100kV, and 9.6% of units exceeded the remedial level for tube potentials below 100kV. The failure rate of fluoroscopy units at the remedial level (table 15) is higher than for radiographic units, however both are in the orange category of the risk matrix. The difference may be related to the test methodology. Measuring kV on a fluoroscopy unit is more challenging than on a radiographic unit. Additional spectral filtration (employed by fluoroscopy units to reduce skin dose) such as copper can produce beam qualities outside the calibration conditions of kV meter resulting in an increased uncertainty in the measurement. This is discussed further in the NEMA standard [8]. It is for the MPE to consider the accuracy of the kV measurement for any given combination of fluoroscopy unit and kV meter.

3.1.9 KERMA Area Product (KAP) meter and reference air KERMA calibration at the PERP

KAP meter accuracy testing is not indicated in IPEM 91 [2]. For the purpose of this analysis a remedial level of ±25% of a correct value was assumed [13, 14]. However, the IEC standards [9,11] indicate acceptable performance to be within 35%. Reference air KERMA calibration accuracy is a test not described for fluoroscopy in IPEM 91 [2] or IPEM report 32 part II [4] since reference air KERMA was not widely available when either report was published. It is thought that most sites that have included the reference air KERMA calibration in their routine test protocol are working to the

same remedial level as adopted for the KAP meter: ±25%. Again, the IEC standards [9,11] indicate acceptable performance to be within 35%.

The results for both KAP and reference air KERMA calibration accuracy are highly dependent upon the measurement technique, which calls into question whether a correct value can be agreed upon. If the tests are not carried out in accordance with the manufacturer's protocol for setting up the unit the results are unlikely to agree. This can cause confusion and disagreement over whether the meters should be recalibrated. When interpreting the reported failure rates for KAP and reference air KERMA meter testing it is important to bear in mind that some will be the result of poor calibration and some the result of a difference in test technique. There is insufficient information to differentiate between the two for the data analysed as part of this work. The results for KAP meter calibration are shown in table 16; the results for reference air KERMA calibration are shown in table 17.

Table 16: the results for the accuracy of KAP meter calibration

Analysis	N	Exceeding remedial level	Largest failure	Mean difference of those passing	3rd quartile result of those exceeding 25%
For all testing sessions on all units	348	10%	55%	12%	38%
For all testing sessions - couch in	96	17%	55%	15%	40%
For all testing sessions - couch out	228	7%	51%	11%	42%
Fixed units (including interventional) – couch in	96	17%	55%	15%	39%
Fixed units (including interventional) – couch out	43	12%	42%	10%	42%
Mobile units – couch out	192	6%	51%	7%	41%

Table 17: the results for the accuracy of reference air KERMA calibration

Analysis	N	Exceeding remedial level	Largest failure	Mean result of those exceeding remedial	3rd quartile result of those exceeding 35%
All testing sessions on all units	13	23%	99%	65%	84%
Most recent testing session on all unique units	10	20%	69%	47%	59%
All commissioning testing sessions on all units	5	40%	99%	84%	92%

The overall failure rate of 9.5% for KAP meter calibration indicates a LR of 3. The SR is 4, as its inaccuracy could lead to the failure to see a need for optimisation or the implementation of inappropriate optimisation strategies (as identified in the tube and generator work [1]) which can

have a potential effect on population dose. This gives a risk matrix category of red indicating it should remain as part of routine testing.

The IEC standard [9] indicates that the KAP meter should be accurate to within ±35% without the table present. This is reflected in the results where units tested without the couch in the beam are less likely to exceed the remedial level used in this paper and have a lower mean difference. Mobile units, all of which were indicated to be tested without a couch present, had a lower failure rate and smaller difference from true. This may be indicative of confusion as to whether KAP meters should be calibrated with the couch in or out for fixed units, with many medical physics departments opting for couch in as being more typical of the clinical setup. The results suggest a need for guidance on test technique and for closer collaboration between MPEs and equipment manufacturers on calibration.

It must be taken into consideration that there was limited data for the results of the reference air KERMA calibration accuracy test. This test stands out as one that is regularly performed (13 sites indicated that it is performed to some degree, as shown in table 5) yet very little data was submitted. This could be indicative of the level of confidence placed in the results of this test at present. There was insufficient data to assign a LR. The SR is dependent upon how the reference air KERMA indicator is used locally. Where not used the SR is not applicable and the test does not need to be performed for that particular unit. Other sites may use it as part of a deterministic effect patient follow up protocol [15], in which case it is desirable to have an accurate reading. There is a risk to the individual in that a poorly calibrated reference air KERMA could lead to a failure to trigger the appropriate patient follow up protocol. As with KAP meters, there is a need for guidance on test technique and for closer collaboration between MPEs and equipment manufacturers on calibration.

Increased importance of the accuracy of the calibration of the reference air KERMA may be assumed if this was to be utilised for equipment QA. In accordance with the IEC standards [9,11] typical dose rates and doses at the PERP are being given in the literature accompanying fluoroscopic units.

3.1.10 Tube leakage

Whilst tube leakage is not indicated in IPEM 91 [2], it is a recommended test during the critical examination of equipment in IPEM Report 107 [16]. The remedial level is 1mGy hr^{-1} at 1m [17]. The analysis for the results that were submitted are shown in table 18.

Table 18: the highest absolute values and the percentage of units exceeding the remedial level for tube leakage

Analysis	Unique tubes	Exceeding remedial level	Greatest leakage measured (mGy hr^{-1} at 1m)
For all testing sessions on all units	52	0%	0.59
For commissioning/new tube/other testing sessions on all units	47	0%	0.59
For routine testing sessions on all units	5	0%	N/A

There were no reported instances of this test having failed. The LR is 1. Although there were no failures to inform the SR, SR is rated 5 due to the potential to significantly expose a member of staff for a prolonged period, or to irradiate members of the public where shielding is designed to attenuate scattered radiation only. This indicates a risk matrix category of orange. The test needs to remain as part of commissioning. The sample submitted for routine checks was too small to draw any conclusions with confidence. The equivalent test for tube and generator testing [1] had a 0% failure rate for 55 responses. The conclusion was that the test need only be retained for commissioning of new units and testing of new x-ray tubes.

3.1.11 Tube filtration

The ICRP recommend in publication 33 [18] that total filtration in the useful beam for normal diagnostic work shall be equivalent to not less than 2.5mm of aluminium equivalent and is echoed in the relevant IEC standard [9]. IPEM report 91[2] does not include a test of tube filtration but this test is included in the recommended content of a critical examination in IPEM report 107 [16]. It is worth noting that many dosemeters return an estimate of half value layer (HVL) and total filtration in terms of aluminium equivalent. It is likely that many sites undertaking this test routinely do so because the result is displayed alongside the other information on the dose meter during kV or output tests. In all, there were 103 results submitted for x-ray tube filtration. The breakdown of units for which tube filtration results were submitted is shown in table 19.

Table 19: The types of unit for which results for tube filtration were reported

	Unit type		
Detector type	Fixed fluoroscopy	Mobile c-arm	Mini c-arm
Flat panel detector	37	1	1
Image Intensifier	12	48	4
Total	49	49	5

Across all 103 testing sessions there were no failures recorded. This gives the test a LR of 1. The SR is 5 however as a tube filtration of less than 2.5mm aluminium equivalence would result in a significant increase in individual and population skin dose. It would also mean deviating from a longstanding ICRP recommendation [18] and a failure to comply with an IEC standard [9]. This

indicates a risk matrix category of orange. The test needs to remain as part of a critical examination, but there is no evidence to suggest it needs to form part of routine testing.

3.1.12 Fluoroscopy timer accuracy and automatic cut-out

Though not described in IPEM report 91[2] or IPEM report 32 [4, 5], a number of sites verify the accuracy of the fluoroscopy timer. It is important that the MPE understands what the unit is displaying in order to give a definitive analysis of the timer accuracy as some units may display a fluoroscopy time representing only the time the x-ray beam is on (i.e. taking the pulse rate and length into account). In this instance the measurement may be useful for characterising the unit, especially if the way fluoroscopy time is measured can be altered in the software. The remedial level assumed for this work is any deviation between measured and displayed fluoroscopy time that is in excess of the uncertainty of the measurement technique.

Data were received for only 8 testing sessions with no failures recorded and it is therefore not possible to assign a LR. The SR is 1 as it is of very limited impact. The fluoroscopy timer could be used to terminate a Cine run during an examination but it is not thought that an extended fluoroscopy time could give a patient any significant overexposure before being manually terminated by the operator.

Fluoroscopy generators meeting the requirements of the IEC standards [9] should be provided with an alarm that sounds after a user defined period and not exceeding 5 minutes of cumulative fluoroscopy. The standard [9] also specifies that exposure automatically terminates after 10 minutes of continuous fluoroscopy. 10 minutes was chosen as being beyond any likely clinical exposure. No failures in the automatic termination were recorded from the data of 22 testing sessions. It is difficult to suggest a SR; it is not thought that operators are solely reliant upon an auditory alarm to alert them to the fluoroscopy time. The termination after 10 minutes of exposure is intended as a safety mechanism in the event of accidental activation of the unit resulting in an extended exposure.

As the SR is most likely low, this test has a risk matrix category of green; there is no evidence to suggest inclusion in a routine test protocol. The Radiation Protection Adviser (RPA) may specify a test of the automatic cut out as part of a critical examination and this test is included in IPEM report 107 on critical examinations [16].

3.2 Fluorography results

IPEM Report 32 (VI) [5] was written at a time when a secondary system was in use for fluorography. This may have been spot-film imaging, or latterly a VHS or similar recording system. As such, a thorough evaluation of the fluorography side of the unit was required as components were largely separate from the fluoroscopic imaging chain. Modern units utilise the same imaging chain for both modes, and often fluorography can be considered to be a high dose mode. The term ‘acquisition mode’ may be more appropriate due to this mode often being available for the recording of images during a procedure, however with the improvements in technology and the decreasing cost of computer memory it is commonly possible to record sections of fluoroscopic imaging. The current IPEM guidance can therefore be considered significantly out of date and is an important area where updated guidance is required.

25 medical physics departments provided data relating to the frequency with which they undertake each fluorography test. The results are summarised in table 20.

Table 20: Frequency with which each fluorography test is undertaken by responding medical physics departments

Fluorography Tests	Type of Test	Indicated as never performed	Critical Exam only	Annual	Biennial
Dose per image at the input face of the image receptor under AEC	FLG04	-	-	24	1
Collimated dose per image at the input face of the image receptor under AEC	Additional	20	-	1	-
Dose per image at the entrance surface of a phantom under AEC	Additional	2	2	20	-
Collimated dose per image at the entrance surface of a phantom under AEC	Additional	19	-	2	-
DIDR acquisition runs	Additional	-	-	2	-
ESDR acquisition runs	Additional	-	-	2	-
kV accuracy	Additional	-	-	1	-
Tube output	Additional	-	-	3	-
Timer accuracy (for tube measurements)	Additional	-	-	1	-

Table 20 highlights some interesting points. One site chooses to duplicate kV testing using both fluoroscopy and fluorography. Where the same tube and generator is used this test does not need to be duplicated.

There is a wide variation in the dosimetry metric being tested. An analysis of the reported test methodology and results for dose per image at the input face of the image receptor under AEC and the dose per image at the entrance surface of a phantom under AEC shows variations in the reported metric including dose per image, total dose and dose rate from an acquisition run. A review of the data suggests the reported metric changes on a unit-to-unit basis in some cases. This could be a necessary change in technique in response to the capabilities of each individual unit, or to reflect how the unit is used clinically.

3.2.1 Dose per image at the entrance surface of a phantom and dose per image at the input face of the image receptor under ADRC

Dose per image at the entrance surface of a phantom under ADRC is not included in IPEM Report 91 [2] or Report 32 part VI [5], though it is clear from table 20 that most of the medical physics departments that submitted data undertake this test routinely. For those medical physics

departments contributing data, the attenuator choice and thickness was as used for ESDR testing under fluoroscopy, as summarised in table 6.

The IPEM 91 [2] remedial level for dose per image at the input face of the image receptor under ADRC is a change from baseline of 25% and suspension is a change from baseline of 50% or more. It is anticipated that medical physics departments contributing data for the dose per image at the entrance surface of a phantom under ADRC use the same remedial and suspension levels for this test and therefore these values were used for estimating the LR.

The analysis of the results for dose per image at the entrance surface of a phantom under ADRC are given in table 21 and for dose per image at the input face of the image receptor under ADRC in table 22. Both exclude commissioning measurements as these measurements form the baseline data for subsequent analysis.

Table 21: Analysis of the results of dose per image at the entrance surface of phantom under ADRC

Analysis	N	Exceeding remedial level	Mean result of those exceeding remedial	3rd quartile result of those exceeding remedial	Maximum deviation from baseline
All testing sessions on all units	139	29.5%	66.1%	67.2%	308%
Most recent testing session on all unique units	73	32.9%	61.5%	61.8%	217%
All testing sessions on all image intensifiers	60	36.7%	71.6%	71.1%	217%
all testing sessions on all flat panel detectors	79	24.1%	59.7%	60.6%	308%

Table 22: Analysis of the results of dose per image to the input face of the image receptor under ADRC

Analysis	n	Exceeding remedial level	Exceeding suspension level	Mean result of those exceeding remedial	3rd quartile result of those exceeding remedial	Maximum deviation from baseline
All testing sessions on all units	234	26.5%	9.4%	22.7%	26.6%	167.2%
Most recent testing session on all unique units	177	28.2%	11.2%	24.3%	27.4%	167.2%
All testing sessions on image intensifiers	145	33.8%	13.8%	27.1%	33.0%	167.2%
All testing sessions on all flat panel detectors	87	14.9%	2.3%	14.8%	20.1%	58%
All testing sessions on all fixed fluoroscopy units	63	12.7%	7.9%	16.3%	19.9%	68.7%
All testing sessions on all interventional units	47	19.1%	0.0%	15.1%	20.3%	45.9%
All testing sessions on all mobile c-arm units	118	36.4%	13.6%	28.5%	33.8%	167.2%
All testing sessions on all mini c-arms	5	40%	20%	-	-	-

For all testing sessions to all units, the failure rates for both tests received a LR of 5 assuming the remedial level is a change from baseline of 25% and the suspension level is a change from baseline of 50% or more. The SR was assigned based on the 3rd quartile result. It was considered that an increase in the dose per image would lead to an increased population dose, whereas a drop in dose per image would have a detrimental effect on image quality. For the dose per image at the entrance surface of phantom under ADRC test the SR is 5, for the dose per image to the input face of the image receptor under ADRC the SR is 4. This gives both tests a risk matrix category of red indicating the importance of remaining part of routine testing and the continued inclusion of dose per image monitoring as part of radiographer QA.

For dose per image at the entrance surface of phantom under ADRC there were no results found to be statistically significant ('equipment age' (p=0.40), 'equipment manufacturer' (p=0.69), 'medical physics department' (p=0.04), 'equipment type' (0.62) and 'detector type' (0.55)). For dose per image to the input face of the image receptor under ADRC there were no results found to be

statistically significant ('equipment age' (p=0.94), 'equipment manufacturer' (p=0.59), 'equipment type' (p=0.96) or 'detector type' (p=0.98)). For II units a statistically significant difference was found for 'medical physics department' (p=0.01) but not for FPD units (p=0.03). This could be related to differences in test technique and the larger number of II in the reported data.

As was the case with fluoroscopy, the high failure rate is not thought to be completely indicative of unit performance. It would seem prudent to only use the ±25% from baseline remedial level on units that have a medical physics or QA program set up on them, or a user quality control mode as described in NEMA XR-27 [8]. For all other programs s on these units and for units on which an unchanging QA program is not an option, representative doses would be more useful as an aid to optimisation. Insufficient consensus on the measurement method and results reported meant that sample sizes were too small to recommend representative ranges.

It is proposed that updated guidance on a relevant, comparative metric and a common technique to measure it (especially with regard to the manner in which a unit is permitted to identify the exposure factors prior to the acquisition) would be welcome and allow for a future exercise in comparing fluorographic exposures with a view to establishing representative ranges. This data would be extremely useful to the MPE when advising on optimisation.

It is possible that intentional reductions in dose rate following improvements to image processing will not affect the image quality. The working party is unable to differentiate between results exceeding the remedial level that are the result of intentional changes to the unit or not. The SR of 5 is assigned in the knowledge that the 3rd quartile result of those exceeding the remedial level could have had little or no effect, but to account for the possibility that it did. In either case, the ESDR and DIDR needs to be routinely tested so changes can be found and investigated.

4. Summary of results

A summary of the LR, SR, risk outcome and recommendations for each test is shown in table 23.

Table 23: A summary of the LR, SR, risk outcome and recommendations for all of the fluoroscopy and fluorography tests discussed throughout

Fluoroscopy						
Test	Reference	LR	SR	Risk outcome	Recommendation	Manuscript reference
ESDR at phantom under ADRC	IPEM 91, FLU06	5	5	Red	Retain in routine testing; include physics QA programme	3.1.2
DIDR under ADRC	IPEM 91, FLU07	5	5	Red	Retain in routine testing; include	3.1.2

					physics QA programme	
Variation with dose rate setting	-	-	-	-	Many instances of poorly set up units observed; include in routine testing	3.1.3
DIDR under ADRC for a collimated beam	IPEM 32(II), 1.2.2	-	3	Yellow – orange	Include at the MPE's discretion	3.1.6
ESDR at phantom under ADRC for a collimated beam	Additional	-	-	-	Include at the MPE's discretion	n/a
kV accuracy	IPEM 91, RAD12	4	3	Orange	Retain in routine testing	3.1.8
Accuracy of KAP meter calibration	BIR report 18	3	4	Red	Retain in routine testing	3.1.9
Accuracy of Reference air KERMA (rate) calibration	BS EN 60601-2-54: 2009	-	Dependent on use	Dependent on use	Include at the MPE's discretion	3.1.9
Tube leakage	IPEM report 107	1	5	Orange	Commissioning only	3.1.10
Tube filtration	ICRP publication 33	1	5	Orange	Commissioning only	3.1.11
Fluoroscopy timer accuracy	BS EN 60601-2-54: 2009	-	1	Green	Unnecessary as a routine test	3.1.12
Fluoroscopy automatic cut-out (10 minutes continuous)	BS EN 60601-2-54: 2009	-	-	Green	Unnecessary as a routine test	3.1.12
<u>Fluorography</u>						
Dose per image at the input face of the image receptor under AEC	IPEM 91, FLG04	5	4	Red	Retain in routine testing; include physics QA programme	3.2.1

Dose per image at the entrance surface of a phantom under AEC	Additional	5	5	Red	Retain in routine testing; include physics QA programme	3.2.1

5. Conclusions

The results presented offer a means of assessing the effectiveness of current testing regimes for the dosimetric and safety aspects of fluoroscopy and fluorographic units. The results highlighted significant differences between the testing regimes and methodology (in terms of phantoms used and metrics reported) employed by different UK based medical physics departments across the UK. This creates uncertainties when analysing trends and increases the width of typical ranges presented. The data therefore serves to highlight that rapid technological advancement has occurred from when the guidance was published and that there has been little consensus on how to adapt test methodologies in response. Despite these difficulties it was still possible to analyse and comment upon the efficacy of a range of tests, to present data on representative ranges for specific circumstances and to offer a commentary based on the results received that has in part been informed by the opinions expressed by a number of MPEs around the UK to provide an evidence base for the development of future guidance.

6. Acknowledgements

The authors would like to thank the Institute of Physics and Engineering in Medicine (IPEM) and the following people who provided some of the data described in this paper: Shaun Beggs (Bradford Teaching Hospitals NHS Foundation Trust), Sarah Cade (Royal United Hospitals Bath NHS Foundation Trust), Paul Colley (Brighton and Sussex University Hospitals NHS Trust), Philip Doyle (Belfast Health and Social Care Trust), Tony Hartley (Cambridge University Hospitals NHS Foundation Trust), Katharine Kenny (Oxford University Hospitals NHS Trust), Laurence King (The Royal Marsden NHS Foundation Trust), Sara Knapp (Gloucestershire Hospitals NHS Foundation Trust), Lera Köhler (NHS Highland), Louise Lindsay (NHS Greater Glasgow and Clyde), John Loveland (Imperial College Healthcare NHS Trust), Susan Maguire (Mater Private Hospital), Stephen McCallum (NHS Grampian), Peter McGookin (University Hospitals Birmingham NHS Foundation Trust), Kevin Munson (Nottingham University Hospitals NHS Trust), Kevin Robson (The Newcastle upon Tyne Hospitals NHS Foundation Trust), Jane Shekhdar (The Hillingdon Hospitals NHS Foundation Trust), Nick Weir (NHS Lothian) and Chris Wood (Northampton General Hospital NHS Trust).

References

[1] Ian Honey et al 2018 IPEM Topical Report: An evidence and risk assessment based analysis of the efficacy of tube and generator quality assurance tests on general x-ray units. *Phys. Med. Biol.* **63** 245011

[2] IPEM91 Recommended Standards for the Routine Performance Testing of Diagnostic X-ray Systems. IPEM Report No 91, York, IPEM (2005).

[3] IPEM32 Measurement of the performance characteristics of diagnostic X-ray systems: X-Ray Tubes and Generators. IPEM Report No 32 (Part I), York, IPEM (1996)

[4] IPEM32 Measurement of the performance characteristics of diagnostic X-ray systems: X-Ray Image Intensifier TV Systems. IPEM Report No 32 (Part II), York, IPEM (1996)

[5] IPEM32 Measurement of the performance characteristics of diagnostic X-ray systems: X-Ray Image Intensifier Fluorography Systems. IPEM Report No 32 (Part VI), York, IPEM (1995)

[6] Winter, B (2013) Linear models and linear mixed effects models in R with linguistic applications. arXiv:1308.5499

[7] Martin CJ, Sutton DG, Workman A, Shaw AJ and Temperton D. Protocol for measurement of patient entrance surface dose rates for fluoroscopic X-ray equipment. British Journal of Radiology 71 1283-1287 (1998).

[8] NEMA XR 27-2013 X-ray Equipment for Interventional Procedures User Quality Control Mode with Amendment 1, NEMA, 2013.

[9] BS EN 60601-2-54:2009 Medical electrical equipment – Part 2-54: Particular requirements for the basic safety and essential performance of X-ray equipment for radiography and radioscopy.

[10] IPEM Medical and Dental Guidance Notes, York, IPEM (2002)

[11] BS EN 60601-2-43:2010 Medical electrical equipment. Part 2-43: Particular requirements for basic safety and essential performance of X-ray equipment for interventional procedures.

[12] Jones, A. K., Balter, S., Rauch, P., Wagner, L. K. Medical imaging using ionizing radiation: Optimization of dose and image quality in fluoroscopy. Med. Phys. 41 (1) 2014

[13] Wall, B.F. Quality control of dose area product meters. (British Institute of Radiology Report 18: Technical and Physical Parameters for Quality Assurance in Medical Diagnostic Radiology) British Institute of Radiology, London, UK; 1989

[14] IEC, 1977. Area Exposure Product Meter, Publication 580, (Bureau Central de la Commission Electrotechnique Internationale, Genève)

[15] ICRP, 2000. Avoidance of radiation injuries from medical interventional procedures. ICRP Publication 85. Ann. ICRP 30 (2)

[16] IPEM The Critical Examination of X-ray Generating Equipment in Diagnostic Radiology, IPEM Report 107, York, IPEM (2012)

[17] BSI. Medical electrical equipment. Particular requirements for safety. Specification for x-ray source assemblies and x-ray tube assemblies for medical diagnosis. BS EN 60601-2-28, London, BSI (1993)

[18] ICRP, 1982. Protection against Ionizing Radiation from External Sources Used in Medicine. ICRP Publication 33. Ann. ICRP 9 (1)