

Superior Diagnostic Accuracy

with Additional and Stand-Alone

Digital Breast Tomosynthesis

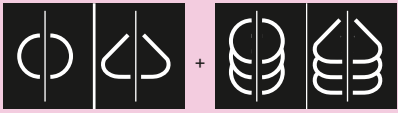



Key findings of the clinical Premarket Approval studies
of Mammomat Inspiration with Tomosynthesis Option

By Johannes Georg Korporaal, PhD, Nancy Nalleweg, PhD,
Thomas Mertelmeier, PhD, and Axel Hebecker, PhD

1 Introduction

This white paper is a summary of two clinical studies that supported the Premarket Approvals (PMA) of the “Mammomat Inspiration with Tomosynthesis Option” by the Food and Drug Administration (FDA) (1, 2)¹. Both clinical studies contributed to the assessment of the safety and effectiveness of the device for breast cancer screening and diagnosis in the US. The main difference between the two studies is how digital breast tomosynthesis (DBT) is used compared to full-field digital mammography (FFDM) alone (see Table 1).

Table 1
Study setups of the
two clinical studies

Adjunct scenario	Replacement scenario
<p>2-view FFDM plus 2-view DBT</p>  <p>vs.</p> <p>2-view FFDM</p> 	<p>Stand-alone 2-view DBT</p>  <p>vs.</p> <p>2-view FFDM</p> 

Both studies prove the superior diagnostic accuracy of Siemens DBT compared to full-field digital mammography (FFDM) alone. In chapter 2, the study design and populations for both studies are described and the key findings are covered in chapters 3–5.

¹ This white paper contains results from both clinical studies beyond those published in the FDA documents.

2 Objectives and methods

The goal of investigating two distinct scenarios was 1) to prove the value of wide-angle DBT in current clinical practice and 2) to demonstrate the potential of wide-angle DBT as a stand-alone modality.

In clinical practice, DBT is currently predominantly used as an **additional modality** to FFDM. Although the DBT slices provide more detail and can increase cancer detection (3–6), there are still clinical and traditional arguments for retaining FFDM with DBT. For example, FFDM images enable comparison with prior exams and give an overview of the whole breast. Further, reimbursement regulations may require that a 2D image is acquired together with DBT.

To demonstrate the potential of wide-angle tomosynthesis systems, Siemens also investigated the use of DBT as a **stand-alone modality**, as this would result in a lower total examination dose by eliminating the radiation dose of the FFDM.

For both scenarios, three objectives were addressed:

1. Superior diagnostic accuracy (Chapter 3)

The primary objective was to show diagnostic superiority of DBT with FFDM and stand-alone DBT compared to FFDM alone, as measured by the area under the receiver operating characteristic curve (AUC ROC) at the breast level.

2. Reduced non-cancer recall rate (Chapter 4)

As a secondary objective, the subject-level non-cancer recall rate was investigated. It was hypothesized that the recall rate would decrease because of higher diagnostic confidence with DBT.

3. Improved reader performance and lower interobserver variability (Chapter 5)

With DBT an improvement in the diagnostic performance of each individual reader was expected as well as a reduction in interobserver variability.

For both studies a multi-reader multi-case (MRMC) study design was chosen.

Methods

2.1 Library of images

Both studies were based on a library of images specifically collected for these DBT PMA studies. This case collection was performed between May 2011 and February 2014 from seven United States clinical sites:

- Duke University, Durham, NC;
- SUNY, Stonybrook, NY;
- Brigham & Women's Hospital, Boston, MA;
- Cleveland Clinic, Cleveland, OH;
- NYU, New York, NY;
- St. Luke's Episcopal Hospital, Houston, TX; and
- Miami Baptist Hospital, Miami, FL.

The library included FFDM images and DBT raw data as well as DBT images for presentation. All images were collected prospectively through a written informed consent process and all local institutional review boards approved the studies prior to enrollment of participants in the study.

The FFDM images were acquired according to the standard of care with various commercially available FFDM systems. In addition to these clinically indicated screening mammograms, DBT images were acquired solely on Siemens Inspiration DBT systems. All breasts were imaged under craniocaudal (CC) as well as mediolateral oblique (MLO) positioning.

For malignant lesions, the ground truth for the type and location of the lesions was 1) based on the mammography findings described by the radiologist at the clinical site according to the standard of care and 2) confirmed by the radiology and pathology reports after biopsy. For cases with a biopsy-confirmed benign finding, follow-up examinations were performed after 6 or 12 months. For normal cases, follow-up was performed after 12 months to confirm the non-cancer status.



Methods

2.2 Study populations

For each scenario, a subset of cases was randomly selected from the library of images (Table 2). Specific mammographic features like breast density and lesion type were taken into account to obtain a distribution of cases similar to that seen in clinical practice.



Adjunct scenario	Replacement scenario
	
<p>Subjects enrolled: 300</p> <ul style="list-style-type: none"> • 50 with cancer • 85 with benign findings • 165 normal 	<p>Subjects enrolled: 330</p> <ul style="list-style-type: none"> • 105 with cancer • 72 with benign findings • 138 normal • 15 special cases
<p>Breasts analyzed: 490</p> <ul style="list-style-type: none"> • 53 biopsy-proven cancerous breasts (10.8%) • 90 biopsy-proven benign breasts (18.4%) • 347 normal breasts with follow-up (70.8%) 	<p>Breasts analyzed: 548</p> <ul style="list-style-type: none"> • 110 biopsy-proven cancerous breasts (20.1%) • 77 biopsy-proven benign breasts (14.1%) • 361 normal breasts with follow-up (65.9%)
<p>Breast density:</p> <ul style="list-style-type: none"> • Almost entirely fat: 5.3% • Scattered fibroglandular: 40.7% • Heterogeneously dense: 48.7% • Extremely dense: 4.7% • Missing data: 0.7% 	<p>Breast density:</p> <ul style="list-style-type: none"> • Almost entirely fat: 5.8% • Scattered fibroglandular: 39.9% • Heterogeneously dense: 47.3% • Extremely dense: 6.4% • Missing data: 1.2%
<p>Lesion type:</p> <ul style="list-style-type: none"> • Mass: 65.2% • Calcification: 24.2% • Architectural distortion: 9.1% • Asymmetric density: 1.5% 	<p>Lesion type:</p> <ul style="list-style-type: none"> • Mass: 63.4% • Calcification: 21.6% • Architectural distortion: 9.0% • Asymmetric density: 6.0%
<p>Cancer type:</p> <ul style="list-style-type: none"> • Invasive cancer: 78.8% • Ductal carcinoma in situ: 21.2% 	<p>Cancer type:</p> <ul style="list-style-type: none"> • Invasive cancer: 80.6% • Ductal carcinoma in situ: 19.4%

Table 2

Subject characteristics

Methods

2.3 Reading and image interpretation

A total of 22 MQSA-qualified radiologists (“readers”) interpreted and scored the clinical images in the adjunct scenario and 31 MQSA-qualified radiologists in the replacement scenario. All radiologists were trained in the principles of breast tomosynthesis and the interpretation of DBT images.

In the adjunct scenario, the readers reviewed the cases sequentially in the following order:

1. FFDM (MLO and CC);
2. FFDM (MLO and CC) + DBT (MLO);
3. FFDM (MLO and CC) + DBT (MLO and CC).

In the replacement scenario, each reader participated in two reading sessions separated by a wash-out period of at least 4 weeks. During each reading session, half of the cases were interpreted with FFDM alone (MLO and CC). The other half of the cases were interpreted with DBT MLO views first and in a second step DBT with MLO and CC views. After the wash-out period, the readers interpreted the opposite modality for each case. Cases were randomized to be read with either FFDM or DBT first and the reading order within each group was randomly assigned.

Readers marked the locations of all suspected lesions on FFDM and DBT and assigned two types of scores for each lesion:

- a probability of malignancy confidence score on a scale from 0.5 to 100;
- a BI-RADS score based on the assumption that the lesion was the only finding (“forced” BI-RADS score)

The non-cancer recall rate was based on all subjects without cancer who received a false positive result (BI-RADS 3, 4 or 5) or a BI-RADS score of 0. Scoring for all subjects was done at the breast-level.



3 Significant improvement in diagnostic accuracy

The diagnostic accuracy in detecting and characterizing breast lesions can be described with the area under the receiver operating characteristic curve (AUC ROC) (see (7) for an explanation of AUC ROC). To summarize the diagnostic accuracy of all readers together, average ROC curves were calculated.

For both the adjunct and the replacement scenario, the average AUC ROC was significantly higher with DBT compared to FFDM alone (Table 3). The relative increase in diagnostic accuracy at the breast level was 13.4% for the adjunct scenario and 5.3% for the replacement scenario.

This demonstrates the **superior diagnostic accuracy** of Siemens' 50° wide-angle DBT in addition to FFDM, as well as for its stand-alone use.

Note: the AUC ROC values and the changes therein cannot be directly compared between the two studies, because the study populations had different sizes, included different cases and were read by different readers.

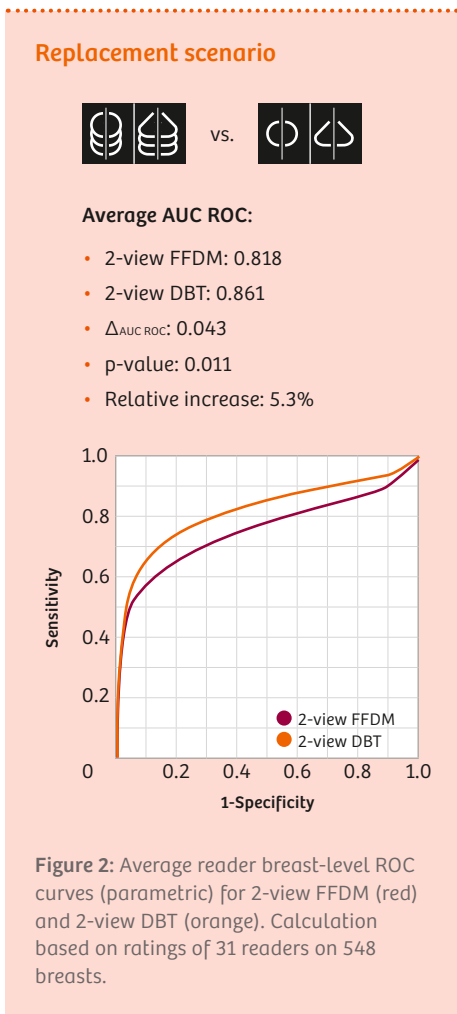
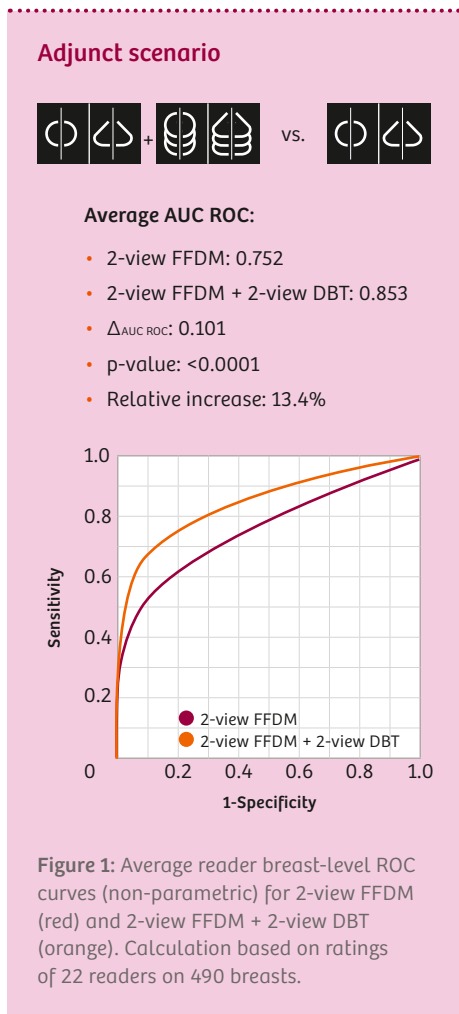


Table 3

Results: Significant improvement in diagnostic accuracy

4 Significant reduction in the non-cancer recall rate

Typically, all women with a positive or inconclusive finding during screening will be recalled for further diagnostic work-up. This group consists of true and false positives, the latter receiving unnecessary follow-up procedures such as ultrasound or biopsy. The non-cancer recall rate describes the number of women with false positive findings as a proportion of all screened women as illustrated in Figure 3.

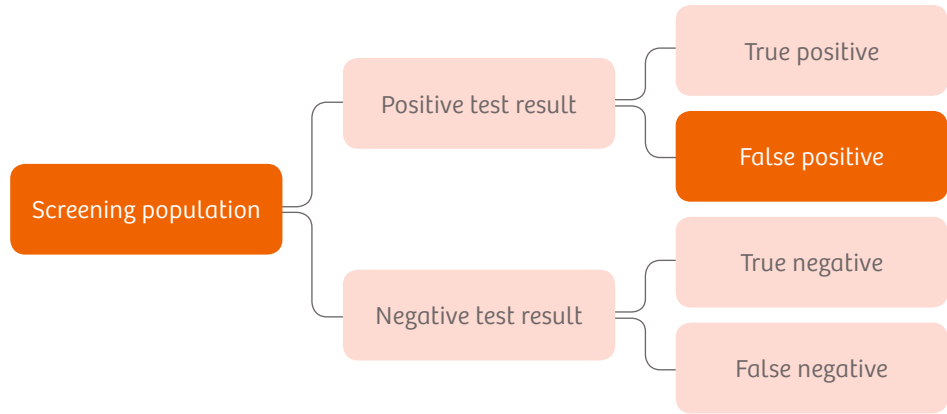


Figure 3:

The non-cancer recall rate is defined as the number of false positive cases as a proportion of all women screened.

For both the adjunct and the replacement scenario, the readers’ average non-cancer recall rate was significantly lower with DBT compared to FFDM alone (Table 4). In the adjunct scenario, for 18 out of 22 readers their non-cancer recall rate with DBT was lower than with FFDM alone. In the replacement scenario this was the case for 28 out of 31 readers.

This demonstrates the **increased confidence** with Siemens’ 50° wide-angle DBT in addition to FFDM, as well as for its stand-alone use.

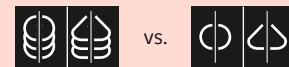
Adjunct scenario



Readers’ average non-cancer recall rate:

- 2-view FFDM: 0.438
- 2-view FFDM + 2-view DBT: 0.355
- $\Delta_{\text{non-cancer recall rate}}$: -0.083
- p-value: 0.0009
- Relative decrease: 18.9%

Replacement scenario



Readers’ average non-cancer recall rate:

- 2-view FFDM: 0.479
- 2-view DBT: 0.386
- $\Delta_{\text{non-cancer recall rate}}$: -0.093
- p-value: 0.0001
- Relative decrease: 19.4%

Table 4:

Results: Significant reduction in the non-cancer recall rate

Note: the reported non-cancer recall rates differ in magnitude from those reported in general population screening studies due to the enriched study populations of both studies (Table 2).



Results

5 Improved reader performance and lower interobserver variability

As shown in chapter 3, the average AUC ROC increased significantly with DBT in both scenarios for all radiologists together. In addition, the performance of each individual reader is also of interest.

In the adjunct scenario, all 22 readers showed improvement in their diagnostic accuracy with the addition of 2-view DBT (Figure 4) and weaker-performing readers showed more improvement compared to readers who were already stronger with 2-view FFDM alone. The differences in AUC ROC values between readers, as described with the coefficient of variation (Table 5, calculated post-hoc), were smaller with the addition of 2-view DBT, indicating lower interobserver variability.

In the replacement scenario, 29 out of 31 readers showed improvement in their diagnostic accuracy with 2-view DBT alone compared to 2-view FFDM alone (Figure 5). The coefficient of variation in AUC ROC values was also smaller with 2-view DBT compared to 2-view FFDM (Table 5, calculated post-hoc), indicating lower interobserver variability.

This demonstrates the **improved reader performance and lower interobserver variability** with Siemens' 50° wide-angle DBT in addition to FFDM, as well as for its stand-alone use.

Note: the AUC ROC values and the changes therein cannot be directly compared between the two studies, because the study populations had different sizes, included different cases and were read by different readers.

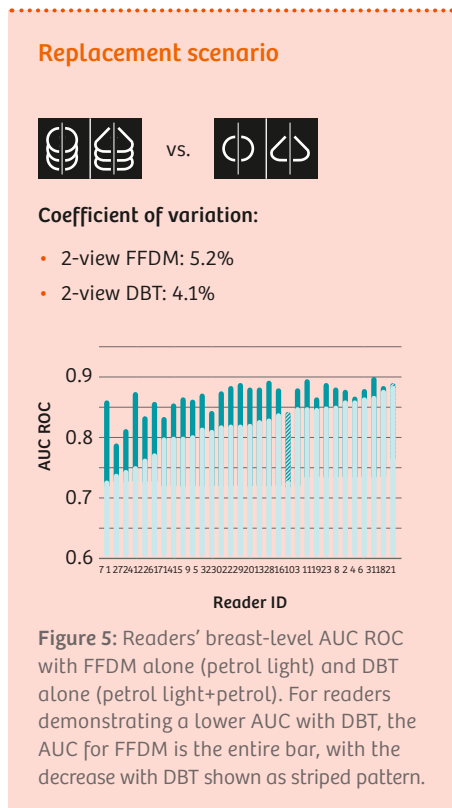
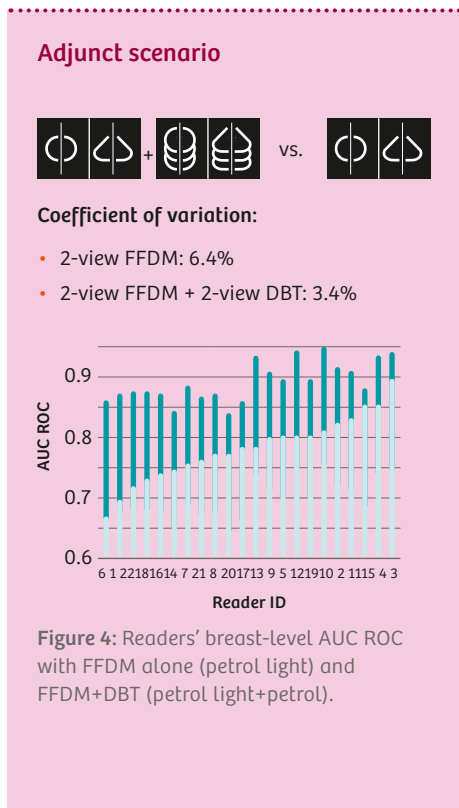


Table 5

Results: Improved reader performance and lower interobserver variability

6 Discussion

Digital breast tomosynthesis (DBT) is a promising technique that might replace FFDM as the gold standard for breast cancer screening and diagnosis. With Siemens' wide-angle DBT, clear advantages compared to FFDM have been proven with the two PMA studies described in this white paper:

- superior diagnostic accuracy,
- a reduced non-cancer recall rate,
- improved reader performance, and
- lower interobserver variability.

All of these improvements are highly relevant to the clinical implementation and are strong drivers for DBT to become the standard of care in screening and diagnostic mammography.

Radiation dose is especially important in breast cancer screening, because the majority of women undergoing screening examinations are healthy. Our replacement study clearly shows that with stand-alone wide-angle DBT the acquisition and radiation dose of FFDM is not necessary to gain the advantages of superior diagnostic accuracy, a reduced non-cancer recall rate, improved reader performance and lower interobserver variability that DBT can offer.

When using DBT as a stand-alone modality, a synthetic mammogram can potentially be of help for the comparison with prior exams and to get a fast overview of the entire breast. Note that synthetic mammograms were not used in either of these studies and in particular, the results of the replacement scenario were achieved without the readers referring to a 2D image when reading the DBT.

Further evidence supporting the use of stand-alone DBT in breast cancer screening comes from the Malmö Breast Tomosynthesis Screening Trial (3). In this trial, Siemens' wide-angle DBT was used as a stand-alone screening modality and examinations were performed with one view (MLO) only, to further reduce the radiation dose and reading time. The interim analysis shows an increased cancer detection rate of more than 40% and a significant increase in sensitivity. Even though the Malmö study used the Siemens wide-angle DBT system with one view only, it delivers further evidence that stand-alone DBT might be feasible as a future screening modality for the general population (8).

Similar to the Malmö study, the Siemens PMA study with the adjunct scenario also included an investigation of one-view DBT (MLO) in addition to FFDM. More details can be found in a separate white paper (7) or in the FDA Summary of Safety and Effectiveness Data (1).

Note that while Siemens performs and supports studies investigating the use of DBT in various protocols, these are not recommendations and a decision on the use of DBT as an additional or stand-alone modality and on the number of views must be made by the examining physician.

7 Summary

This white paper describes two clinical studies with Siemens' wide-angle tomosynthesis system which supported the Premarket Approvals by the Food and Drug Administration (FDA). DBT was used in addition to FFDM (adjunct scenario) and as a stand-alone modality (replacement scenario) and compared with FFDM. For the adjunct as well as the replacement scenario, superior diagnostic accuracy, a reduced non-cancer recall rate, improved reader performance, and lower interobserver variability were reported compared to FFDM alone. These results show that Siemens' wide-angle tomosynthesis system can be used with a high level of confidence in screening and diagnostic mammography.

Glossary

AUC ROC	Area under the ROC curve
BI-RADS	Breast Imaging Reporting and Data System
CC	Craniocaudal
DBT	Digital Breast Tomosynthesis
FDA	Food and Drug Administration
FFDM	Full-Field Digital Mammography
MLO	Mediolateral oblique
MQSA	Mammography Quality Standards Act and Program
PMA	Premarket Approval
ROC	Receiver Operating Characteristic

References

1. Siemens Medical Solutions USA Inc. (2015) PMA P140011: FDA Summary of Safety and Effectiveness Data.
2. Siemens Medical Solutions USA Inc. (2016) Clinical Study Report for Protocol SMS-SP09-01: Retrospective, multi-reader, multi-case study (MRMC) to demonstrate the superior accuracy of Siemens Digital Breast Tomosynthesis to Full-Field Digital Mammography (FFDM) as a replacement for FFDM screening mammography.
3. Lång K, Andersson I, Rosso A, Tingberg A, Timberg P, Zackrisson S (2016) Performance of one-view breast tomosynthesis as a stand-alone breast cancer screening modality: results from the Malmo Breast Tomosynthesis Screening Trial, a population-based study. *European radiology* 26(1):184–90.
4. Skaane P, Bandos AI, Gullien R, Eben EB, Ekseth U, Haakenaasen U et al. (2013) Comparison of digital mammography alone and digital mammography plus tomosynthesis in a population-based screening program. *Radiology* 267(1):47–56.
5. Ciatto S, Houssami N, Bernardi D, Caumo F, Pellegrini M, Brunelli S et al. (2013) Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study. *The Lancet. Oncology* 14(7):583–9.
6. Friedewald SM, Rafferty EA, Rose SL, Durand MA, Plecha DM, Greenberg JS et al. (2014) Breast cancer screening using tomosynthesis in combination with digital mammography. *JAMA* 311(24):2499–507.
7. Nalleweg N, Mertelmeier T, Korporaal JG, Hebecker A (2016) Superior Diagnostic Accuracy with Additional Digital Breast Tomosynthesis: Main findings of the clinical study for Premarket Approval of Mammomat Inspiration with Tomosynthesis Option. Siemens White Paper.
8. Sardaneli F, Aase HS, Alvarez M, Azavedo E, Baarslag HJ, Balleyguier C et al. (2016) Position paper on screening for breast cancer by the European Society of Breast Imaging (EUSOBI) and 30 national breast radiology bodies from Austria, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Israel, Lithuania, Moldova, The Netherlands, Norway, Poland, Portugal, Romania, Serbia, Slovakia, Spain, Sweden, Switzerland and Turkey. *European radiology*. doi: 10.1007/s00330-016-4612-z.

On account of certain regional limitations of sales rights and service availability, we cannot guarantee that all products included in this brochure are available through the Siemens sales organization worldwide. Availability and packaging may vary by country and are subject to change without prior notice.

Some / All of the features and products described herein may not be available in the United States or other countries. The information in this document contains general technical descriptions of specifications and options as well as standard and optional features that do not always have to be present in individual cases.

Siemens reserves the right to modify the design, packaging, specifications and options described herein without prior notice.

Please contact your local Siemens sales representative for the most current information.

In the interest of complying with legal requirements concerning the environmental compatibility of our products (protection of natural resources and waste conservation), we recycle certain components.

Using the same extensive quality assurance measures as for factory-new components, we guarantee the quality of these recycled components.

Note: Any technical data contained in this document may vary within defined tolerances. Original images always lose a certain amount of detail when reproduced.