

目次

1.	背景及び概観.....	4
1.1	緒言	4
1.1.1	骨巨細胞腫.....	4
1.1.2	骨巨細胞腫におけるデノスマブの作用	4
1.2	骨巨細胞腫を対象とした臨床開発計画の概要.....	5
1.2.1	骨巨細胞腫を対象とした臨床試験	6
1.2.2	客観的抗腫瘍効果の解析	13
1.2.2.1	客観的抗腫瘍効果の評価方法	13
1.2.2.2	客観的抗腫瘍効果の解析における有効性評価項目	16
1.2.2.3	客観的抗腫瘍効果の解析における統計手法.....	17
1.3	有効性評価デザインの概要	20
2.	個々の試験結果の要約	21
2.1	試験 20040215（外国第Ⅱ相臨床試験）	21
2.2	試験 20062004（外国第Ⅱ相臨床試験）	22
2.3	試験 AMG162-B-J201（国内第Ⅱ相臨床試験）：6 ヶ月カットオフ解析.....	24
2.4	試験 AMG162-B-J201（国内第Ⅱ相臨床試験）：12 ヶ月カットオフ解析.....	26
3.	全試験を通しての結果の比較と解析	28
3.1	試験対象集団	28
3.1.1	客観的抗腫瘍効果の解析対象集団	28
3.1.2	客観的抗腫瘍効果の解析における被験者の内訳	30
3.1.3	客観的抗腫瘍効果の解析における人口統計学的特徴及びベースラインの疾患特性	33
3.1.4	客観的抗腫瘍効果の解析対象集団におけるデノスマブの曝露状況	39
3.2	客観的抗腫瘍効果の解析結果	39
3.2.1	有効性主要評価項目：客観的奏効.....	39
3.2.2	有効性副次評価項目	46
3.2.2.1	客観的奏効の持続期間	46
3.2.2.2	客観的奏効までの期間	49
3.2.2.3	客観的奏効の持続性	52
3.2.2.4	腫瘍コントロールの持続性	54
3.2.2.5	客観的奏効の経時推移	56
3.2.3	有効性探索的評価項目	57
3.2.3.1	あらゆる評価基準を用いた最良効果別の被験者の割合	57
3.2.3.2	病勢の進行までの期間	58

3.2.4	探索的評価項目：客観的奏効を認めた被験者と認めなかった被験者における痛みの転帰、鎮痛薬の使用、及び臨床的有益性.....	58
3.2.4.1	客観的奏効を認めた被験者と認めなかった被験者における BPI-SF の痛みスコア	58
3.2.4.2	客観的奏効を認めた被験者と認めなかった被験者における鎮痛薬使用の変化.....	61
3.2.4.3	客観的奏効を認めた被験者と認めなかった被験者における臨床的有益性.....	61
3.2.4.4	痛みの転帰、鎮痛薬の使用、及び臨床的有益性に関する探索的解析の要約.....	61
3.2.5	画像コントロール群における客観的抗腫瘍効果の解析結果	61
3.3	外科的に切除可能な GCTB 患者に対するデノスマブの有効性	63
3.3.1	被験者の背景.....	64
3.3.2	外科的に切除可能な GCTB 患者での手術施行状況.....	64
3.3.3	手術までの期間.....	65
3.3.4	外科的に切除可能な GCTB 患者での臨床的有益性.....	65
3.4	個々の試験ごとの結果	65
3.5	部分集団における結果の比較	68
3.6	全試験を通しての結果の概要	68
4.	推奨用法・用量に関する臨床情報の解析	71
5.	効果の持続、耐薬性.....	73
6.	参考文献	74
7.	付録	76

略語一覧

略語	略していない表現（英）	略していない表現（日）
BPI-SF	Brief Pain Inventory – Short Form	簡易疼痛調査用紙（縮小版）
BSAP	bone-specific alkaline phosphatase	骨型アルカリフォスファターゼ
CI	confidence interval	信頼区間
CR	complete response	完全奏効
CT	computed tomography;	コンピュータ断層撮影（法）
CTX1	C-telopeptide-1	C-テロペプチド-I
ECOG	Eastern Cooperative Oncology Group	－
EORTC	European Organisation for Research and Treatment of Cancer	－
EU	European Union	欧州連合
FDA	United States Food and Drug Administration	米国食品医薬品局
¹⁸ FDG-PET	fluorodeoxyglucose positron emission tomography	フルオロデオキシグルコース陽電子放射型断層撮影（法）
GCTB	giant cell tumor of bone	骨巨細胞腫
IEC	Independent Ethics Committee	独立倫理委員会
IRB	Institutional Review Board	治験審査委員会
MRI	magnetic resonance imaging	磁気共鳴画像法
MPA	Medical Products Agency	-
NTX (uNTX)	N-telopeptide (urinary NTX)	N-テロペプチド（尿中 NTX）
uNTX/Cr	uNTX corrected for urine creatinine	クレアチニン補正尿中 NTX
ONJ	osteonecrosis of the jaw	顎骨壊死
PD	progressive disease	病勢の進行
PEI	Paul-Ehrlich-Institut	-
PET	positron emission tomography	陽電子放射型断層撮影（法）
PR	partial response	部分奏効
PRO	patient-reported outcome	患者の報告による転帰
Q4W	every 4 weeks	4 週間に 1 回
RANKL	RANK ligand	RANK リガンド
RECIST	Response Evaluation Criteria in Solid Tumors	固形癌の効果判定基準
sBLA	supplemental Biologics License Application	-
SD	stable disease	安定
SLD	sum of the longest diameter	最長径の和
sNDS	supplemental New Drug Submission	-
SRE	skeletal-related event(s)	骨関連事象
SUV _{bw}	Standardized Uptake Value by body weight	Standardized Uptake Value (FDG の集積を表す指標。PET カメラで測定した集積放射能〔減衰補正後〕を体重 1 kg あたりの投与放射能で割った値)
SUV _{max}	maximum Standardized Uptake Value	Standardized Uptake Value の最大値
TRAP 5b	tartrate-resistant acid phosphatase 5b	酒石酸抵抗性酸性フォスファターゼ 5b
UE	unable to evaluate	評価不能

1. 背景及び概観

1.1 緒言

1.1.1 骨巨細胞腫

骨巨細胞腫 (giant cell tumor of bone: GCTB) は極めてまれな原発性骨腫瘍であり、主に長骨の骨幹端部・骨端部、又は脊椎もしくは仙骨に偏心性の溶骨性病変として発現する。本疾患患者の多くは、20代～30代の若年成人である。小児でのGCTBの報告はごくまれであり、未成年の場合は通常、骨格が成熟した後（すなわち、骨端成長板の閉鎖後）に発現する。米国、欧州連合 (European Union: EU)、カナダ、及びオーストラリアで新たに診断される症例は、それぞれ毎年約800名、800名、80名、及び30名である。国内における発症報告数は、1972年～2003年で2126名（電子化集計前）、2006年～2008年で415名（電子化集計）であり、1972年～2008年（2004年及び2005年を除く）で合計2541名が報告されている。2008年の報告は単年で142名であった（[全国骨腫瘍登録一覧表, 2008](#)）。

GCTBは、急速な増殖、重度の骨破壊、及び周辺軟部組織への広がりの特徴とする。未治療のまま放置するとGCTBは進行し、罹患骨を完全に破壊して巨大腫瘍を形成し、その結果、患者は肉体的変形に至り、場合によっては四肢を失う可能性がある。したがって、治療の目的は、腫瘍を除去し症状を軽減すること、隣接関節や周囲の解剖学的構造を保存すること、並びに罹患した四肢や脊椎の機能を回復させることである。手術では、腫瘍が適切に切除できれば根治が可能である（[Civista Health Library, 2011](#)、[Malawer et al, 2011](#)、[Singer et al, 2011](#)）。しかし、侵襲性の高い外科的処置（関節切除、関節／人工関節置換術、切断、片側骨盤切断）や広範な切除が必要となる場合があり、これらはいずれも重大な後遺症が残るおそれがある。原発腫瘍を完全に一塊切除しなかった場合、再発の頻度は高く（[Becker et al, 2008](#)）、一塊切除後の再発リスクは10%～20%、搔爬術後の再発は40%～75%と推定されている（[Malawer et al, 2005](#)）。腫瘍の再発の大部分は、外科的手術後2年以内に認められ（[Yasko, 2002](#)）、一度、外科的切除を行った場合、再発後の外科的治療には限界が生じる。再発したGCTB患者では、最大6%で転移が認められる（[Szendrői, 2004](#)）。

病変の位置や程度によっては、外科的切除が不可能な場合が多い。根治切除が不可能な患者での治療目標は、腫瘍の進行を遅らせるとともに、生活の質を維持することである。

切除不能GCTBの治療として、塞栓術、放射線療法、化学療法、ビスフォスフォネート投与、及びその他の薬剤の投与などが行われている（[Caudell et al, 2003](#)、[Leggon et al, 2004](#)、[Malawer et al, 2011](#)、[Singer et al, 2011](#)）。しかし、これらによる治療の転帰（腫瘍縮小効果、疾患コントロール期間など）に関するデータは少なく、いずれの治療法でも持続的効果は示されていないため、GCTBに対する新たな治療選択肢が必要とされている。

1.1.2 骨巨細胞腫におけるデノスマブの作用

GCTBはRANK リガンド (RANK ligand: RANKL) を産生しており、GCTBの増殖はRANKLに依存している（[Szendrői et al, 2003](#)）。GCTBは、顕著な血管新生を伴う著明な溶骨性病変を

呈し、辺縁は不明瞭で移行域が広い。GCTB 病変は、組織学的には RANK を発現する破骨細胞様巨細胞及びその前駆細胞、並びに破骨細胞の活性化に必須なメディエータである RANKL を発現する単核間質細胞で構成されている。RANKL を発現する単核間質細胞は、GCTB における腫瘍病変を構成するとともに、腫瘍組織で著明な溶骨活性を示す破骨細胞様巨細胞を動員すると考えられている (Branstetter et al, 2012、Thomas et al, 2010)。

デノスマブは RANKL に対するヒト型モノクローナル IgG2 抗体であり、高い親和性 ($K_d: 3 \times 10^{-12}$ M) で可溶性及び細胞膜結合型のヒト RANKL に特異的に結合する (Kostenuik et al, 2009)。デノスマブが RANKL に結合すると RANK の活性化が阻害され、破骨細胞の形成、機能、及び生存が抑制される (Kostenuik et al, 2009)。GCTB 患者では、腫瘍の間質細胞に発現している RANKL がデノスマブによって阻害されることによって、破骨細胞様の腫瘍関連巨細胞が消失し、骨溶解と巨細胞腫瘍の進行が抑制される。増殖性間質に代わり、非増殖性で分化型の密度の高い新たな線維性骨が形成され、その結果として臨床転帰が改善する。

RANKL 阻害薬としての特異的な作用機序を考慮すると、デノスマブによる治療は GCTB の標的療法と認識され、アンメットメディカルニーズに対する重要な治療選択肢となる可能性がある。

1.2 骨巨細胞腫を対象とした臨床開発計画の概要

アムジェン社は、GCTB を対象として 2 つの国際共同、非盲検、単一群、第 II 相試験 (試験 20040215 及び 20062004) を実施した。これらの試験データを用いて、中央読影機関による客観的抗腫瘍効果のレトロスペクティブな評価を行い、デノスマブの GCTB に対する有効性を検討した。放射線画像を用いた本評価は、米国食品医薬品局 (United States Food and Drug Administration: FDA) と協議の上合意を得て実施した (20 年 月 日 会議議事録、20 年 月 日 会議議事録)。なお、本評価方法については、カナダ保健省 (20 年 月 日 会議、Control No.) 及び EU 報告者 (20 年 月 日の 会議; 20 年 月 日の 会議) とも協議の上合意を得ている。

国内では第一三共株式会社が、試験 AMG162-B-J201 (非盲検、単一群、第 II 相試験) を実施中である。本試験では、試験 20040215 及び 20062004 と同じ用法・用量を用い、FDA と合意した評価基準による客観的抗腫瘍効果のプロスペクティブな評価を行った。本申請資料では、6 ヶ月カットオフ解析 (データカットオフ日: 20 年 月 日) 及び 12 ヶ月カットオフ解析 (データカットオフ日: 20 年 月 日) の結果をまとめた。

本申請では、試験 AMG162-B-J201、20040215、及び 20062004 の 3 試験を評価資料とした。GCTB 患者に対するデノスマブの有効性は、主に中央読影機関による客観的抗腫瘍効果をもとに評価した。試験 AMG162-B-J201 での客観的抗腫瘍効果 (6 ヶ月カットオフ解析結果) と試験 20040215 及び 20062004 の外国 2 試験併合解析 (以下、外国 2 試験併合解析) による客観的抗腫瘍効果を示し、試験 AMG162-B-J201 と外国 2 試験併合解析で客観的抗腫瘍効果が一貫し

ていること示す。また、上記3試験併合解析（以下、3試験併合解析）による客観的抗腫瘍効果及び試験 AMG162-B-J201 の12ヵ月カットオフ解析の結果も示す。さらに、客観的抗腫瘍効果に関する結果以外の各試験の重要な有効性評価項目の結果もまとめた。

1.2.1 骨巨細胞腫を対象とした臨床試験

試験 20040215 及び 20062004（コホート1）では切除不能 GCTB 患者を対象とした。また、試験 20062004（コホート2）では切除可能 GCTB 患者を対象に、デノスマブの治療が、手術の代替選択肢となり得るか、あるいは比較的低侵襲性の外科的処置により手術の後遺症を低減できるかどうかを探索的に評価した。国内第II相臨床試験である AMG162-B-J201 では、切除不能 GCTB 患者及び切除可能 GCTB 患者の両方を対象とした。これら GCTB 患者を対象とした国内外の第II相臨床試験のデザインを表 2.7.3.1-1 に要約する。

試験 20040215（外国第II相臨床試験）

試験 20040215 は、再発又は切除不能な成人の GCTB 患者を対象とした非盲検、単一群、第II相試験であり、病理組織学的検査（巨細胞の消失）又は画像検査（治験責任医師の評価で標的病変の進行を示す所見なし）により、デノスマブの抗腫瘍効果を評価することを目的とした。また、デノスマブの安全性を評価した。本試験では、デノスマブ 120 mg の4週間に1回（Q4W）の皮下投与に、第8日及び第15日の120 mg 負荷投与を加えた投与スケジュールとした。デノスマブ投与は、腫瘍の完全切除、病勢の進行、治験責任医師又は治験依頼者による中止の判断、被験者による中止の申し出、又はビスフォスフォネート製剤、カルシトニンもしくはインターフェロン α -2a 投与のいずれかに至るまで継続した。ただし、デノスマブ投与が有効であった被験者には、再発した場合などにデノスマブの再投与を可能とした。安全性データは試験期間を通じて収集し、デノスマブ最終投与後2年まで6ヵ月間隔で収集した。本試験終了時点で本試験に参加していた被験者は試験 20062004 に移行した。

試験 20040215 の主要解析（データカットオフ日: 20██年██月██日）の結果及び最終解析（全試験期間及び2年間の安全性追跡調査における安全性評価）の結果について、それぞれ治験総括報告書にまとめ、本申請資料に含めた。

試験 20062004（外国第II相臨床試験）

試験 20062004 は、成人及び骨格が成熟した未成年の GCTB 患者を対象とした進行中の非盲検、単一群、第II相試験であり、デノスマブの安全性、並びに治験責任医師が判定した抗腫瘍効果を評価するためにデザインされた。GCTB は重篤であり治療選択肢が限られていることから、骨格が成熟した未成年の GCTB 患者も本試験に参加できることとした。未成年の被験者を組み入れる場合は、放射線撮影で1つ以上の長骨が成熟していること（例: 閉鎖した骨端成長板を有する上腕骨）を確認した。試験 20040215 と同様、デノスマブ 120 mg Q4W の皮下投与に、第8日及び第15日の120 mg 負荷投与を加えた投与スケジュールとした。本試験には、以下の3つのコホートを組み入れた。

- ・ コホート 1: 切除不能な GCTB 患者（例: 仙骨又は脊椎の GCTB、あるいは肺転移などの多発性病巣を有する患者）
- ・ コホート 2: 切除可能な GCTB を有し、重度の後遺症を残す手術（例: 関節切除、手足の切断、又は片側骨盤切断）の実施が試験中に予定されている患者
- ・ コホート 3: 試験 20040215 から移行した被験者

コホート 2 では、デノスマブ投与の抗腫瘍効果及び安全性に加え、手術の種類及び実施の有無を集計し、手術可能な被験者に対するデノスマブの有効性を評価した。コホート 2 に組み入れられた切除可能な被験者の、ベースライン及び試験中に必要な外科的処置は治験責任医師が判断した。治験責任医師は、ベースライン時点で計画し症例報告書に記載した選択肢の中から、試験中に実施する手術を選択した。

デノスマブ投与は、コホート 2 の被験者では腫瘍の完全切除後 6 回まで継続した。他のすべての被験者では、病勢の進行の確認、治験責任医師又は治験依頼者による中止の判断、被験者による中止の申し出、治験責任医師による臨床的有益性なしの判断、あるいは併用禁止療法の実施のいずれかに至るまでデノスマブ投与を継続した。なお、デノスマブ投与が有効であった被験者には、再発した場合などにデノスマブの再投与を可能とした。

事前に計画した試験 20062004 の第 3 回中間解析（データカットオフ日: 2011 年 3 月 25 日）の結果を治験総括報告書としてまとめ、本申請資料に含めた。

試験 AMG162-B-J201（国内第 II 相臨床試験）

試験 AMG162-B-J201 は、成人又は骨格が成熟した未成年の GCTB 患者を対象とした進行中の非盲検、単一群、第 II 相試験であり、デノスマブの客観的抗腫瘍効果をプロスペクティブに、中央読影機関で評価するためにデザインされた。また、本試験ではデノスマブの薬物動態及び安全性を評価した。未成年の被験者を組み入れる場合は、放射線撮影で 1 つ以上の長骨が成熟していること（例: 閉鎖した骨端成長板を有する上腕骨）を確認した。試験 20040215 及び 20062004 と同様、デノスマブ 120 mg Q4W の皮下投与に、第 8 日及び第 15 日の 120 mg 負荷投与を加えた投与スケジュールとした。腫瘍を完全切除した被験者でのデノスマブ投与は、完全奏効（complete response: CR）又は部分奏効（partial response: PR）を病理学的に確認後 6 回まで継続した。他のすべての被験者では、病勢の進行の確認、治験責任医師又は治験依頼者による中止の判断、被験者による中止の申し出、治験責任医師による臨床的有益性なしの判断、あるいは併用禁止療法の実施のいずれかに至るまでデノスマブ投与を継続した。なお、デノスマブ投与が有効であった被験者には、再発した場合などにデノスマブの再投与を可能とした。

事前に計画した試験 AMG162-B-J201 の 6 ヶ月カットオフ解析（データカットオフ日: 20 年 月 日）及び 12 ヶ月カットオフ解析（データカットオフ日: 20 年 月 日）の結果を治験総括報告書としてまとめ、本申請資料に含めた。

表 2.7.3.1-1 GCTB を対象とした臨床試験

試験番号	試験の目的	試験デザイン 及び対照の種類	地域	組み入れ被験者数 (曝露被験者数)	主要な選択基準	試験期間 (追跡調査を含む)	試験の状態; 報告書の種類
20040215	有効性 (巨細胞の消失、病勢の進行なし)、薬力学、薬物動態、安全性、抗体反応	第 II 相、非盲検、単一群	米国、オーストラリア、フランス	37 名 (37 名にデノスマブを皮下投与 120 mg Q4W、第 8 日及び第 15 日に 120 mg 負荷投与 ^{a)})	組織学的に GCTB が確認された男女 年齢: 18 歳以上	腫瘍の完全切除、病勢の進行、中止、ビスフォスフォネート、カルシトニン、又はインターフェロン α -2a の投与、あるいは 試験 20062004 への移行まで	完了; 主要解析の治験総括報告書及び最終解析の治験総括報告書

Page 1 of 3

GCTB = 骨巨細胞腫; Q4W = 4 週間に 1 回

^a 高カルシウム血症が認められない限り、500 mg/日以上のカルシウム及び 400 IU/日以上ビタミン D を毎日投与することを強く推奨した。^b 事前に計画された第 3 回中間解析の時点で曝露された被験者の数。^c 事前に計画された 6 ヶ月カットオフ解析の時点で曝露された被験者の数。^d 高カルシウム血症が認められない限り、600 mg/日以上のカルシウム及び 400 IU/日以上ビタミン D を毎日投与した。

表 2.7.3.1-1 GCTB を対象とした臨床試験 (続き)

試験番号	試験の目的	試験デザイン 及び対照の種類	地域	組み入れ被験者数 (曝露被験者数)	主要な選択基準	試験期間 (追跡調査を含む)	試験の状態; 報告書の種類
20062004	安全性、有効性 (コホート 1 では病勢の進行までの期間、コホート 2 では第 6 ヶ月時点で手術を必要としなかった被験者の割合)	第 II 相、非盲検、単一群	北米、欧州、オーストラリア	予定数 500 名 (286 名 ^b にデノスマブを皮下投与 120 mg Q4W、第 8 日及び第 15 日に 120 mg 負荷投与 ^{a)})	成人又は骨格が成熟した未成年 (骨格成熟とは、放射線画像にて成熟した長骨が 1 つ以上と定義); 未成年は 12 歳以上、体重 45 kg 以上 コホート 1: 切除不能な GCTB の患者 コホート 2: 切除可能な GCTB であるが、重度の後遺症を残す手術が予定されている患者 コホート 3: 試験 20040215 から移行した被験者	腫瘍の完全切除後に CR 又は PR が病理学的に確認されてから 6 回投与まで、あるいは病勢の進行、中止、又は併用禁止療法の実施まで	進行中; 事前に計画された第 3 回中間解析の治験総括報告書

Page 2 of 3

GCTB = 骨巨細胞腫; Q4W = 4 週間に 1 回

^a 高カルシウム血症が認められない限り、500 mg/日以上のカルシウム及び 400 IU/日以上ビタミン D を毎日投与することを強く推奨した。^b 事前に計画された第 3 回中間解析の時点で曝露された被験者の数。^c 事前に計画された 6 ヶ月カットオフ解析の時点で曝露された被験者の数。^d 高カルシウム血症が認められない限り、600 mg/日以上のカルシウム及び 400 IU/日以上ビタミン D を毎日投与した。

2.7.3 臨床的有効性の概要

デノスマブ

ランマーク皮下注 120mg

表 2.7.3.1-1 GCTB を対象とした臨床試験 (続き)

試験番号	試験の目的	試験デザイン及び対照の種類	地域	組み入れ被験者数 (曝露被験者数)	主要な選択基準	試験期間 (追跡調査を含む)	試験の状態; 報告書の種類
AMG162-B-J201	有効性 (客観的奏効を認めた被験者の割合、奏効期間、客観的奏効までの期間、客観的奏効が 4、12、及び 24 週間以上持続した被験者の割合、各時期における客観的奏効を認めた被験者の割合)、薬物動態、安全性、抗体反応	第 II 相、非盲検、単一群	日本	予定数 10 名以上 (17 名 ^c にデノスマブを皮下投与 120 mg Q4W、第 8 日及び第 15 日に 120 mg 負荷投与 ^d)	成人又は骨格が成熟した未成年 (骨格成熟とは、放射線画像にて成熟した長骨が 1 つ以上と定義) ; 未成年は 12 歳以上、体重 45 kg 以上; 切除不能な GCTB の患者、切除可能な GCTB であるが、重度の後遺症を残す手術が予定されている患者	腫瘍の完全切除後に CR 又は PR が病理学的に確認されてから 6 回投与まで、あるいは病勢の進行、中止、又は併用禁止療法の実施まで	進行中; 事前に計画された 6 ヶ月カットオフ解析の治験総括報告書及び事前に計画された 12 ヶ月カットオフ解析の治験総括報告書

Page 3 of 3

GCTB = 骨巨細胞腫; Q4W = 4 週間に 1 回

^a 高カルシウム血症が認められない限り、500 mg/日以上のカルシウム及び 400 IU/日以上ビタミン D を毎日投与することを強く推奨した。

^b 事前に計画された第 3 回中間解析の時点で曝露された被験者の数。

^c 事前に計画された 6 ヶ月及び 12 ヶ月カットオフ解析の時点で曝露された被験者の数。

^d 高カルシウム血症が認められない限り、600 mg/日以上のカルシウム及び 400 IU/日以上ビタミン D を毎日投与した。

個々の試験（20040215、20062004、及び AMG162-B-J201）における有効性評価項目

外国第 II 相試験（20040215 及び 20062004）では、病理組織学的評価項目、治験責任医師の評価による画像評価項目、及び臨床的評価項目を用いてデノスマブの有効性を評価した。

試験 20040215 の主要目的は、再発又は切除不能な GCTB 患者を対象にデノスマブの治療効果を評価することであった。効果の評価は、投与前後の生検検体の病理組織学的分析により実施した。治験責任医師が投与後の生検検体を採取できなかった場合のみ、画像評価を実施した。画像評価では、コンピュータ断層撮影 (computed tomography: CT) 又は磁気共鳴画像法 (magnetic resonance imaging: MRI) によって測定した標的病変のベースラインからの増加 (最長径の変化) が、第 25 週時点で 20% 以上の場合を病勢の進行と定義した。固形癌の効果判定基準 (Response Evaluation Criteria in Solid Tumors: RECIST) は、骨腫瘍の評価用としては限界があるため (Costelloe et al, 2010)、試験 20040215 では適用しなかった。

試験 20062004 の主要目的は、GCTB 患者でのデノスマブの安全性プロファイルを評価することであり、有効性評価は副次目的であった。本試験の有効性評価項目は、コホート 1 における病勢の進行までの期間 (副次評価項目)、コホート 2 における第 6 ヶ月の時点で手術が施行されなかった被験者の割合 (副次評価項目)、並びに病勢の経時推移 (探索的評価項目) であった。有効性評価は、主に測定可能かつ生検可能であるという基準に従って治験責任医師が選択した標的病変を対象とした。病勢は、治験責任医師による評価に基づき、CR、PR、安定 (stable disease: SD)、及び病勢の進行 (progressive disease: PD) に分類された。

試験 20040215 及び 20062004 では有効性に関し、その他の副次及び探索的評価項目も設定した。病勢の進行、抗腫瘍効果、後遺症を伴う手術の遅延/減少、及び痛みの減少は、切除可能及び切除不能 GCTB 患者における有効性の適切な尺度になると考えた。これら評価項目の解析結果は、モジュール 5 の各試験の治験総括報告書に示した。両試験では、試験デザイン及び有効性評価項目が異なるため、治験実施計画書に規定された有効性評価項目に関する併合解析は実施しなかった。なお、客観的抗腫瘍効果の併合解析については、[第 1.2.2 項](#)で考察する。

国内第 II 相試験 AMG162-B-J201 の主要目的は、切除可能又は切除不能 GCTB 患者を対象にデノスマブの客観的抗腫瘍効果を評価することである。薬物動態及び安全性の評価は副次目的とした。本試験の有効性評価項目は、客観的奏効 (CR 又は PR) を認めた被験者の割合 (主要評価項目)、客観的奏効が 4、12、又は 24 週間以上持続した被験者の割合、各時期における客観的奏効を認めた被験者の割合、各画像評価 (modified RECIST 基準、modified European Organisation for Research and Treatment of Cancer [EORTC] 基準、及び density/size 基準) における客観的奏効を認めた被験者の割合、客観的奏効までの期間、及び奏効期間 (副次評価項目)、並びにすべての時点の中での最良効果、ベースライン時に計画していた手術よりも軽度の手術で済んだ被験者の割合、疼痛スコアの変化、骨代謝マーカーの変化、病理組織学的検査を実施した被験者における病理組織学的変化、病勢の進行までの期間、及び臨床的有益性を認めた被験者の割合 (探索的評価項目) とした。標的病変及び非標的病変の画像評価は、CT、MRI、又は全身フルオロデオキシグルコース陽電子放射型断層撮影 (fluorodeoxyglucose positron emission tomography: ¹⁸FDG-PET) によって実施した。客観的抗腫瘍効果は外国 2 試験併合解析

と同様の評価基準によって評価した（詳細は第 1.2.2 項参照）。

個々の試験における主な評価項目を表 2.7.3.1-2 に示す。

表 2.7.3.1-2 個々の試験の治験実施計画書で規定した主な有効性評価項目

評価項目	評価項目の位置づけ 試験 20040215	評価項目の位置づけ 試験 20062004 ^a	評価項目の位置づけ 試験 AMG162-B-J201
抗腫瘍効果に関する評価項目			
客観的奏効 ^b を認めた被験者の割合（客観的奏効率）	-	-	主要
奏効期間 ^b	-	-	副次
客観的奏効までの期間 ^b	-	-	副次
客観的奏効が 4、12、及び 24 週間以上持続した被験者の割合 ^b	-	-	副次
各時期における客観的奏効を認めた被験者の割合 ^b	-	-	副次
奏効率；奏効は以下の定義： - 巨細胞の 90%消失又は巨細胞が腫瘍の 5%未満であった場合は巨細胞の完全消失 - 病理組織学的評価未実施の場合、第 25 週時点の画像評価で、ベースラインに比べて治験責任医師の評価による標的病変の進行がみられない	主要 ^c	-	- ^d
病勢の進行までの期間	-	副次（コホート 1） 探索（全被験者）	探索
骨病変の評価：治験責任医師の評価による臨床的有益性、骨石灰化、骨修復	探索	探索 （臨床的有益性のみ）	探索 （臨床的有益性のみ）
生活の質に関する評価項目			
BPI-SF における「最もひどい」痛みスコアのベースラインからの変化	-	探索	探索
鎮痛薬スコアのベースラインからの変化	-	探索	- ^d
手術の回避に関する評価項目			
第 6 ヶ月の時点で手術が施行されなかった被験者の割合	-	副次（コホート 2）	- ^d
ベースライン時に計画していた手術よりも軽度の手術で済んだ被験者の割合	-	探索（コホート 2）	探索
骨代謝マーカーの変化（sCTX、uNTX、BSAP、TRAP-5b、OC）	副次／探索	-	探索

BPI-SF = 簡易疼痛調査用紙（縮小版）；BSAP = 骨型アルカリフォスファターゼ；sCTX = 血清中 C-テロペプチド；uNTX = 尿中 N-テロペプチド；OC = オステオカルシン；TRAP-5b = 酒石酸抵抗性酸性フォスファターゼ 5b

^a 試験 20062004 の主要評価項目は安全性評価である。

^b 客観的抗腫瘍効果は中央読影機関において、modified RECIST 基準、modified EORTC 基準、及び density/size 基準によって評価した（第 1.2.2 項参照）。

^c 主要な効果の測定には、投与前後の生検検体の病理組織学的分析を用いた。投与後の生検検体を採取できなかった場合のみ、画像評価を実施した。

^d Statistical analysis plan（6-month cut-off analysis）で規定した解析、又は追加解析にて実施した。

1.2.2 客観的抗腫瘍効果の解析

試験 20040215 及び 20062004 に組み入れられた被験者におけるデノスマブの客観的抗腫瘍効果を評価するため、中央読影機関によって、レトロスペクティブに独立した画像評価を実施した。本画像評価のデザイン及び評価項目を以下に示す（第 1.2.2.1 項及び第 1.2.2.2 項）。両試験データの併合解析（外国 2 試験併合解析）の統計手法については、第 1.2.2.3 項で述べる。客観的抗腫瘍効果に関する独立した評価の統計解析計画書は、これらの解析前に確定した。

試験 AMG162-B-J201 ではプロスペクティブに、中央読影機関によって客観的抗腫瘍効果を評価した。また、試験 20040215、20062004、及び AMG162-B-J201（6 ヶ月カットオフ解析時）における客観的抗腫瘍効果の併合解析（3 試験併合解析）を実施した。これらの試験及び解析での客観的抗腫瘍効果の評価のデザイン及び評価項目、並びに統計手法は、外国 2 試験併合解析の場合（第 1.2.2.1 項、第 1.2.2.2 項、及び第 1.2.2.3 項）と概ね同様である。試験 AMG162-B-J201 及び 3 試験併合解析の統計解析計画書は、これらの解析前に確定した。

1.2.2.1 客観的抗腫瘍効果の評価方法

試験 20040215 及び 20062004 における客観的抗腫瘍効果の独立した評価は、規制当局（FDA 等）との協議により合意を得て実施した。

両試験に組み入れられた被験者の投与前及び試験期間中の画像を入手するため、同意書を修正し、該当する治験審査委員会（IRB）／独立倫理委員会（IEC）から承認を得た上で、すべての被験者から再度同意を試みた。

抗腫瘍効果及び病勢の進行の評価のため、入手したすべての画像（CT、MRI、及び／又は¹⁸F-FDG-PET による）を中央読影機関に提供した。なお、本画像評価では、単純 X 線フィルム、骨スキャン検査結果、又は超音波検査結果は対象としなかった。

GCTB に関し、確立した抗腫瘍効果の評価基準はない。抗腫瘍効果の評価基準である RECIST は、主に軟部組織における固形癌のサイズの縮小を基準としている。軟部組織成分のない腫瘍は通常、RECIST 基準では評価されない。すべての GCTB 患者が測定可能な軟部組織病変を有するわけではなく、また、抗腫瘍効果は腫瘍サイズの減少とは相関しない場合もあるため、RECIST のみでは、GCTB に対する抗腫瘍効果を評価することには限界がある（Eisenhauer et al, 2009）。原発性骨腫瘍の場合には、腫瘍サイズの減少ではなく、腫瘍代謝活性の低下又は腫瘍における骨形成（石灰化）の増加を指標として抗腫瘍効果を評価することは可能である。このため、新たに 2 つの評価基準、すなわち EORTC 基準（¹⁸F-FDG-PET）及び modified inverse Choi（density/size）基準を追加して、客観的抗腫瘍効果を評価することとした。これらの基準を用いることにより、腫瘍細胞の活動性の尺度としての代謝活性の変化と、組織学的変化の尺度としての GCTB 病変の密度上昇という、異なるアプローチで抗腫瘍効果を評価することができる。以上のように、GCTB については効果判定基準が十分に確立されていないため、RECIST、EORTC、及び inverse Choi 基準を含めた包括的な腫瘍縮小効果判定を用いることとした。

これまで GCTB では、¹⁸F-FDG-PET 及び Choi 基準を抗腫瘍効果の評価に用いたという報告はない。しかし、いくつかの臨床研究から、治療効果の初期バイオマーカーとして ¹⁸F-FDG-PET

による評価が利用可能であることが示されている。¹⁸FDG-PETによるEORTC効果判定基準は、化学療法施行後にStandardized Uptake Value by Body Weight (SUV_{bw})の平均値が大きく低下したことを示した複数の臨床研究結果に基づいて設定されたものである(Young et al, 1999)。SUV_{bw}の平均値の低下は、腫瘍の種類や化学療法の種類によらず認められている。¹⁸FDG-PETによって評価した代謝活性の変化によって、臨床転帰を予想することも可能である。Zerizerらは、肝転移患者において、SUV_{bw}の平均値の低下と腫瘍マーカーの減少との間に有意な相関があり、無増悪生存期間が予測可能であることを報告している(Zerizer et al, 2012)。同様に、複数の臨床研究からChoi基準は、消化管間質腫瘍(Choi et al, 2007)及び軟部組織肉腫(Stacchiotti et al, 2009)において、化学療法に対する病理学的反応の予測性がRECISTよりも高いことが示されている。これらの臨床研究では、効果判定において腫瘍サイズ以外の変化を測定することの付加価値が示されており、特殊な腫瘍やGCTBのような骨腫瘍の評価において、特に意義があると考えられる。

客観的奏効は、以下の評価基準による最良効果を用いて判定したCR又はPRと定義した。

- CT/MRIにより腫瘍量を評価するmodified RECIST 1.1 基準
- ¹⁸FDG-PETにより代謝活性を評価するmodified EORTC 基準
- CT/MRIによる腫瘍サイズとCTのHounsfield単位による密度を評価するmodified inverse Choi (density/size) 基準

これら3つの評価基準を用い、GCTB患者における客観的抗腫瘍効果を総合的に評価した。本評価に用いた基準のオリジナルの基準からの改変箇所を表2.7.3.7-8、表2.7.3.7-9、及び表2.7.3.7-10に示す。

Modified RECIST 1.1 基準による客観的抗腫瘍効果は、表2.7.3.1-3に示した評価基準(Eisenhauer et al, 2009)に従い、標的病変及び非標的病変を評価・判定した。抗腫瘍効果の確定のための追跡評価は不要とし、2名の読影者が本基準を用いて判定した。なお、新規病変の出現はPDとした。

表 2.7.3.1-3 Modified RECIST 1.1 評価基準

効果	標的病変	非標的病変
完全奏効 (CR)	すべての標的病変の消失。標的病変としたすべてのリンパ節病変の短径が 10 mm 未満。	すべての非標的病変の消失。すべての非標的リンパ節病変の短径が 10 mm 未満。
部分奏効 (PR)	ベースライン SLD に比して、SLD が 30% 以上減少。	-
安定 (SD)	最小の SLD に比して、標的病変で PR に相当する縮小がなく、PD に相当する増大がない。	CR 又は PD に相当しない非標的病変の 1 つ以上の残存。
病勢の進行 (PD)	最小の SLD に比して、標的病変の SLD が 20%以上増大、かつ、SLD の絶対値でも 5 mm 以上増加。	既存の非標的病変の明らかな増悪。
評価不能 (UE)	ベースライン時に存在していたが、その後 UE になった標的病変。	ベースライン時に存在していたが、その後 UE になったあらゆる非標的病変。

SLD = 最長径の和

^{18}F FDG-PET 又は ^{18}F FDG-PET/CT の結果が評価可能な場合は、 SUV_{bw} を用いて効果を評価した。Modified EORTC 基準による客観的抗腫瘍効果は、表 2.7.3.1-4 に示した評価基準 (Young et al, 1999) に従い、標的病変を評価・判定した。なお、新規病変の出現は PD とした。

表 2.7.3.1-4 Modified EORTC 評価基準

効果	PET 標的病変
CR	すべての標的病変の腫瘍容積内における異常な FDG 取り込みが、周囲の正常組織と区別できないレベルまで、完全に消失。
PR	SUV_{max} 合計の変化率 ($\%\Delta\Sigma \text{SUV}_{\text{max}}$) が、ベースラインに比して 25%以上減少。
SD	ベースラインに比して、 $\%\Delta\Sigma \text{SUV}_{\text{max}}$ の増加が 25%未満、又は減少が 25%未満。
PD	$\%\Delta\Sigma \text{SUV}_{\text{max}}$ が、ベースラインに比して 25%以上増加。
UE	^{18}F FDG-PET 検査が実施できなかった、又は実施したが UE と判断された場合。標的病変のうちの 1 つが UE と判断され、かつ PD に相当しない場合は、評価可能な標的病変で明らかな増悪と判定されない限り、当該時点での CR、PR、又は SD の判定はできず、UE と評価。

CR = 完全奏効; ^{18}F FDG-PET = フルオロデオキシグルコース陽電子放射型断層撮影; PD = 病勢の進行; PR = 部分奏効; SD = 安定; SUV_{max} = SUV の最大値; UE = 評価不能

CT 又は MRI 画像が評価可能な場合は、CT Hounsfield 単位のベースラインからの増加（密度上昇）を用いて病変の骨化を評価し（CT 画像のみ適応）、最長径を用いて病変のサイズを評価した。この評価（以下、density/size 基準による評価）には、Choi 基準 (Choi et al, 2007) を改変して適用した。Choi らは、消化管間質腫瘍の患者において CT Hounsfield 単位の変化と ^{18}F FDG-PET スキャンの変化の関連性を評価し、CT Hounsfield 単位による密度の減少と抗腫瘍効果との関係を示した。デノスマブは破骨活性を阻害するため、デノスマブ投与による GCTB 病変の骨化及び石灰化が期待される（すなわち、腫瘍組織の石灰化骨への置換）。したがって、本評価では、病変の密度の増加（Hounsfield 単位の変化率により測定）を基準として Choi 基準を改変して（すなわち、Choi 基準とは逆の反応を効果として）適用した。また、病変のサイズ変化については、Choi 基準に準拠して評価した。Density/size 基準による標的病変の客観的抗腫瘍効果は、表 2.7.3.1-5 に示す基準を用いて 2 名の読影者が判定した。なお、新規病変の出現は PD とした。

表 2.7.3.1-5 Density/Size 基準による評価

効果	標的病変
CR	すべての腫瘍の消失
PR	ベースラインに比して、サイズ (%Δ Choi SLD) が 10%以上縮小、又は CT 密度 (%ΔHU _{mea}) が 15%以上増加。
SD	CR、PR、又は PD に相当しない場合。
PD	一次元の腫瘍サイズ (Choi SLD) の増加が 10%を超え、CT 密度が PR に相当しない場合。CT/MRI によって新規病変の出現が確認された場合。
UE	CT/MRI 検査が実施できなかった、又は実施したが UE と判断された場合。標的病変が、密度及びサイズの測定によって評価できないと判断され、PD に相当しない場合は、当該時点に CR、PR、又は SD の判定はできず、UE と評価。

CR = 完全奏効; CT = コンピュータ断層撮影; MRI = 核磁気共鳴画像法; PD = 病勢の進行; PR = 部分奏効; SD = 安定; SLD = 最長径の和; UE = 評価不能

客観的抗腫瘍効果の評価に用いた各評価基準に関する詳細は、[Imaging Review Charter](#) に示した (モジュール 5、第 5.3.5.3 項)。

判定プロセスにおけるバイアス、病勢の自然経過、平均値への回帰などのアーチファクトとデノスマブの真の効果を区別する一助として、20 名以上の被験者のデノスマブ投与開始前の画像 (通常の治療経過中に測定したもの) を用いて、画像コントロール群を構成した。これらの投与開始前の画像データは、試験期間中の抗腫瘍効果又は病勢の進行の評価に用いた基準と同じ画像評価基準を用いて評価した。なお、コントロール群の被験者は、試験期間中の投与群にも含まれているため、画像読影者が当該被験者を認知できないよう盲検下で評価した。同一被験者におけるコントロール群の画像と試験期間中の画像は、それぞれ異なる読影者が評価した。

試験 AMG162-B-J201 における客観的腫瘍効果の評価方法は、外国 2 試験併合解析に用いた評価方法と同じである。ただし、試験 AMG162-B-J201 では、画像コントロール群を設定しなかった。客観的抗腫瘍効果の評価に用いた各評価基準に関する詳細は、[Imaging Review Charter for Japan Study](#) に示した (モジュール 5、第 5.3.5.3 項)。

1.2.2.2 客観的抗腫瘍効果の解析における有効性評価項目

複数のアプローチでデノスマブの抗腫瘍効果を評価したことから、客観的抗腫瘍効果を認めた被験者の割合 (主要評価項目) を示すとともに、各評価基準ごとに判定した客観的抗腫瘍効果の結果も示し、総合的に評価した。また、デノスマブの抗腫瘍効果の持続性を評価するため、副次及び探索的評価項目として、客観的抗腫瘍効果の発現時期及び持続期間、並びに病勢の進行までの期間を評価した。外国 2 試験併合解析での客観的抗腫瘍効果の解析における有効性評価項目を表 2.7.3.1-6 に要約する。これらの解析の結果を第 3.2 項に要約する。すべての図表及び一覧は、モジュール 5.3.5.3 (有効性の併合解析) に示した。

試験 AMG162-B-J201 の客観的抗腫瘍効果に関する評価項目は、外国 2 試験併合解析の評価項目 (表 2.7.3.1-6) と同様であるため、本申請資料では外国 2 試験併合解析の評価結果とともに、当該試験及び 3 試験併合解析の結果を第 3.2 項に提示する。これらのすべての図表及び一

覧は、モジュール 5.3.5.3（有効性の併合解析）に示した。

表 2.7.3.1-6 客観的抗腫瘍効果の解析における有効性評価項目：外国 2 試験併合解析

主要 ^a :	客観的奏効を認めた被験者の割合（客観的奏効率）
副次 ^a :	- 客観的奏効の持続期間 - 客観的奏効までの期間 - 客観的奏効が 4 週間以上持続した被験者の割合 - 客観的奏効が 8 週間以上持続した被験者の割合 - 客観的奏効が 12 週間以上持続した被験者の割合 - 客観的奏効が 24 週間以上持続した被験者の割合 - 腫瘍コントロールが 12 週間以上持続した被験者の割合 - 各抗腫瘍効果評価基準を用いて判定した場合と、あらゆる評価基準を用いた最良効果に基づき判定した場合に、客観的奏効を認めた被験者の累積割合の経時推移
探索:	- あらゆる評価基準を用いた最良効果別の被験者の割合 - あらゆる評価基準を用いた病勢の進行までの期間 あらゆる評価基準を用いた最良効果に基づき、客観的奏効を認めた被験者と認めなかった被験者の間で、以下の項目を検討: - 治験責任医師が判断した臨床的有益性 - 痛みのスコア（最もひどい痛み、痛みの重症度、及び痛みによる機能障害）及び鎮痛薬スコアの変化（いずれも試験 20062004） - 最もひどい痛みのスコアが 2 ポイント以上減少した被験者の割合（試験 20062004） - 最もひどい痛みのスコアが 4 ポイント以下であった被験者の割合（試験 20062004） - 最もひどい痛みのスコアが 2 ポイント以上増加した被験者の割合（試験 20062004） - 最もひどい痛みのスコアが 4 ポイントを超えた被験者の割合（試験 20062004） - 最もひどい痛みのスコアが 2 ポイント以上減少するまでの期間（試験 20062004） - 最もひどい痛みのスコアが 4 ポイント以上減少するまでの期間（試験 20062004） - 最もひどい痛みのスコアが 2 ポイント以上増加するまでの期間（試験 20062004） - 最もひどい痛みのスコアが 4 ポイントを超えるまでの期間（試験 20062004） - 鎮痛薬スコアがベースライン時の 2 ポイント以下から試験期間中に 3 ポイント以上に变化した被験者の割合（試験 20062004） - 鎮痛薬スコアがベースライン時の 3 ポイント以上から試験期間中に 2 ポイント以下に変化した被験者の割合（試験 20062004）

^aあらゆる評価基準（modified RECIST 1.1 基準、modified EORTC 基準、及び density/size 基準）を用いた最良効果による客観的奏効とともに、各評価基準ごとの客観的奏効についても示す。

EORTC = European Organisation for Research and Treatment of Cancer; RECIST = Response Evaluation Criteria in Solid Tumors

臨床的有効性の概要のための統計解析計画書 - 骨巨細胞腫（モジュール 5、第 2.0 版）

1.2.2.3 客観的抗腫瘍効果の解析における統計手法

客観的抗腫瘍効果の評価における解析は記述的なものとし、仮説検定は行わなかった。外国 2 試験併合解析での画像評価には、試験 20040215 の最終解析データのカットオフ日（2010 年 11 月 日）までに収集されたデータと、試験 20062004 の事前に計画された第 3 回中間解析のカットオフ日（2011 年 3 月 25 日）までに収集されたデータを用いた。客観的抗腫瘍効果の解析結果は、試験コホートごと（試験 20062004 のコホート 1 及び 2）及び個々の試験ごとの結果とともに、両試験のデータを併合した解析の結果を示す。未成年被験者の集団についても同様の解析を実施した（試験 20062004 のみ）。レトロスペクティブな画像データの収集及び評価は、両試験で同様に実施した。したがって、客観的抗腫瘍効果の解析のためにこれら試験のデータを併合することは適切であると考えられた。

客観的抗腫瘍効果の解析における有効性評価項目の統計手法を表 2.7.3.1-7 に要約する。詳細

は、モジュール 5 の臨床的有効性の概要のための統計解析計画書に示した。客観的抗腫瘍効果に関する統計解析計画書は、中央読影機関から受領した画像評価結果の解析を実施する前に確定した。

試験 AMG162-B-J201 の画像評価では、事前に計画された 6 ヶ月カットオフ解析のカットオフ日（20 年 月 日）及び 12 ヶ月カットオフ解析のカットオフ日（20 年 月 日）までに収集されたデータを用いた。3 試験併合解析では当該データと、レトロスペクティブに収集した試験 20040215 及び 20062004 の画像データを用いた。客観的抗腫瘍効果に関する評価項目の解析方法は、外国 2 試験併合解析と同じ方法（表 2.7.3.1-7）とした。

表 2.7.3.1-7 客観的抗腫瘍効果の解析における統計手法

評価項目のレベル	評価項目	方法
主要	客観的奏効を認めた被験者の割合 (客観的奏効率)	粗推定値及び両側 95% CI を要約
副次	客観的奏効までの期間及び客観的奏効の持続時間	該当する場合は、四分位範囲 (中央値、25 及び 75 パーセントイル値) の Kaplan-Meier 推定値及び両側 95% CI を算出
	客観的奏効が 4、8、12、及び 24 週間以上持続した被験者の割合	粗推定値及び両側 95% CI を要約
	客観的奏効を認めた被験者の累積割合の経時推移 (6 ヶ月間隔)	客観的奏効に関する Kaplan-Meier イベント発生率の経時推移及び両側 CI を要約
	腫瘍コントロールが 12 週間以上持続した被験者の割合	粗推定値及び両側 95% CI を要約
探索	試験期間中の最良効果別の被験者の割合	CR、PR、SD、PR 以上、及び SD 以上の粗推定値及び両側 95% CI を要約
	病勢進行までの期間	該当する場合は、四分位範囲 (中央値、25 及び 75 パーセントイル値) の Kaplan-Meier 推定値及び両側 95% CI を算出
その他の探索的評価項目: あらゆる評価基準を用いた最良効果に基づき、客観的奏効を認めた被験者と認めなかった被験者:		
	治験責任医師が判断した臨床的有益性が得られた被験者の割合	記述統計量を要約
試験 20062004 及び試験 AMG162-B-J201:		
	痛みのスコア (最もひどい痛み、痛みの重症度、及び痛みによる機能障害) の変化	記述統計量を要約
	最もひどい痛みのスコアが 2 ポイント以上減少した被験者と、最もひどい痛みのスコアが 4 ポイント以下であった被験者の割合の経時推移	記述統計量を要約
	最もひどい痛みのスコアが 2 ポイント以上増加した被験者と、最もひどい痛みのスコアが 4 ポイントを超えた被験者の割合の経時推移	記述統計量を要約
	最もひどい痛みのスコアが 2 ポイント以上減少するまでの期間又は最もひどい痛みのスコアが 4 ポイント以下になるまでの期間	該当する場合は、中央値の Kaplan-Meier 推定量及び両側 95% CI を算出
	最もひどい痛みのスコアが 2 ポイント以上増加するまでの期間又は最もひどい痛みのスコアが 4 ポイントを超えるまでの期間	該当する場合は、中央値の Kaplan-Meier 推定量及び両側 95% CI を算出
	鎮痛薬スコアの変化及び鎮痛薬スコアが変化した被験者の割合の経時推移	記述統計量を要約

CI = 信頼区間; CR = 完全奏効; PR = 部分奏効; SD = 安定

1.3 有効性評価デザインの概要

アムジェン社は、GCTB 患者を対象として 2 つの国際共同、非盲検、単一群、第 II 相臨床試験を実施した。規制当局（FDA 等）との協議後、有効性評価のため客観的抗腫瘍効果に関する独立したレトロスペクティブな画像評価を実施した。GCTB については、確立した評価基準がない。そこで今回は、デノスマブの客観的抗腫瘍効果を評価するため、3 つの評価基準を用いた。この独立した画像評価の主要目的は、各評価基準(modified RECIST 基準、modified EORTC 基準、及び density/size 基準)を用いて、GCTB 患者における客観的抗腫瘍効果を評価することとした。副次目的は、客観的奏効の持続期間及び持続性、客観的奏効までの期間、並びに腫瘍コントロールを評価することとした。

主な有効性評価デザインの概要を以下に示す。

- 2 つの第 II 相試験のいずれにおいても、被験者集団は申請する適応症の患者集団を代表するものであり、検討した被験者数は、このまれな疾患としては最大である。
- 独立した画像評価によって 2 試験で一貫した評価を行い、客観的抗腫瘍効果の併合解析を可能とした。
- デノスマブの客観的抗腫瘍効果を評価するため、3 つの評価基準を用いた。
 - 1) Modified RECIST 1.1 基準は、抗腫瘍効果の評価として確立した評価基準であり、主に軟部組織における固形癌のサイズ縮小に基づく基準である。
 - 2) 客観的抗腫瘍効果の評価ではさらに、modified EORTC 基準 (^{18}F FDG-PET) 及び density/size 基準を用いた。これらの基準により、代謝活性の変化と腫瘍密度の上昇という異なるアプローチで抗腫瘍効果を評価することを可能とした。
 - 3) あらゆる評価基準を用いた最良効果の評価及び個々の評価基準ごとの評価を実施し、総合的に抗腫瘍効果を評価した。
- 判定プロセスにおけるバイアス、病勢の自然経過、平均値への回帰などのアーチファクトとデノスマブの真の効果を区別する一助として、デノスマブ投与開始前の画像を用いて画像コントロール群を設定し、通常の治療経過中の画像を評価した。
- 各試験であらかじめ規定していた有効性の評価結果は、独立した画像評価に基づく客観的抗腫瘍効果に関する併合解析結果を裏付けるための有効性データとした。

試験 AMG162-B-J201 における客観的抗腫瘍効果に関する評価デザインも上述と同様とした。ただし、試験 AMG162-B-J201 では、画像コントロール群を設定しなかった。

2. 個々の試験結果の要約

試験 20040215、20062004、及び AMG162-B-J201 の結果の要約を、それぞれ第 2.1 項、第 2.2 項、第 2.3 項、及び第 2.4 項に示す。

2.1 試験 20040215（外国第 II 相臨床試験）

再発又は切除不能な GCTB 患者を対象としてデノスマブ（AMG 162）の安全性及び有効性を評価する非盲検多施設共同第 II 相試験

方法

試験 20040215 は、再発又は切除不能な GCTB の成人患者を対象とした非盲検、単一群、第 II 相試験である。本試験では、デノスマブ 120 mg Q4W の皮下投与に、第 8 日及び第 15 日の 120 mg 負荷投与を加えた投与スケジュールとした。本試験の目的は、成人の GCTB 患者におけるデノスマブの抗腫瘍効果（病理組織学的検査又は画像検査により判定）及び安全性を評価することであった。主要評価項目は抗腫瘍効果とした。当該効果は、巨細胞の 90%以上の消失、又は腫瘍細胞に占める巨細胞の割合が 5%未満の場合は巨細胞の完全な消失、あるいは第 25 週までに画像検査で標的病変の進行を示す所見なし、と定義した。安全性は、有害事象の種類、頻度、及び重症度、並びに臨床検査値異常によって評価した。その他の副次及び探索的評価項目は、骨代謝マーカー、骨病変部の石灰化及び修復、臨床的有益性、抗デノスマブ抗体、並びにデノスマブの薬物動態とした。また、デノスマブの最終投与後 2 年間、安全性データを 6 ヶ月間隔で収集した（安全性の追跡調査期）。

結果

合計 37 名の再発又は切除不能な GCTB の患者が本試験に組み入れられ、全被験者が 1 回以上のデノスマブ投与を受けた。主要解析では、データカットオフ（20 年 月 日）の時点で有効性の結果をまとめた。有効性の解析対象集団に組み入れられた 35 名のうち、85.7% (95% 信頼区間 [CI] : 69.7~95.2) で抗腫瘍効果がみられた。十分な病理組織学的検査データが得られた 20 名では全員（100.0%）が効果判定基準を満たし、画像検査データのみに基づき評価した 15 名中 10 名（66.7%）が効果判定基準を満たした（第 7 項、表 2.7.3.7-4 及び表 2.7.3.7-5）。画像検査による病変の最長径の変化は、主要評価項目の解析結果と概ね一致した。奏効率は、年齢又はビスフォスフォネート使用歴にかかわらず一貫していた。クレアチニン補正尿中 N-テロペプチド（urinary N-telopeptide corrected for urine creatinine: uNTX/Cr）及び血清中 C-テロペプチド-I（C-telopeptide-1: CTX1）は第 5 週以降、一貫して抑制された（ベースラインから約 80%減少）。その他の骨代謝マーカー（骨型アルカリフォスファターゼ [bone-specific alkaline phosphatase: BSAP]、オステオカルシン、及び酒石酸抵抗性酸性フォスファターゼ 5b

[tartrate-resistant acid phosphatase 5b: TRAP 5b]）も減少し、試験期間を通じてベースラインを下回った。デノスマブの投与によって、骨石灰化、病変部の骨修復、及び臨床的有益性が認められた。負荷投与終了時（第 5 週）、血清中デノスマブ濃度トラフ値の平均値及び中央値は初

回投与後の約 2 倍であり、負荷投与を加えた投与スケジュールによって、全身曝露量が投与後 1 ヶ月で目標値まで増加したことが示された。

37 名の全被験者が安全性の解析対象集団に含まれた。21 名が安全性追跡調査期間に移行し、安全性の追跡調査解析対象集団に含まれた。以下には、最終解析（2010 年 11 月 ■ 日）時点での安全性の結果をまとめた（試験 20040215 治験総括報告書[最終解析]）。試験期間中、89.2% の被験者で少なくとも 1 件の有害事象が発現した。比較的良好に見られた有害事象は、関節痛 arthralgia、背部痛 back pain、及び四肢痛 pain in extremity であった。9 名（24.3%）で重篤な有害事象が認められた。10 名（27.0%）で CTCAE グレード 3、4、又は 5 の有害事象が認められた。1 名が悪性新生物進行 malignant neoplasm progression のために死亡し、2 名（5.4%）が有害事象のために試験を中止した（肺転移 metastases to lung: 関連なし、顎骨壊死 osteonecrosis of the jaw [ONJ]：関連あり）。

安全性の追跡調査期間中、57.1%の被験者で有害事象が発現した。比較的良好に見られた有害事象は、筋力低下 muscular weakness、貧血 anemia、関節痛 arthralgia、及び悪心 nausea であった。6 名（28.6%）で重篤な有害事象が報告された。6 名（28.6%）で CTCAE グレード 3、4、又は 5 の有害事象が認められた。本期間中、5 名の被験者が死亡し（うっ血性心不全 cardiac failure congestive、疾患進行 disease progression、肺転移 metastases to the lung、及び心室性頻脈 ventricular tachycardia）、2 名（9.5%）が有害事象のために試験を中止した（疾患進行 disease progression、肺転移 metastases to the lung: いずれも関連なし）。

試験期間中又は安全性の追跡調査期間中、低カルシウム血症 hypocalcemia の有害事象は認められなかった。約 18 ヶ月間の投与後に、ONJ と判定された事象が 1 名で発現した。試験期間中又は追跡調査期間中に新規原発性悪性腫瘍の報告はなかった。予想された血清中カルシウム及びリンの変化を除き、臨床検査値の変化は認められなかった。バイタルサインへの影響を示す変化は観察されなかった。抗デノスマブ抗体の検査を行った被験者に、抗デノスマブ抗体が陽性の被験者は認められなかった。

結論

デノスマブを GCTB 患者に投与した結果、良好なベネフィット／リスクプロファイルが確認された。デノスマブを投与した GCTB 患者では、治療効果及び骨代謝の抑制が認められ、忍容性は良好であった。

2.2 試験 20062004（外国第 II 相臨床試験）

GCTB 患者を対象としたデノスマブの非盲検多施設共同第 II 相試験

方法

試験 20062004 は、切除可能な GCTB 又は切除不能な GCTB の成人患者及び骨格が成熟した未成年患者を対象とした進行中の非盲検、単一群、第 II 相試験である。本試験では、デノスマブ 120 mg Q4W の皮下投与に、第 8 日及び第 15 日の 120 mg 負荷投与を加えた投与スケジ

ルールとした。次の3つのコホートを組み入れた。すなわち、コホート1は、手術により切除不能な GCTB 患者（例：仙骨又は脊椎の GCTB、あるいは肺転移などの多発性病巣を有する患者）、コホート2は、切除可能な GCTB を有し、重度の後遺症が残る手術（例：関節切除、手足の切断、又は片側骨盤切断）の実施が試験中に予定されている患者、コホート3は、試験 20040215 から移行した被験者とした。本試験の目的は、デノスマブ投与の安全性を評価することであった。安全性は、有害事象の種類、頻度、及び重症度、並びに臨床検査値異常によって評価した。副次及び探索的目的は、デノスマブ投与の効果を評価することであった。副次評価項目は、コホート1では病勢の進行までの期間とし、コホート2では第6ヵ月の時点で手術が施行されなかった被験者の割合とした。探索的評価項目は、病理組織学的検査、痛みの尺度、及び治験責任医師の評価による臨床的有益性とした。

結果

2011年3月25日（事前に計画された第3回中間解析カットオフ日）までに、286名の被験者が試験 20062004 に組み入れられた。試験 20040215 から移行した被験者のうち4名は、安全性の追跡調査期に直接組み入れられたため、以下に示す投与期の解析には含まれない。

コホート1（n=170）では、1回以上投与された169名中6名（3.6%）が、治験責任医師によって病勢の進行と判断された。病勢の進行までの期間は中央値に達しなかった。コホート2（n=101）では、6ヵ月以上投与された71名中64名（90.1%; 95% CI: 80.7~95.9）で、第6ヵ月までに手術は施行されなかった。本中間解析の時点で、コホート2において1回以上投与された被験者100名のうち90名（90.0%; 95% CI: 82.4~95.1）では、手術が施行されなかったか（74名）、ベースライン時に計画していた手術よりも軽度の手術が施行された（16名）。コホート1と2を併せて、治験責任医師が報告した最良効果がSD以上であった被験者は、252名中250名（99.2%）であった（CR 25名 [9.9%]、PR 94名 [37.3%]、SD 131名 [52.0%]）。コホート3（n=11）では、治験責任医師が報告した最良効果がPRの被験者が2名、SDの被験者が9名であった。

評価可能な被験者の多くでは、デノスマブ投与によって速やかに痛みが改善した。ベースライン時に最もひどい痛みのスコアが2ポイント以上であった被験者の31.4%では、投与開始後1週間以内に、臨床的に意味のある最もひどい痛みの減少（最もひどい痛みのスコアがベースラインから2ポイント以上の減少）が認められた。また、これらの被験者の50%以上では、第5週及びその後の全評価時点で臨床的に意味のある最もひどい痛みの減少が認められ、痛みの改善は、試験期間中にわたって持続した。平均鎮痛薬スコアは、試験期間中を通じて低かったため（2ポイント以下）、痛みの改善は、鎮痛薬使用の増加に関連したものではないと考えられた。なお、patient-reported outcome（PRO）解析対象集団において、ベースライン時の鎮痛薬使用がなし又は少量であり、いずれかの来院時点で強力なオピオイド（スコア: 3ポイント以上）使用に移行した被験者の割合は、全被験者の4.7%以下であった。

合計281名の被験者が1回以上のデノスマブ投与を受けた。84.0%の被験者で少なくとも1件の有害事象が発現した。比較的によく見られた有害事象は、関節痛 arthralgia、頭痛 headache、

悪心 nausea、及び疲労 fatigue であった。25 名 (8.9%) で重篤な有害事象が認められた。CTCAE グレード 3、4、又は 5 の有害事象の発現率は 17.8% であった。本試験中、1 名 (0.4%) が呼吸不全 respiratory failure のために死亡し、13 名 (4.6%) が有害事象のために試験を中止した。

15 名 (5.3%) で低カルシウム血症 hypocalcemia の有害事象が認められたが、重篤なものはなく、投与中止に至った被験者はいなかった。ONJ と判定された事象が 3 名 (1.1%) で認められ、それぞれ投与開始から約 13、14、及び 20 ヶ月後に発現した。5 名において、試験中に悪性骨腫瘍 bone malignancies 又は病勢の進行 disease progression が認められた。2 名は GCTB が悪性転換したものと考えられ、このうち 1 名は過去の放射線療法との関連が示唆された。残り 3 名では、治験責任医師から提供された情報及び事象の発現時期を踏まえると、新規悪性腫瘍の発現ではないと考えられた。

データカットオフ日までに組み入れられた未成年被験者 (18 歳未満) は 10 名と少なかったが、未成年被験者における安全性及び有効性の結果は、成人被験者と同様であった。

結論

GCTB 患者におけるデノスマブの忍容性は概ね良好であった。デノスマブの安全性プロファイルは、これまでに確認されたデノスマブ 120 mg の安全性プロファイルと同様であった。治験責任医師による疾患の状態の評価の結果、病勢の進行を認めた被験者は少なく、多くの被験者で計画した手術が延期された。以上のことから、GCTB 患者におけるデノスマブの高い臨床的意義が示唆された。

2.3 試験 AMG162-B-J201 (国内第 II 相臨床試験) : 6 ヶ月カットオフ解析

GCTB 患者を対象とした AMG 162 (デノスマブ) の非盲検多施設共同第 II 相臨床試験

方法

試験 AMG162-B-J201 は、切除可能な GCTB 又は切除不能な GCTB の成人患者及び骨格が成熟した未成年患者を対象とした進行中の非盲検、単一群、第 II 相試験である。本試験では、デノスマブ 120 mg Q4W の皮下投与に、第 8 日及び第 15 日の 120 mg 負荷投与を加えた投与スケジュールとした。本試験の主目的は、デノスマブの客観的抗腫瘍効果を評価することである。薬物動態及び安全性の評価は副次目的とした。本試験の有効性評価項目は、客観的奏効を認めた被験者の割合 (主要評価項目)、客観的奏効が 4、12、又は 24 週間以上持続した被験者の割合、各時期における客観的奏効を認めた被験者の割合、各画像評価における客観的奏効を認めた被験者の割合、客観的奏効までの期間、及び奏効期間 (副次評価項目)、並びにすべての時点の中での最良効果、疼痛スコアの変化、骨代謝マーカーの変化、及び臨床的有益性を認めた被験者の割合など (探索的評価項目) とした。腫瘍病変の画像評価は、CT、MRI、及び ^{18}F FDG-PET によって実施した。客観的抗腫瘍効果は中央読影機関において、3 つの基準 (modified RECIST 基準、modified EORTC 基準、及び density/size 基準) によって評価した。薬物動態評価項目は、血清中デノスマブ濃度トラフ値とした。安全性は、有害事象の種類、頻度、及び重症度、並び

に臨床検査値異常によって評価した。

結果

20 年 月 日（事前に計画された 6 ヶ月カットオフ日）までに、17 名の被験者が試験 AMG162-B-J201 に組み入れられ、全被験者が 1 回以上のデノスマブ投与を受けた。このうち、13 名が 6 ヶ月以上のデノスマブ投与を受けた。

あらゆる評価基準を用いた最良効果に基づく客観的奏効率は 82.4% (14/17) であり、個々の評価基準による客観的奏効率は、modified RECIST 基準で 29.4% (5/17)、modified EORTC 基準で 70.6% (12/17)、density/size 基準で 64.7% (11/17) であった。客観的奏効が 4、12、又は 24 週間持続した被験者の割合は、それぞれ 85.7% (12/14)、83.3% (10/12)、又は 100% (7/7) であった。最良効果に基づく客観的奏効率の経時推移について Kaplan-Meier 法で推定すると、第 25 週で 87.4% であり、あらゆる評価基準を用いた最良効果に基づく客観的奏効までの期間の中央値 (95% CI) は 3.0 ヶ月 (2.9~3.1) であった。また、あらゆる評価基準を用いた最良効果による CR は 17.6% (3 名)、PR は 64.7% (11 名)、SD は 17.6% (3 名) であり、いずれの評価基準又はいずれの評価時点でも PD と判定された被験者はいなかった。治験責任医師が判断した臨床的有益性が得られた被験者の割合は 82.4% (14/17) であった。痛みの改善に関しては、半数以上の被験者で、第 5 週以降に臨床的に意味のある最もひどい痛みの減少が認められた。また、骨代謝マーカー (uNTX/Cr) は第 5 週目以降の評価時点で約 80% の抑制が認められた。その他の骨代謝マーカー (CTX1、BSAP、オステオカルシン、及び TRAP 5b) も減少し、試験期間を通じてベースラインを下回った (オステオカルシンの第 5 週を除く)。負荷投与終了時 (第 5 週)、血清中デノスマブ濃度の平均値及び中央値は初回投与後の約 2.5 倍であり、その後はほぼ一定となったことから、投与後 1 ヶ月で定常状態に達したことが示された。

17 名全員が安全性の解析対象集団に含まれ、6 ヶ月カットオフ日 (20 年 月 日) までに、16 名 (94.1%) で少なくとも 1 件の有害事象が発現した。2 名以上に認められた有害事象 (10% 以上の被験者に発現) は、鼻咽頭炎 nasopharyngitis 4 名 (23.5%)、注射部位反応 injection site reaction 4 名 (23.5%)、倦怠感 malaise 4 名 (23.5%)、齲歯 dental caries 3 名 (17.6%)、膀胱炎 cystitis 2 名 (11.8%)、発熱 pyrexia 2 名 (11.8%)、及び悪心 nausea 2 名 (11.8%) であった。重篤な有害事象が発現した被験者は 2 名であった。CTCAE グレード 3 又は 4 の有害事象が 2 名に発現し、疼痛 pain 及び気胸 pneumothorax であった。死亡は認められず、有害事象によって試験を中止又は治験薬投与を中止した被験者は認められなかった。低カルシウム血症 hypocalcemia の有害事象は 1 名に認められ、非重篤、中等度の有害事象であった。ONJ と判定された有害事象又は新規原発性悪性腫瘍は認められなかった。血清中カルシウム及びリンの変化を除き、臨床検査値の変化は認められなかった。バイタルサインへの影響を示す変化は観察されなかった。抗デノスマブ抗体が陽性の被験者は認められなかった。

結論

デノスマブの投与を受けた GCTB 患者 17 名中 14 名 (82.4%) で客観的奏効が認められた。

その他の有効性評価項目でも、デノスマブの良好な効果が示された。本試験での血清中デノスマブ濃度トラフ値は、同じ用法・用量による他の試験と同程度であった。日本人 GCTB 患者におけるデノスマブの忍容性は概ね良好であり、安全性プロファイルは、これまでに確認されたデノスマブ 120 mg Q4W の皮下投与時のプロファイルと同様であった。以上のことから、GCTB 患者におけるデノスマブの高い臨床的意義が示唆された。

2.4 試験 AMG162-B-J201（国内第 II 相臨床試験）：12 ヶ月カットオフ解析

GCTB 患者を対象とした AMG 162（デノスマブ）の非盲検多施設共同第 II 相臨床試験

方法

第 2.3 項に示した。

結果

20 年 月 日（事前に計画された 12 ヶ月カットオフ日）までに、17 名の被験者が試験 AMG162-B-J201 に組み入れられ、全被験者が 1 回以上のデノスマブ投与を受けた。このうち、10 名が 12 ヶ月以上のデノスマブ投与を受けた。12 ヶ月カットオフ日時点で、1 名が試験及びデノスマブの投与を中止し、16 名が試験を継続中であった。なお、データカットオフ以降、20 年 月 日時点で、さらに 1 名が試験及びデノスマブの投与を中止している。

あらゆる評価基準を用いた最良効果に基づく客観的奏効率は 88.2%（15/17）であり、個々の評価基準による客観的奏効率は、modified RECIST 基準で 35.3%（6/17）、modified EORTC 基準で 82.4%（14/17）、density/size 基準で 70.6%（12/17）であった。客観的奏効が 4、12、又は 24 週間持続した被験者の割合は、それぞれ 88.2%（15/17）、82.4%（14/17）、又は 86.7%（13/15）であった。最良効果に基づく客観的奏効率の経時推移について Kaplan-Meier 法で推定すると、第 25 週で 82.4%、第 49 週で 88.2%であり、あらゆる評価基準を用いた最良効果に基づく客観的奏効までの期間の中央値（95% CI）は 3.0 ヶ月（2.9～3.1）であった。また、あらゆる評価基準を用いた最良効果による CR は 23.5%（4 名）、PR は 64.7%（11 名）、SD は 11.8%（2 名）であった。あらゆる評価基準を用いて客観的奏効が認められた 15 名のうち、1 名の被験者がその後 PD と判定された。また、個々の評価基準によって客観的奏効の後に PD と判定された被験者は、modified RECIST 基準で 0/6 名、modified EORTC 基準で 1/14 名、density/size 基準で 1/12 名であった。あらゆる評価基準を用いて客観的奏効の後に PD と判定された被験者では、CT/MRI により新たな病変（神経膠芽細胞腫 glioblastoma）が認められており、このため治験責任医師は、この被験者を PD ではないと判定した。治験責任医師が判断した臨床的有益性が得られた被験者の割合は 82.4%（14/17）であった。痛みの改善に関しては、半数以上の被験者で、第 5 週及びその後の全評価時点で臨床的に意味のある最もひどい痛みの減少が認められた。また、骨代謝マーカー（uNTX/Cr）は第 5 週目以降の評価時点で約 80%の抑制が認められた。その他の骨代謝マーカー（CTX1、BSAP、オステオカルシン、及び TRAP 5b）も減少し、試験期間を通じてベースラインを下回った（オステオカルシンの第 5 週を除く）。

17名全員が安全性の解析対象集団に含まれ、12ヵ月カットオフ日（20■■年■■月■■日）までに、すべての被験者（100.0% [17/17]）で1件以上の有害事象が認められた。2名以上に認められた有害事象（10%以上の被験者に発現）は、鼻咽頭炎 nasopharyngitis（29.4% [5/17]）、齲歯 dental caries（23.5% [4/17]）、インフルエンザ influenza（23.5% [4/17]）、注射部位反応 injection site reaction（23.5%[4/17]）、倦怠感 malaise（23.5%[4/17]）、悪心 nausea（17.6%[3/17]）、発熱 pyrexia（17.6% [3/17]）、関節痛 arthralgia（11.8% [2/17]）、膀胱炎 cystitis（11.8% [2/17]）、頭痛 headache（11.8% [2/17]）、歯周炎 periodontitis（11.8% [2/17]）、気胸 pneumothorax（11.8% [2/17]）、及び歯痛 toothache（11.8% [2/17]）であった。重篤な有害事象が発現した被験者は4名であった。CTCAE グレード3、4、又は5の有害事象が4名に発現し、グレード3の気胸 pneumothorax が2名、グレード3の疼痛 pain、及びグレード3の神経膠芽細胞腫 glioblastoma が各1名であった。死亡は認められず、有害事象によって試験を中止又は治験薬投与を中止した被験者は認められなかった。低カルシウム血症 hypocalcemia の有害事象は1名に認められ、非重篤、中等度の有害事象であった。ONJ と判定された有害事象は認められず、新規原発性悪性腫瘍（神経膠芽細胞腫 glioblastoma）が1名に認められた。血清中カルシウム及びリンの変化を除き、臨床検査値の変化は認められなかった。バイタルサインへの影響を示す変化は観察されなかった。抗デノスマブ抗体が陽性の被験者は認められなかった。12ヵ月データカットオフ日以降に、神経膠芽細胞腫 glioblastoma を発現した1名が試験及び治験薬の投与を中止し、また、1名が ONJ と判定された。

結論

デノスマブの投与を受けた GCTB 患者 17 名中 15 名（88.2%）で客観的奏効が認められた。その他の有効性評価項目でも、デノスマブの良好な効果が示された。本試験での血清中デノスマブ濃度トラフ値は、同じ用法・用量による他の試験と同程度であった。日本人 GCTB 患者におけるデノスマブの忍容性は概ね良好であり、本解析において新たに特定されたリスクは認められなかった。また、12ヵ月カットオフ解析の結果は、6ヵ月カットオフ解析の結果と大きな違いは認められなかった。

3. 全試験を通しての結果の比較と解析

本項では、試験 20040215、20062004、及び AMG162-B-J201（6 ヶ月カットオフ解析）における客観的抗腫瘍効果に関する独立した画像評価の結果を示す。また、外科的に切除可能な GCTB 患者に対するデノスマブの有効性の結果を第 3.3 項に、客観的抗腫瘍効果評価結果を裏付ける有効性の結果として、個々の第 II 相試験のその他の解析結果を第 3.4 項に示す。なお、試験 AMG162-B-J201 の 12 ヶ月カットオフ解析の結果は、第 5 項に示した。

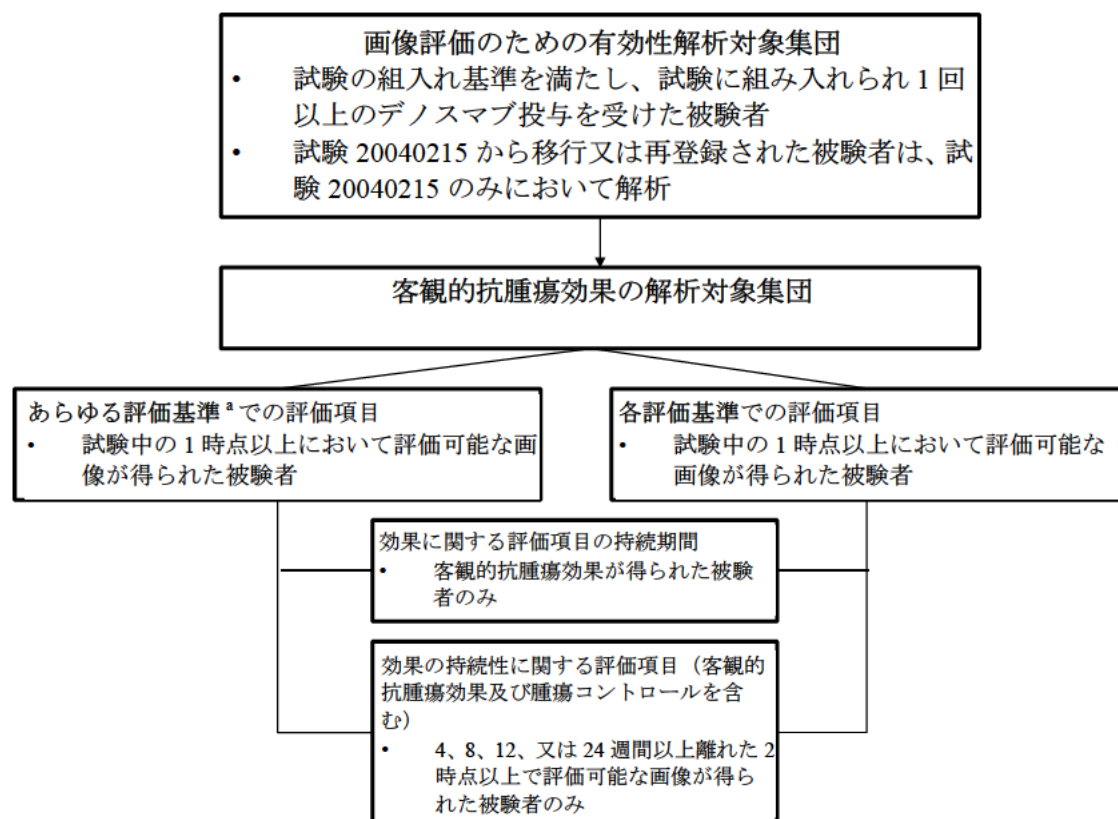
3.1 試験対象集団

3.1.1 客観的抗腫瘍効果の解析対象集団

客観的抗腫瘍効果の外国 2 試験併合解析のための有効性解析対象集団には、試験 20040215 又は 20062004 の組み入れ基準を満たし、試験に組み入れられ 1 回以上のデノスマブ投与を受けたすべての被験者（n = 303）を含めた。客観的抗腫瘍効果の解析対象集団（n = 190）には、有効性解析対象集団のうち、試験中の少なくとも 1 時点で評価可能な画像が得られたすべての被験者を含めた。あらゆる評価基準（modified RECIST 基準、modified EORTC 基準、density/size 基準）又は各評価基準を用いた最良効果の評価には、客観的抗腫瘍効果の解析対象集団を用いた（図 2.7.3.3-1）。

図 2.7.3.3-1 画像評価のための有効性解析対象集団及び客観的抗腫瘍効果の解析対象集団:

外国 2 試験併合解析



^a Modified RECIST 基準、modified EORTC 基準、又は density/size 基準

外国 2 試験併合解析における客観的抗腫瘍効果の評価はレトロスペクティブな評価であり、画像の評価は試験 20062004 の治験実施計画書に規定されていなかったため、一部の被験者では同意が得られなかったか、すべての画像データが得られなかったことから、抗腫瘍効果の評価できなかった。除外理由はいずれも、疾患の転帰に関連したものではないと考えられたため、解析対象集団には評価可能な被験者のみを含めることとした。同様に、抗腫瘍効果の持続性及び腫瘍コントロールの持続性の解析対象集団には、これらの評価項目の評価が可能な画像データが得られた被験者を含めた。

画像コントロール群の解析対象集団は、中央読影機関による評価に用いる投与開始前の画像データが 3 時点以上で得られた被験者 26 名とした。なお、画像コントロール群の投与開始前の画像について、画像間の期間、画像様式、又は測定手法に関する規定は設定しなかった。

試験 AMG162-B-J201 の有効性解析対象集団及び客観的抗腫瘍効果の解析対象集団の定義は、基本的に外国 2 試験併合解析の場合と同様である。いずれの解析対象集団にも、試験に組み入れられた全被験者（n = 17）が含まれた。3 試験併合解析では、外国 2 試験併合解析と試験 AMG162-B-J201 の各解析対象集団を併合し、有効性の併合解析対象集団（n = 320）及び客観的抗腫瘍効果の併合解析対象集団（n = 207）とした。外国 2 試験併合解析（図 2.7.3.3-1）と同様、あらゆる評価基準（modified RECIST 基準、modified EORTC 基準、density/size 基準）又は

各評価基準を用いた最良効果の評価には、客観的抗腫瘍効果の解析対象集団を用いた。試験 AMG162-B-J201 では、画像コントロール群を設定しなかった。

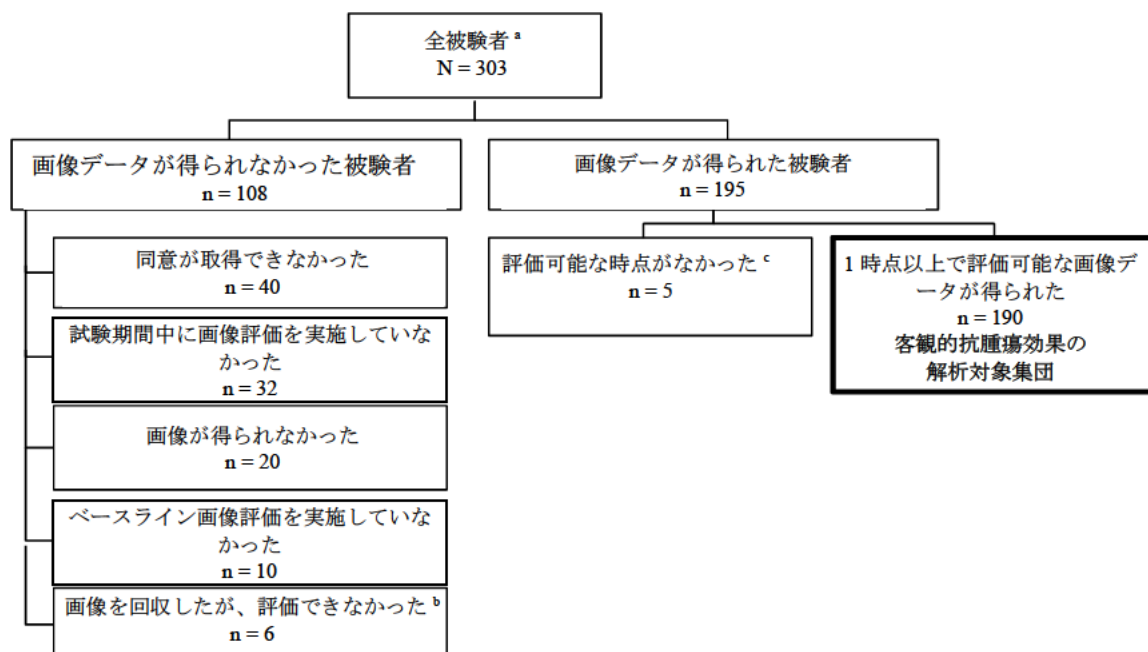
3.1.2 客観的抗腫瘍効果の解析における被験者の内訳

外国 2 試験併合解析での独立した画像評価のための画像データの収集対象は、試験 20040215 の完了時点における 37 名及び試験 20062004 の事前に計画した第 3 回中間解析のデータカットオフ日時点における 266 名の合計 303 名であった。試験 20062004 の 2 名が、有効性解析対象集団に含まれなかった。1 名はデノスマブ投与を受けておらず、1 名は試験の組み入れ基準を満たさなかった。試験 20040215 由来の試験 20062004 被験者（移行した被験者 [コホート 3; n = 11] 及び再登録被験者 [n = 3]) は、試験 20040215 で収集された画像データのみを用いて評価した。

303 名中 108 名（試験 20040215 の 10 名及び試験 20062004 の 98 名）は、[Listing tie1.1.5](#) に示した理由で画像データが得られなかった。主な理由は、画像評価に関する同意を取得できなかった (37.0% [40/108])、試験期間中に画像評価を実施していなかった (29.6% [32/108])、及び画像が全く得られなかった (18.5% [20/108]) などであった ([Table tie1.1.2](#))。これはレトロスペクティブな評価であったため、約 1/3 の被験者で画像データが利用できないことは、想定範囲であった。なお、画像データが評価可能な被験者と画像データがない被験者で、個々の試験で治験責任医師が評価した治療効果（試験 20040215 では治験実施計画書に定められた病理組織学検査／画像検査上の効果、試験 20062004 では病勢の進行までの期間、第 6 ヶ月の時点で手術が施行されなかった被験者の割合、及び疾患の状態）に関する結果は類似していた（事後解析; [Table tie4.4.503](#)、[Table tie4.4.504](#)、[Table tie4.4.505](#)、及び [Table tie4.4.506](#)）。

1 時点以上の評価可能な画像データが得られた 190 名が、外国 2 試験併合解析の客観的抗腫瘍効果の解析対象集団に含まれた。画像データ回収状況の内訳を [図 2.7.3.3-2](#) 及び [Table tie1.1.2](#) に示す。独立した画像評価の対象としたすべての被験者（評価可能な被験者と評価不能な被験者を含む）の臨床的叙述をモジュール 5、第 5.3.5.3 項（[画像の叙述](#)）に示す。

図 2.7.3.3-2 画像データの有無による被験者の内訳: 外国 2 試験併合解析



^a 試験 20040215 から試験 20062004 に移行又は再登録した被験者は、試験 20040215 においてのみ画像解析に組み入れた。

^b X 線のみ

^c 画質のため病変が評価不能、又は被験者が評価時点の前に外科的切除を受けていたことが判明。

出典 Table tiael.1.1.1; Table tiael.1.1.2

客観的抗腫瘍効果の解析対象集団のうち、38 名（20.0%）が試験を中止した。試験中止の理由を表 2.7.3.3-1 に示す。また、39 名（20.5%）がデノスマブの投与を中止した（Table tiael.1.2.2）。

試験 AMG162-B-J201 では、全被験者（n = 17）が有効性解析対象集団及び客観的抗腫瘍効果の解析対象集団に含まれた。3 試験併合解析は、有効性の併合解析対象集団（n = 320）及び客観的抗腫瘍効果の併合解析対象集団（n = 207）を用いて実施した。試験 AMG162-B-J201 では、6 ヶ月データカットオフ日までにデノスマブの投与を中止した被験者はいなかった。

表 2.7.3.3-1 試験中止理由: 試験 20040215 及び 20062004

(1 時点以上で評価可能な画像データが得られた被験者)

(Efficacy Analysis Set)

	Study 20040215 n (%)	Study 20062004			Overall n (%)
		Cohort 1 n (%)	Cohort 2 n (%)	Cohorts 1 and 2 n (%)	
Subjects with at least one evaluable time point assessment	27	114	49	163	190
Ongoing	0 (0.0)	108 (94.7)	44 (89.8)	152 (93.3)	152 (80.0)
Discontinued study	27 (100.0)	6 (5.3)	5 (10.2)	11 (6.7)	38 (20.0)
Protocol-specified criteria ^a	8 (29.6)	2 (1.8)	3 (6.1)	5 (3.1)	13 (6.8)
Rollover to other study	9 (33.3)	0 (0.0)	0 (0.0)	0 (0.0)	9 (4.7)
Administrative decision	1 (3.7)	1 (0.9)	2 (4.1)	3 (1.8)	4 (2.1)
Other ^b	4 (14.8)	0 (0.0)	0 (0.0)	0 (0.0)	4 (2.1)
Adverse event	1 (3.7)	2 (1.8)	0 (0.0)	2 (1.2)	3 (1.6)
Disease progression	2 (7.4)	1 (0.9)	0 (0.0)	1 (0.6)	3 (1.6)
Consent withdrawn	1 (3.7)	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.5)
Noncompliance	1 (3.7)	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.5)

Page 1 of 1

Percentages based on number of enrolled subjects who were eligible for the study, received at least one dose of denosumab, and had at least one evaluable time point assessment

^a Complete resection

^b Discontinued study at investigator's discretion.

Modified from source: Table tiae1.2.1 and Listing 1.1.502.

3.1.3 客観的抗腫瘍効果の解析における人口統計学的特徴及びベースラインの疾患特性

客観的抗腫瘍効果の解析対象集団の被験者における人口統計学的特徴及び特定のベースラインの疾患特性を表 2.7.3.3-2 及び表 2.7.3.3-3 に示す。

試験 20040215 及び 20062004（外国 2 試験併合解析）

被験者の 55%が女性であった。大多数の被験者は白人であった（78.9%）。年齢の中央値（範囲）は 33（13～76）歳であり、6 名（3.2%）は 18 歳未満であった。ベースライン時の身体的特徴（身長、体重、及び体格指数）を Table tie2.1.2 に要約する。

ほとんどの被験者（95.8%）は、試験組み入れ時の Eastern Cooperative Oncology Group (ECOG) performance status が 0 又は 1 であった。

客観的抗腫瘍効果を評価した被験者の約 29%は、初発又は再発の切除可能な GCTB を有し、71%は、初発又は再発の切除不能な GCTB であった（表 2.7.3.3-3）。個々の試験において治験責任医師が選択した標的病変を表 2.7.3.3-3 に示す。なお、客観的抗腫瘍効果の評価に用いられた病変は、中央読影機関が独自に最も評価に適した病変を選択したため、治験責任医師が選択・評価した標的病変と異なる可能性がある。治験責任医師が選択した標的病変で最も多かったのは骨盤（32.1%）であった。次に多かったのは「その他」（25.8%）であった。「その他」に含まれた標的病変部位は、主に肺であった。また、骨盤領域、後大腿部軟部組織腫瘍、小菱形骨、腹膜後腔、錐体骨、右後胸膜、肘、頸部軟部組織、舌骨、踵骨及び指などが含まれた（Listing tie2.1.500.1）。

痛みの評価は試験 20062004 においてのみ実施した。ベースライン時点における最もひどい痛みは、客観的抗腫瘍効果の解析対象集団（1 時点以上で評価可能な画像データが得られた被験者）で、痛みの評価データが得られたコホート 1 及び 2 の被験者のうち、約半数（54.5%、[84/154]）ではなし又は軽度（スコア: 0～4 ポイント）、28.6%（44/154）では重度（スコア: 7～10 ポイント）であった（Table tie2.1.5）。0 ポイント（「痛みなし」）から 10 ポイント（「想像できる最悪の痛み」）までのスケールで、ベースライン時の最もひどい痛みのスコアの平均値（標準偏差）は、コホート 1 と 2 の全体で 4.1（3.0）であった（Table tie2.1.4）。ほとんどの被験者（72.3% [118/163]）では、ベースライン時の鎮痛薬スコアは 0 ポイント又は低値（2 ポイント以下）であった（Table tie2.1.6）。ベースライン時の鎮痛薬スコア（スケール: 0 ポイント [鎮痛薬なし] ～7 ポイント [600 mg/日超の経口モルヒネ相当]）の平均値（標準偏差）は 1.9（2.5）であった（Table tie2.1.6）。

試験 AMG162-B-J201

被験者 17 名中 9 名（52.9%）が女性であった。年齢の中央値（範囲）は 30（18～66）歳であり、1 名が 18 歳であった。ベースライン時の身体的特徴（身長、体重、及び体格指数）を Table 15.1-11（試験 AMG162-B-J201 治験総括報告書）に要約する。

ほとんどの被験者（88.2%）は、試験組み入れ時の ECOG performance status が 0 又は 1 であ

った。

被験者 17 名中 4 名 (23.5%) は、初発又は再発の切除可能な GCTB を有し、13 名 (76.5%) は、初発又は再発の切除不能な GCTB であった (表 2.7.3.3-3)。治験責任医師が選択した標的病変で最も多かったのは骨盤 (35.3%) であり、次に多かったのは「その他」 (23.5%) であった (表 2.7.3.3-3)。「その他」に含まれた標的病変部位は、主に肺であった。なお、客観的抗腫瘍効果の評価では、中央読影機関が独自に最も評価に適した病変を選択したため、治験責任医師が選択・評価した標的病変と異なる可能性がある。

ベースライン時点における最もひどい痛みは、76.5% (13/17) ではなし又は軽度 (スコア: 0 ~4 ポイント)、5.9% (1/17) では重度 (スコア: 7~10 ポイント) であった (試験 AMG162-B-J201 治験総括報告書 Table 15.1-19)。ベースライン時の最もひどい痛みのスコアの平均値 (標準偏差) は 2.2 (2.5) であった (試験 AMG162-B-J201 治験総括報告書 Table 15.1-17)。ほとんどの被験者 (82.4%、14/17) では、ベースライン時の鎮痛薬スコアは 0 ポイント又は低値 (2 ポイント以下) であった (Table 15.1-27)。ベースライン時の鎮痛薬スコア (スケール: 0 ポイント [鎮痛薬なし] ~7 ポイント [600 mg/日超の経口モルヒネ相当]) の平均値 (標準偏差) は 1.4 (2.4) であった (Table 15.1-27)。

表 2.7.3.3-2 ベースライン時の人口統計学的特性（記述統計量）

（1 時点以上で評価可能な画像データが得られた被験者）

（Efficacy Analysis Set）

	Study 20040215 (N = 27)	Study 20062004			215/2004 Overall (N = 190)	Study AMG162-B-J201 (N = 17)
		Cohort 1 (N = 114)	Cohort 2 (N = 49)	Cohorts 1 and 2 (N = 163)		
Sex - n (%)						
Female	15 (55.6)	66 (57.9)	24 (49.0)	90 (55.2)	105 (55.3)	9 (52.9)
Male	12 (44.4)	48 (42.1)	25 (51.0)	73 (44.8)	85 (44.7)	8 (47.1)
Ethnic group / race - n (%)						
White or Caucasian	21 (77.8)	86 (75.4)	43 (87.8)	129 (79.1)	150 (78.9)	0 (0.0)
Black or African American	1 (3.7)	9 (7.9)	2 (4.1)	11 (6.7)	12 (6.3)	0 (0.0)
Hispanic or Latino	2 (7.4)	11 (9.6)	0 (0.0)	11 (6.7)	13 (6.8)	0 (0.0)
Asian	3 (11.1)	6 (5.3)	2 (4.1)	8 (4.9)	11 (5.8)	17 (100) ^a
Other	0 (0.0)	2 (1.8)	2 (4.1)	4 (2.5)	4 (2.1)	0 (0.0)
Age (years)						
n	27	114	49	163	190	17
Mean	36.2	34.8	34.7	34.7	34.9	36.3
SD	12.5	12.9	11.9	12.6	12.6	15.6
Median	31.0	33.0	34.0	33.0	33.0	30.0

Page 1 of 2

N = number of enrolled subjects who were eligible for the study, received at least one dose of denosumab, and had at least one evaluable time point assessment

^a Japanese

For Studies 20040215 and 20062004,



Source: CSR Table 15.1-9 for Study AMG162-B-J201

表 2.7.3.3-2 ベースライン時の人口統計学的特性（記述統計量）

（1 時点以上で評価可能な画像データが得られた被験者）

（Efficacy Analysis Set）（続き）

	Study 20040215 (N = 27)	Study 20062004			215/2004 Overall (N = 190)	Study AMG162-B-J201 (N = 17)
		Cohort 1 (N = 114)	Cohort 2 (N = 49)	Cohorts 1 and 2 (N = 163)		
Age (years)						
Q1, Q3	27.0, 46.0	26.0, 41.0	25.0, 43.0	25.0, 43.0	26.0, 43.0	24.0, 47.0
Min, Max	19, 63	13, 76	18, 69	13, 76	13, 76	18, 66
Age group - n (%)						
< 18 Years	0 (0.0)	6 (5.3)	0 (0.0)	6 (3.7)	6 (3.2)	0 (0.0)
18 - 40 Years	20 (74.1)	77 (67.5)	34 (69.4)	111 (68.1)	131 (68.9)	10 (58.8)
41 - 60 Years	6 (22.2)	26 (22.8)	14 (28.6)	40 (24.5)	46 (24.2)	5 (29.4)
> 60 Years	1 (3.7)	5 (4.4)	1 (2.0)	6 (3.7)	7 (3.7)	2 (11.8)
Geriatric age group - n (%)						
≥ 65 years	0 (0.0)	4 (3.5)	1 (2.0)	5 (3.1)	5 (2.6)	1 (5.9)
≥ 75 years	0 (0.0)	1 (0.9)	0 (0.0)	1 (0.6)	1 (0.5)	0 (0.0)

Page 2 of 2

N = number of enrolled subjects who were eligible for the study, received at least one dose of denosumab, and had at least one evaluable time point assessment

For Studies 20040215 and 20062004,



Source: CSR Table 15.1-9 for Study AMG162-B-J201

表 2.7.3.3-3 CRF データに基づくベースライン時の疾患特性（記述統計量）

（1 時点以上で評価可能な画像データが得られた被験者）

（Efficacy Analysis Set）

	Study 20040215 (N = 27)	Study 20062004 Cohort 1 (N = 114)	Cohort 2 (N = 49)	Cohorts 1 and 2 (N = 163)	215/2004 Overall (N = 190)	Study AMG162-B-J201 (N = 17)
ECOG Status - n(%)						
0	9 (33.3)	73 (64.0)	24 (49.0)	97 (59.5)	106 (55.8)	9 (52.9)
1	16 (59.3)	35 (30.7)	25 (51.0)	60 (36.8)	76 (40.0)	6 (35.3)
2	0 (0.0)	6 (5.3)	0 (0.0)	6 (3.7)	6 (3.2)	2 (11.8)
Missing	2 (7.4)	0 (0.0)	0 (0.0)	0 (0.0)	2 (1.1)	0 (0.0)
GCT disease type - n(%)						
Primary resectable	0 (0.0)	0 (0.0)	26 (53.1)	26 (16.0)	26 (13.7)	2 (11.8)
Primary unresectable	9 (33.3)	34 (29.8)	0 (0.0)	34 (20.9)	43 (22.6)	5 (29.4)
Recurrent resectable	6 (22.2)	0 (0.0)	23 (46.9)	23 (14.1)	29 (15.3)	2 (11.8)
Recurrent unresectable	12 (44.4)	80 (70.2)	0 (0.0)	80 (49.1)	92 (48.4)	8 (47.1)

Page 1 of 2

N = number of enrolled subjects who were eligible for the study, received at least one dose of denosumab, and had at least one evaluable time point assessment
Target lesion was identified and evaluated by investigators through imaging procedures including X-ray, CT, MRI etc.

Modified from Source: Table tiae2.1.3 and Listing tiae2.1.500.1 for Studies 20040215 and 20062004; and CSR Table 15.1-14 for Study AMG162-B-J201

表 2.7.3.3-3 CRF データに基づくベースライン時の疾患特性（記述統計量）

（1 時点以上で評価可能な画像データが得られた被験者）

（Efficacy Analysis Set）（続き）

	Study 20040215 (N = 27)	Study 20062004 Cohort 1 (N = 114)	Cohort 2 (N = 49)	Cohorts 1 and 2 (N = 163)	215/2004 Overall (N = 190)	Study AMG162-B-J201 (N = 17)
Longest dimension of target lesion (mm)						
n	26	114	49	163	189	17
Mean	52.5	62.4	69.7	64.6	63.0	70.6
SD	32.9	43.7	32.0	40.6	39.8	58.6
Median	44.5	50.5	70.0	60.0	56.4	44.0
Q1, Q3	29.0, 70.0	29.0, 90.0	49.9, 83.0	34.0, 89.0	33.8, 85.0	28.9, 98.3
Min, Max	6, 130	7, 240	10, 200	7, 240	6, 240	15, 228
Location of target lesion - n(%)						
Pelvis	7 (25.9)	46 (40.4)	8 (16.3)	54 (33.1)	61 (32.1)	6 (35.3)
Other ^a	8 (29.6)	36 (31.6)	5 (10.2)	41 (25.2)	49 (25.8)	4 (23.5)
Lower extremities	6 (22.2)	7 (6.1)	26 (53.1)	33 (20.2)	39 (20.5)	3 (17.6)
Spine	2 (7.4)	14 (12.3)	2 (4.1)	16 (9.8)	18 (9.5)	2 (11.8)
Upper extremities	3 (11.1)	6 (5.3)	8 (16.3)	14 (8.6)	17 (8.9)	1 (5.9)
Head/neck	0 (0.0)	5 (4.4)	0 (0.0)	5 (3.1)	5 (2.6)	1 (5.9)
Missing	1 (3.7)	0 (0.0)	0 (0.0)	0 (0.0)	1 (0.5)	0 (0.0)

Page 2 of 2

N = number of enrolled subjects who were eligible for the study, received at least one dose of denosumab, and had at least one evaluable time point assessment
Target lesion was identified and evaluated by investigators through imaging procedures including X-ray, CT, MRI etc.

^a The category of “other” included target lesions that were predominantly located in the lung. Other regions included pelvic region, a retrocrural soft tissue mass, trapezoid, retroperitoneum, petrous bone, right posterior pleura, cubital, cervical soft tissue, hyoid bone, calcaneum, and finger.

Modified from Source: Table tiae2.1.3 and Listing tiae2.1.500.1 for Studies 20040215 and 20062004; and CSR Table 15.1-14 for Study AMG162-B-J201

3.1.4 客観的抗腫瘍効果の解析対象集団におけるデノスマブの曝露状況

客観的抗腫瘍効果の解析対象集団における試験期間の中央値は、外国 2 試験併合解析では 13.4 カ月 (Table 15.3-1.1)、試験 AMG162-B-J201 では 7.56 カ月 (治験総括報告書 Table 15.3-1.1) であった。デノスマブ投与回数の中央値 (範囲) は、外国 2 試験併合解析では 16 (4~54) 回、試験 AMG162-B-J201 では 11 (6~16) 回であった。客観的抗腫瘍効果を評価した被験者における試験期間は Table 15.3-1.2 及び Table 15.3-1.1 に要約した。

3.2 客観的抗腫瘍効果の解析結果

3.2.1 有効性主要評価項目: 客観的奏効

外国 2 試験併合解析では、ベースライン評価時のほか試験期間中の少なくとも 1 時点の評価可能な画像データを有する合計 190 名の被験者を、あらゆる評価基準を用いた最良効果の評価対象とした。これらの被験者のうち、modified RECIST 基準、modified EORTC 基準、及び density/size 基準による評価が可能であり、これら評価の対象となった被験者は、それぞれ 187 名、26 名、及び 176 名であった (Table 15.3-1.1)。2 つ以上の評価基準で評価可能な被験者も含まれた。RECIST 基準では非標的病変も評価を行うのに対し、density/size 基準では非標的病変は評価を行わないため、評価可能であった被験者は、modified RECIST 基準による評価の方が多かった。

あらゆる評価基準を用いた最良効果に基づき客観的奏効 (CR 又は PR) を認めた被験者の割合 (客観的奏効率) は 71.6% (136/190) であり、個々の評価基準による客観的奏効率は、modified RECIST 基準で 25.1% (47/187)、modified EORTC 基準で 96.2% (25/26)、density/size 基準で 76.1% (134/176) であった (表 2.7.3.3-4)。客観的奏効率は試験間で同程度であり、試験 20062004 のコホート 1 とコホート 2 でも同程度であった。最良効果に基づく客観的奏効率は、試験 20040215 で 74.1% (再発又は切除不能な GCTB: 評価可能な被験者 27 名中 20 名)、試験 20062004 のコホート 1 で 66.7% (切除不能な GCTB: 評価可能な被験者 114 名中 76 名)、コホート 2 で 81.6% (切除可能な GCTB: 評価可能な被験者 49 名中 40 名) であった (Table 15.3-1.1)。病変を切除する外科的処置が施行されたコホート 2 の被験者では、切除後は評価不可能とし、病勢の進行についてのみ追跡調査した。ただし、非標的病変のみを切除した場合は、標的病変の RECIST 基準による評価を可能とした。

試験 AMG162-B-J201 では全被験者 17 名を、あらゆる評価基準を用いた最良効果の評価対象とした。いずれの被験者も、modified RECIST 基準、modified EORTC 基準、及び density/size 基準による評価が可能であり、各評価の対象とした (試験 AMG162-B-J201 治験総括報告書 Table 15.1-2)。あらゆる評価基準を用いた最良効果に基づく客観的奏効率は 82.4% (14/17) であり、個々の評価基準による客観的奏効率は、modified RECIST 基準で 29.4% (5/17)、modified EORTC 基準で 70.6% (12/17)、density/size 基準で 64.7% (11/17) であった (表 2.7.3.3-5)。これらの結果は、外国 2 試験併合解析の結果 (表 2.7.3.3-4) と類似していた。なお、3 試験併合解析の結果は Table 15.3-1.1 に示した。

表 2.7.3.3-4 客観的奏効率: 外国 2 試験併合解析
(1 時点以上で評価可能な画像データが得られた被験者)
(Efficacy Analysis Set)

	n	N1	Percent	95% CI ^a
Proportion of subjects with an objective tumor response (CR, PR)				
Based on best response	136	190	71.6	(64.6, 77.9)
RECIST 1.1	47	187	25.1	(19.1, 32.0)
EORTC	25	26	96.2	(80.4, 99.9)
Density/size	134	176	76.1	(69.1, 82.2)

Page 1 of 1

n = number of subjects with a response

N1 = number of subjects with at least one evaluable time point assessment using the respective tumor response criteria

^a Exact confidence interval

Source: Table tiae4.1.1; results by individual study are provided in source table.

表 2.7.3.3-5 客観的奏効率: 試験 AMG162-B-J201
(Efficacy Analysis Set)
(6-month Cutoff Analysis)

	n	N1	Percent	95% CI ^a
Proportion of subjects with an objective tumor response (CR, PR)				
Based on best response	14	17	82.4	(56.6, 96.2)
RECIST 1.1	5	17	29.4	(10.3, 56.0)
EORTC	12	17	70.6	(44.0, 89.7)
Density/size	11	17	64.7	(38.3, 85.8)

Page 1 of 1

n = number of subjects with a response

N1 = number of subjects with at least one evaluable time point assessment using the respective tumor response criteria

^a Exact confidence interval

Source: Table15.2-1.1.1

外国 2 試験併合解析の対象とした 190 名の多くでは、modified RECIST 基準と density/size 基準の両方による評価が可能であった。これらの評価による客観的奏効率の違いは (RECIST 25.1%、density/size 76.1%)、主に以下のことに起因した。

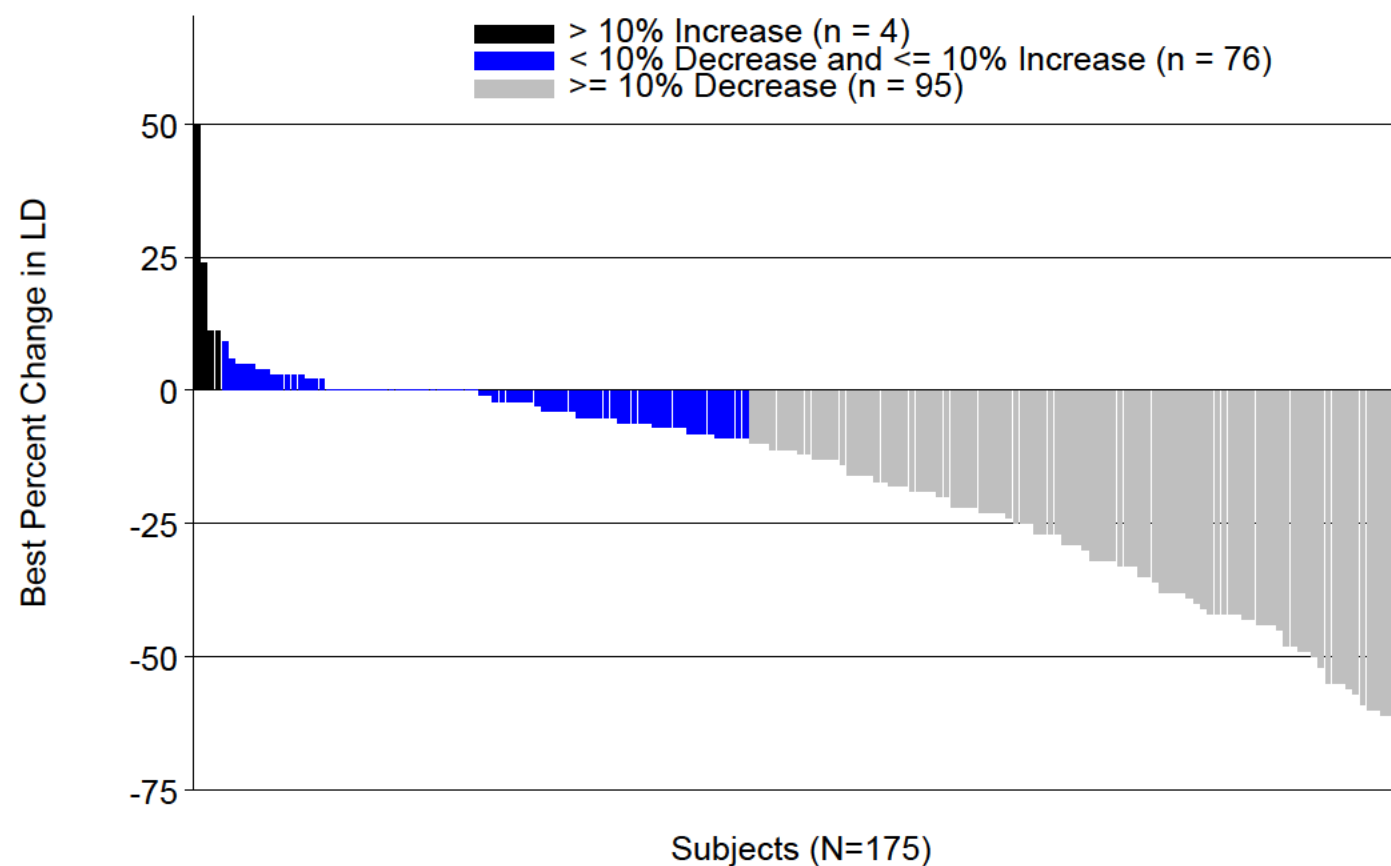
- 1) RECIST 基準では病変サイズを評価するが、density/size 基準では腫瘍密度と病変サイズの両方を評価する。Density/size 基準を用いた場合、密度評価のみで PR と判定されることがある。Density/size 基準による評価で奏効を認めた被験者を対象にした事後解析では、密度とサイズの両方に基づく奏効が 41.8%、密度評価のみに基づく奏効が 30.6%、サイズ評価のみに基づく奏効が 27.6%であった (Table tiae4.4.502)。Density/size 基準によって評価した標的病変における密度の合計の最大変化率を図 2.7.3.3-4 及び Table tiae4.9.7 に要約する。
- 2) Density/size 基準による評価では、サイズに基づく奏効の基準値が RECIST 基準に比べ

て低い。Modified RECIST 基準によって PR と判定されるためには、標的病変の最長径の和 (sum of the longest diameter: SLD) が 30%以上減少する必要がある。一方、density/size 基準による評価では、SLD の 10%以上の縮小で PR と判定される。Density/size 基準による病変径の合計の最大変化率を図 2.7.3.3-3 及び Table 4.9.6 に示す。Modified RECIST 基準による標的病変径の合計の最大変化率を Table 4.9.5 に要約する。

このように、modified RECIST 基準に比べて density/size 基準による評価で奏効率が高かったのは、サイズ評価の基準値の違いと、密度評価のみに基づく結果が反映されたためと考えられる。

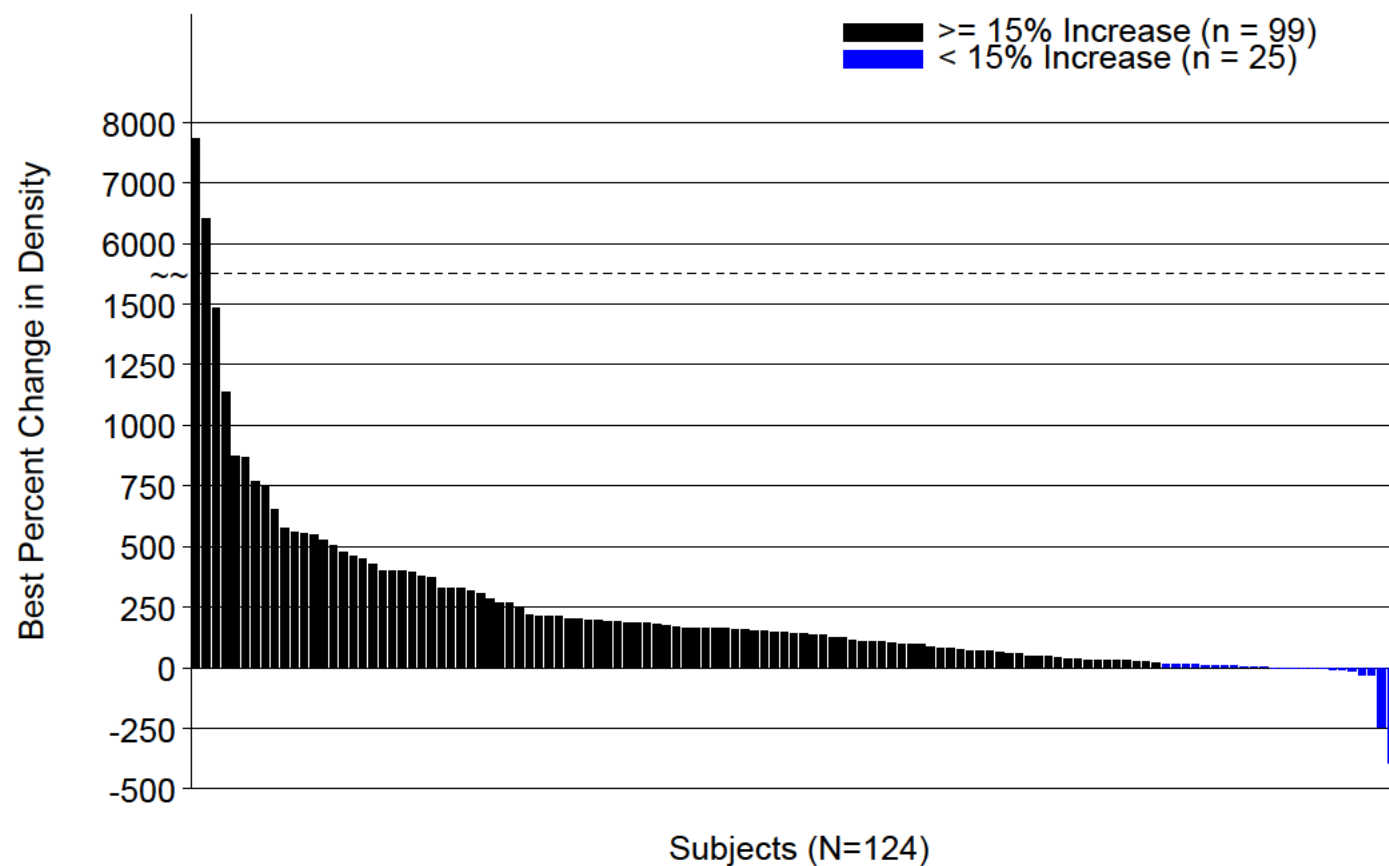
腫瘍部位の密度増加は、病変部位における密度の高い新たな線維性骨の形成を示しており、期待されたデノスマブの効果と一致する。疾患特性とデノスマブの作用機序を考慮すると、骨形成は期待された効果であり、これには、病変部位のサイズの縮小を伴う場合と伴わない場合があると考えられる。

図 2.7.3.3-3 Density/Size 基準による標的病変径の合計の最大変化率を示すウォーターフォールプロット: 外国 2 試験併合解析
(1 時点以上で評価可能な画像データが得られた被験者)



N = number of subjects with data

図 2.7.3.3-4 Density/Size 基準による標的病変密度の合計の最大変化率を示すウォーターフォールプロット：外国 2 試験併合解析
(1 時点以上で評価可能な画像データが得られた被験者)
(Efficacy Analysis Set)



N = number of subjects with data



外国 2 試験併合解析では、modified EORTC 基準による評価が可能な画像が得られた被験者数は、RECIST 基準又は density/size 基準による評価に比べて少数であったが、1 名を除くすべての被験者で客観的奏効が認められた（表 2.7.3.3-4）。この結果は、GCTB 病変における巨細胞の減少と、それに対応した腫瘍代謝活性の減少を示していると考えられる。

上述のように、骨病変では必ずしも病変サイズの縮小が認められない場合があり、軟部組織成分を含まない腫瘍に対する効果判定では、RECIST 基準には限界がある。Modified RECIST 基準を用いて標的病変のタイプ別（軟部組織〔又は軟部組織成分を含む病変〕又は骨組織）に事後解析を実施した結果、奏効率は骨組織病変で 15.4%（19/123）、軟部組織病変又は軟部組織成分を含む病変で 57.1%（28/49）であった（表 2.7.3.3-6）。このように、軟部組織病変の評価法として確立された RECIST 基準で評価した場合、デノスマブによる客観的奏効率は、軟部組織病変を有する被験者で高かった。骨組織病変を有する被験者では、病変サイズの縮小が認められた被験者は一部であったが、多くの被験者で腫瘍代謝活性の減少及び GCTB 病変の密度増加が認められた。また、骨組織病変を有する被験者では、RECIST 基準によって奏効を認めた 19 名全員において、density/size 基準での奏効が確認された。この 19 名のうち 14 名は密度に関する奏効基準（CT 密度の 15%以上の増加）を満たし、5 名は密度評価結果が得られなかった（ad hoc Table 4.4.512）。

なお、この病変タイプ別の解析では、軟部組織病変（又は軟部組織成分を含む病変）、骨組織病変、あるいは両成分を含む病変のいずれかを判断するのに十分な情報が得られた被験者を対象とした。また、これらの解析では、純粋な軟部組織病変を区別していないため、軟部組織成分を含む病変には骨成分が含まれている可能性があり、その場合、density/size 基準では新たな骨形成によって奏効と判定される可能性がある。

試験 AMG162-B-J201 では、modified RECIST 基準を用いて標的病変のタイプ別（軟部組織〔又は軟部組織成分を含む病変〕又は骨組織）に解析を実施した結果、奏効率は骨組織病変で 16.7%（2/12）、軟部組織病変又は軟部組織成分を含む病変で 60.0%（3/5）であった（表 2.7.3.3-7）。骨組織病変を有する被験者では、病変サイズの縮小が認められた被験者は一部であったが、多くの被験者で腫瘍代謝活性の減少及び density/size 基準での奏効が認められた。これらの結果は、外国 2 試験併合解析の結果（表 2.7.3.3-6）と同様の傾向を示した。

表 2.7.3.3-6 病変のタイプ別に示した客観的奏効率: 外国 2 試験併合解析

(1 時点以上で評価可能な画像データが得られた被験者)

(Efficacy Analysis Set)

	Soft Tissue Lesion or Lesion With Soft Tissue Component		Bone Lesion		Bone and Soft Tissue Containing Lesions	
	n/N1 (%)	95% CI ^a	n/N1 (%)	95% CI ^a	n/N1 (%)	95% CI ^a
Proportion of subjects with an objective tumor response (CR, PR)						
Based on best response	46/49 (93.9)	(83.1, 98.7)	89/126 (70.6)	(61.9, 78.4)	1/2 (50.0)	(1.3, 98.7)
RECIST 1.1	28/49 (57.1)	(42.2, 71.2)	19/123 (15.4)	(9.6, 23.1)	0/2 (0.0)	(0.0, 84.2)
EORTC	4/4 (100.0)	(39.8, 100.0)	21/21 (100.0)	(83.9, 100.0)	0/1 (0.0)	(0.0, 97.5)
Density/size	46/48 (95.8)	(85.7, 99.5)	87/125 (69.6)	(60.7, 77.5)	1/2 (50.0)	(1.3, 98.7)

Page 1 of 1

n = number of subjects with a response

N1 = number of subjects with at least one evaluable time point assessment with the respective lesion type using the respective tumor response criteria.

^a Exact confidence interval.

Source: Table tia4.4.501; results by individual study are provided in source table.

表 2.7.3.3-7 病変のタイプ別に示した客観的奏効率: 試験 AMG162-B-J201

(Efficacy Analysis Set)

(6-month Cutoff Analysis)

	Soft Tissue Lesion or Lesion With Soft Tissue Component		Bone Lesion		Bone and Soft Tissue Containing Lesions	
	n/N1 (%)	95% CI ^a	n/N1 (%)	95% CI ^a	n/N1 (%)	95% CI ^a
Proportion of subjects with an objective tumor response (CR, PR)						
Based on best response	5/5 (100.0)	(47.8, 100.0)	9/12 (75.0)	(42.8, 94.5)	0	(-, -)
RECIST 1.1	3/5 (60.0)	(14.7, 94.7)	2/12 (16.7)	(2.1, 48.4)	0	(-, -)
EORTC	4/5 (80.0)	(28.4, 99.5)	8/12 (66.7)	(34.9, 90.1)	0	(-, -)
Density/size	5/5 (100.0)	(47.8, 100.0)	6/12 (50.0)	(21.1, 78.9)	0	(-, -)

Page 1 of 1

n = number of subjects with a response

N1 = number of subjects with at least one evaluable time point assessment with the respective lesion type using the respective tumor response criteria.

^a Exact confidence interval.

Source: Table 15.2-1.14.1

個々の効果判定基準間の相関性

外国 2 試験併合解析データを対象として事後的な感度分析を実施し、客観的抗腫瘍効果の評価に用いた 3 つの評価方法の相関性を検討した。RECIST 基準による評価と density/size 基準による評価の間には、中程度の有意な相関が認められた。両基準で客観的奏効が認められた被験者は 27.2%であった (ϕ 相関係数 = 0.35; $P < 0.001$) (ad hoc Table tia4.4.520)。RECIST 基準と EORTC 基準での相関は弱かった (ϕ 相関係数 = 0.09; $P = 1.00$) (ad hoc Table tia4.4.519)。

Density/size 基準と EORTC 基準では中程度の相関が認められたが、有意ではなかった (ϕ 相関

係数 = 0.55; $P = 0.115$) (ad hoc [Table tie4.4.521](#))。これは、EORTC 基準の評価対象被験者数が少なかったためであると考えられる。

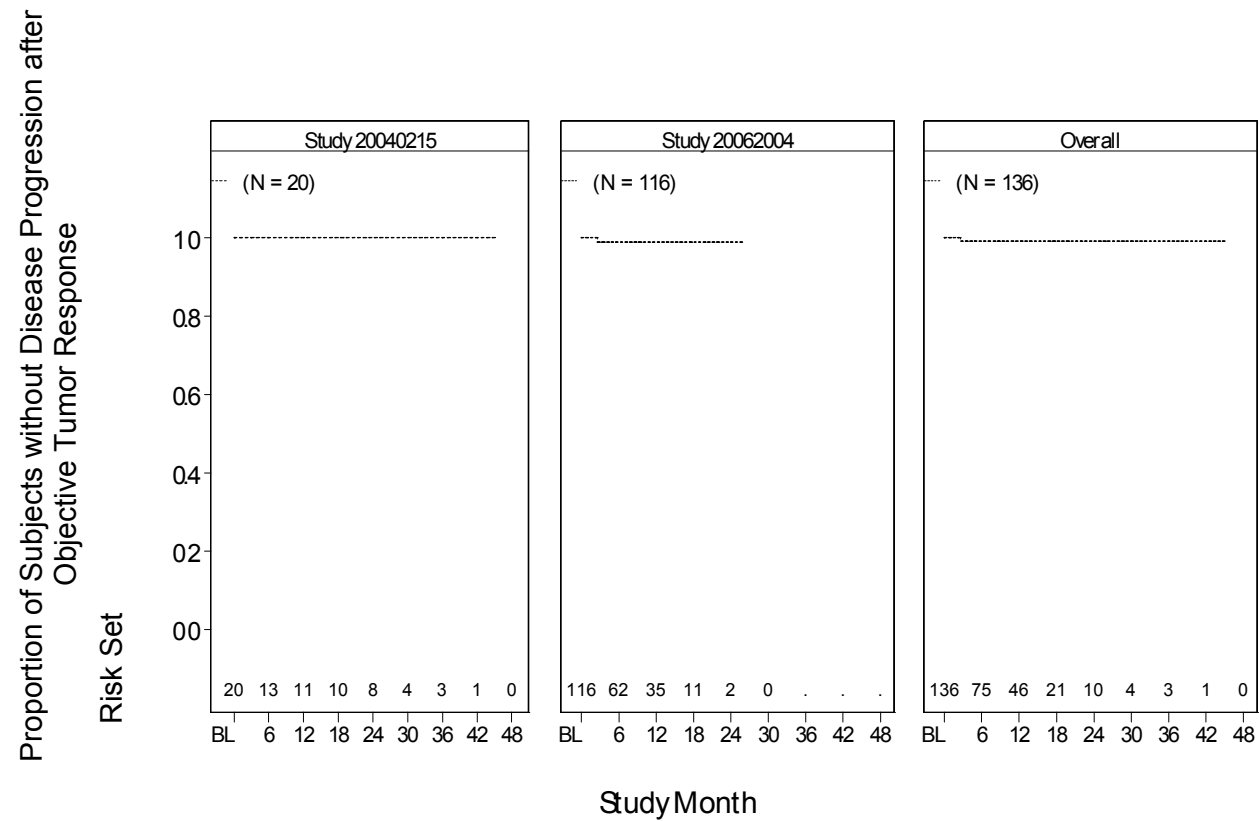
3.2.2 有効性副次評価項目

3.2.2.1 客観的奏効の持続期間

外国 2 試験併合解析では、客観的奏効の後に PD と判定された被験者は非常に少なかった。あらゆる評価基準を用いた最良効果で客観的奏効がみられた 136 名のうち、その後、最良効果の評価に基づき PD と判定された被験者は 1 名であった。個々の評価基準ごとでは、modified RECIST 基準で 47 名中 3 名、modified EORTC 基準で 25 名中 0 名、density/size 基準で 134 名中 1 名において、客観的奏効の後に PD が認められた ([Table tie4.2.1](#))。客観的奏効の後に PD が認められた被験者の数は RECIST 基準と density/size 基準で異なるが、これは RECIST 基準によって PD と判定された 3 名の被験者が、density/size 基準ではすべて PR と判定されたためであった (被験者 002*、003*、004*)。Density/size 基準によって客観的奏効の後に PD と判定された 1 名の被験者は、RECIST 基準ではすべての時点で PD と判定された (被験者 005*)。病勢の進行を認めた被験者は少ないため、Kaplan-Meier 法で客観的奏効の持続期間の中央値を推定することはできなかった。最良効果に基づく客観的奏効の持続期間に関する Kaplan-Meier 曲線を [図 2.7.3.3-5](#) に示す。Modified RECIST 基準、modified EORTC 基準、及び density/size 基準によって評価した個々の試験及び全体の客観的奏効の持続期間に関する Kaplan-Meier 曲線を、それぞれ [Figure tie4.1.1.3](#)、[Figure tie4.1.1.5](#)、及び [Figure tie4.1.1.7](#) に示す。試験 20062004 におけるコホートごとの客観的奏効の持続期間に関する Kaplan-Meier 曲線を [Figure tie4.1.1.2](#) (最良効果)、[Figure tie4.1.1.4](#) (modified RECIST 基準)、[Figure tie4.1.1.6](#) (modified EORTC 基準)、及び [Figure tie4.1.1.8](#) (density/size 基準) に示す。

試験 AMG162-B-J201 では、客観的奏効の後に PD と判定された被験者はいなかったため、Kaplan-Meier 法で客観的奏効の持続期間の中央値を推定することはできなかった。最良効果に基づく客観的奏効の持続期間に関する Kaplan-Meier 曲線を [図 2.7.3.3-6](#) に示す。これらの結果は、上述の外国 2 試験併合解析の結果と同じ傾向を示した。なお、3 試験併合解析の結果を [Table 1tie4-2.1](#)、[Figure 1tie4-3.1](#)、及び [Figure 1tie4-3.2](#) に示した。

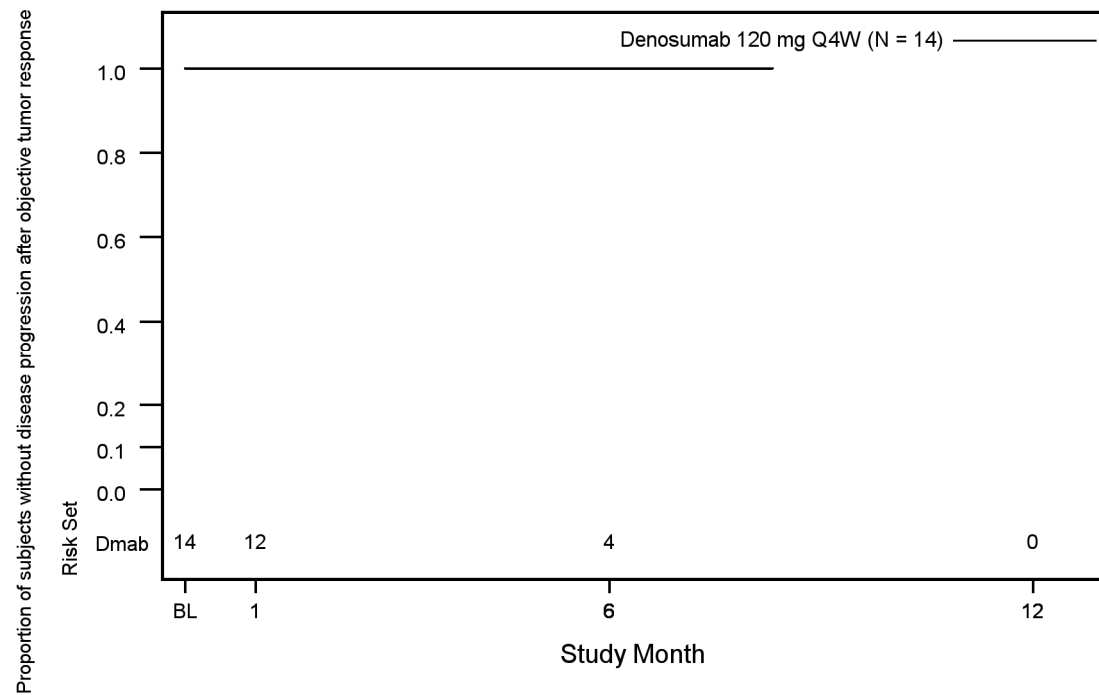
図 2.7.3.3-5 最良効果に基づく客観的奏効の持続期間: 外国 2 試験併合解析
(客観的奏効がみられた被験者)
(Efficacy Analysis Set)



N = number of enrolled subjects who were eligible for the study, received at least 1 dose of denosumab, and had an objective tumor response

Source: Figure 4.1.1.1 ()

図 2.7.3.3-6 最良効果に基づく客観的奏効の持続期間: 試験 AMG162-B-J201
(客観的奏効がみられた被験者)
(Efficacy Analysis Set)
(6-month Cutoff Analysis)



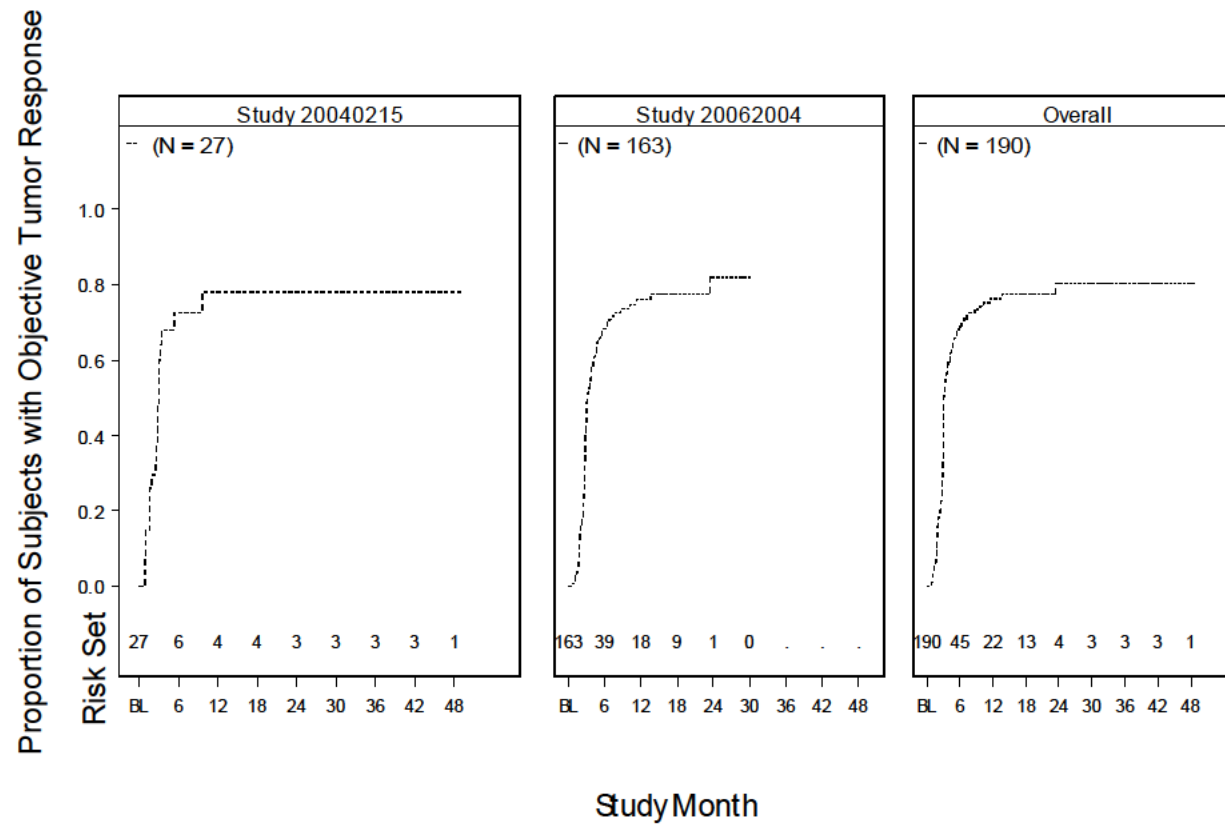
N = Number of subjects who achieved Objective Tumor Response at least once
The horizontal axis of this graph shows the duration after an objective tumor response.
Source: Study AMG162-B-J201 CSR Figure 15.2-1.2

3.2.2.2 客観的奏効までの期間

外国 2 試験併合解析において、あらゆる評価基準を用いた最良効果に基づく客観的奏効が認められた被験者では、その奏効までの期間の中央値（95% CI）は 2.8 カ月（2.76～2.89）であった（ad hoc Table tiae4.4.510）。評価可能な全被験者では、デノスマブの投与開始後、あらゆる評価基準を用いた最良効果に基づく客観的奏効までの期間の中央値（95% CI）は 3.1 カ月（2.89～3.65）であった（Table tiae4.2.2）。評価可能な全被験者における、あらゆる評価基準を用いた最良効果に基づく客観的奏効までの期間に関する Kaplan-Meier 曲線を図 2.7.3.3-7 に示す。Modified RECIST 基準、modified EORTC 基準、及び density/size 基準によって評価した個々の試験及び全体の客観的奏効までの期間に関する Kaplan-Meier 曲線を、それぞれ Figure tiae4.1.2.3、Figure tiae4.1.2.5、及び Figure tiae4.1.2.7 に示す。試験 20062004 におけるコホート別の客観的奏効までの期間に関する Kaplan-Meier 曲線を Figure tiae4.1.2.2（最良効果）、Figure tiae4.1.2.4（modified RECIST 基準）、Figure tiae4.1.2.6（modified EORTC 基準）、及び Figure tiae4.1.2.8（density/size 基準）に示す。

試験 AMG162-B-J201 において、評価可能な全被験者でのあらゆる評価基準を用いた最良効果に基づく客観的奏効までの期間の中央値（95% CI）は 3.0 カ月（2.9～3.1）であった（試験 AMG162-B-J201 治験総括報告書 Table 15.2-1.3.1）。評価可能な全被験者における、あらゆる評価基準を用いた最良効果に基づく客観的奏効までの期間に関する Kaplan-Meier 曲線を図 2.7.3.3-8 に示す。これらの結果は、外国 2 試験併合解析の結果と同様であった。なお、3 試験併合解析の結果を Table 1tiae4-3.1、Figure 1tiae4-2.1、及び Figure 1tiae4-2.2 に示した。

図 2.7.3.3-7 最良効果に基づく客観的奏効までの期間に関する Kaplan-Meier 曲線: 外国 2 試験併合解析
 (1 時点以上で評価可能な画像データが得られた被験者)
 (Efficacy Analysis Set)



N = number of subjects who were eligible for the study, received at least 1 dose of denosumab, and had at least 1 evaluable time point assessment.

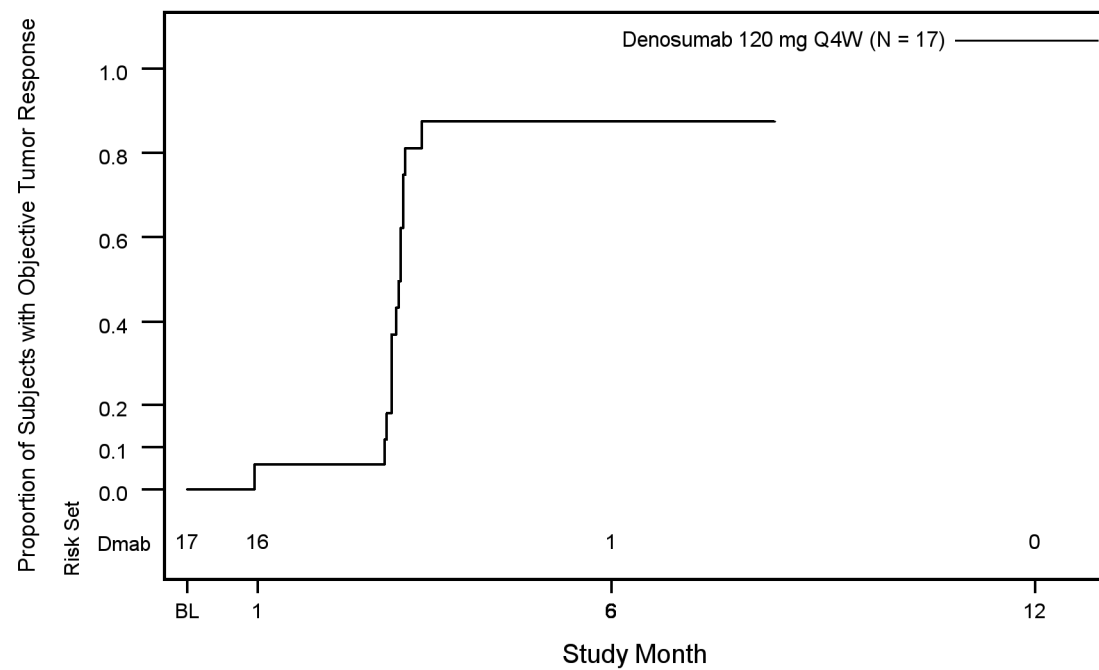
Source: [Figure tiae4.1.2.1](#) ()

図 2.7.3.3-8 最良効果に基づく客観的奏効までの期間に関する Kaplan-Meier 曲線: 試験 AMG162-B-J201

(1 時点以上で評価可能な画像データが得られた被験者)

(Efficacy Analysis Set)

(6-month Cutoff Analysis)



N = Number of enrolled subjects who were eligible for the study, received at least one dose of denosumab, and had at least one evaluable time point assessment

Source: Study AMG162-B-J201 CSR Figure 15.2-1.1

3.2.2.3 客観的奏効の持続性

投与中の 4、8、12、又は 24 週間以上離れた 2 時点以上で評価可能な画像データが得られた被験者において、効果の持続性を評価した。

外国 2 試験併合解析では、これらの期間、あらゆる評価基準を用いた最良効果に基づく客観的奏効が持続した被験者の割合は、66.7%~68.5%であった（表 2.7.3.3-8）。個々の評価基準ごとの効果が持続した被験者の割合も、同評価期間を通して同程度であった。最良効果に基づき、4、8、12、及び 24 週間にわたって効果が持続した被験者の割合は、個々の試験、並びに試験 20062004 のコホート 1 及び 2 を通じ、概ね類似していた（Table tie4.2.5、Table tie4.2.6、Table tie4.2.3、及び Table tie4.2.4）。

この客観的奏効の持続性の解析は、それぞれの評価期間で評価可能な画像データが得られた被験者を対象とした。外国 2 試験併合解析における客観的抗腫瘍効果の解析対象集団のうち評価可能な画像データが得られなかった被験者を打ち切りとした Kaplan-Meier 曲線を、ad hoc Figure tie4.4.502（4 週間）、Figure tie4.4.503（8 週間）、Figure tie4.4.501（12 週間）、及び Figure tie4.4.504（24 週間）に示した（あらゆる評価基準を用いた最良効果の Kaplan-Meier 曲線と個々の評価基準ごとの Kaplan-Meier 曲線を含む）。

試験 AMG162-B-J201 では、あらゆる評価基準を用いた最良効果に基づく客観的奏効が持続した被験者の割合は、83.3%~100.0%であった（表 2.7.3.3-9）。個々の評価基準ごとの効果が持続した被験者の割合も、同評価期間を通して同程度であった。これらの結果は、外国 2 試験併合解析の結果と同じ傾向を示した。なお、3 試験併合解析の結果を Table tie4-4.1、Table tie4-4.2、Table tie4-4.3、及び Table tie4-4.4 に示した。

表 2.7.3.3-8 客観的奏効が持続した被験者の割合: 外国 2 試験併合解析

(Efficacy Analysis Set)

	n	N1	Percent	95% CI ^a
Sustained for at least 4 weeks				
Based on best response	102	153	66.7	(58.6, 74.1)
RECIST 1.1	32	150	21.3	(15.1, 28.8)
EORTC	18	20	90.0	(68.3, 98.8)
Density/size	101	143	70.6	(62.4, 77.9)
Sustained for at least 8 weeks				
Based on best response	99	147	67.3	(59.1, 74.8)
RECIST 1.1	32	144	22.2	(15.7, 29.9)
EORTC	17	18	94.4	(72.7, 99.9)
Density/size	99	138	71.7	(63.5, 79.1)
Sustained for at least 12 weeks				
Based on best response	98	144	68.1	(59.8, 75.6)
RECIST 1.1	32	141	22.7	(16.1, 30.5)
EORTC	16	17	94.1	(71.3, 99.9)
Density/size	97	135	71.9	(63.5, 79.2)
Sustained for at least 24 weeks				
Based on best response	76	111	68.5	(59.0, 77.0)
RECIST 1.1	26	109	23.9	(16.2, 33.0)
EORTC	11	12	91.7	(61.5, 99.8)
Density/size	76	102	74.5	(64.9, 82.6)

Page 1 of 1

n = number of subjects with a response

N1 = number of subjects with at least 2 evaluable time point assessments that were at least 4, 8, 12, or 24 weeks apart using the respective tumor response criteria.

^a Exact confidence interval.

Source: Tables tia4.2.5, tia4.2.6, tia4.2.3, tia4.2.4; results by individual study are provided in source tables.

表 2.7.3.3-9 客観的奏効が持続した被験者の割合: 試験 AMG162-B-J201

(Efficacy Analysis Set)

(6-month Cutoff Analysis)

	n	N1	Percent	95% CI ^a
Sustained for at least 4 weeks				
Based on best response	12	14	85.7	(57.2, 98.2)
RECIST 1.1	1	14	7.1	(0.2, 33.9)
EORTC	8	14	57.1	(28.9, 82.3)
Density/size	9	14	64.3	(35.1, 87.2)
Sustained for at least 8 weeks				
Based on best response	12	14	85.7	(57.2, 98.2)
RECIST 1.1	1	14	7.1	(0.2, 33.9)
EORTC	8	14	57.1	(28.9, 82.3)
Density/size	9	14	64.3	(35.1, 87.2)
Sustained for at least 12 weeks				
Based on best response	10	12	83.3	(51.6, 97.9)
RECIST 1.1	1	12	8.3	(0.2, 38.5)
EORTC	5	11	45.5	(16.7, 76.6)
Density/size	8	12	66.7	(34.9, 90.1)
Sustained for at least 24 weeks				
Based on best response	7	7	100.0	(59.0, 100.0)
RECIST 1.1	0	7	0.0	(0.0, 41.0)
EORTC	4	7	57.1	(18.4, 90.1)
Density/size	6	7	85.7	(42.1, 99.6)

Page 1 of 1

n = number of subjects with a response

N1 = number of subjects with at least 2 evaluable time point assessments that were at least 4, 8, 12, or 24 weeks apart using the respective tumor response criteria.

^a Exact confidence interval.

Source: Study AMG162-B-J201 CSR Table 15.2-1.4.1

3.2.2.4 腫瘍コントロールの持続性

外国 2 試験併合解析では、あらゆる評価基準を用いた最良効果に基づき腫瘍コントロール (CR、PR、又は SD) が 12 週以上持続した被験者の割合は 96.5% (95% CI: 92.1~98.9) であった (表 2.7.3.3-10)。個々の評価基準ごとの評価においても、12 週以上の腫瘍コントロールが高い割合で確認された (97%以上)。腫瘍コントロールが 4、8、又は 24 週以上持続した被験者の割合は 12 週以上の場合と同様であり、96.7%~98.2%であった (表 2.7.3.3-10)。個々の試験においても同様の結果が認められた。切除可能な GCTB 患者と切除不能な GCTB 患者 (試験 20062004 のコホート 1 と 2) の間でも、腫瘍コントロールの持続性は同様であった (Table tiae4.9.1、Table tiae4.9.2、Table tiae4.9.3、及び Table tiae4.9.4)。

試験 AMG162-B-J201 では、あらゆる評価基準を用いた最良効果に基づき腫瘍コントロール (CR、PR、又は SD) が 12 週以上持続した被験者の割合は 100.0% (12/12) であった (表 2.7.3.3-11)。個々の評価基準ごとの評価においても、12 週以上の腫瘍コントロールが全被験者で確認され

た。6ヵ月カットオフ解析日時点で、いずれかの評価基準又はいずれかの評価時点でも PD と判定された被験者はいなかった。腫瘍コントロールが 4、8、又は 24 週以上持続した被験者の割合は、12 週以上の場合と同様であり、100.0%であった（表 2.7.3.3-11）。これらの結果は、外国 2 試験併合解析の結果と類似していた。なお、3 試験併合解析の結果は Table 1tia4-9.1、Table 1tia4-9.2、Table 1tia4-9.3、及び Table 1tia4-9.4 に示した。

表 2.7.3.3-10 腫瘍コントロールが持続した被験者の割合：外国 2 試験併合解析
(Efficacy Analysis Set)

	n	N1	Percent	95% CI ^a
Proportion of subjects with tumor control (CR, PR, or SD) sustained for at least 4 weeks				
Based on best response	148	153	96.7	(92.5, 98.9)
RECIST 1.1	145	150	96.7	(92.4, 98.9)
EORTC	19	20	95.0	(75.1, 99.9)
Density/size	139	143	97.2	(93.0, 99.2)
Proportion of subjects with tumor control (CR, PR, or SD) sustained for at least 8 weeks				
Based on best response	143	147	97.3	(93.2, 99.3)
RECIST 1.1	140	144	97.2	(93.0, 99.2)
EORTC	18	18	100.0	(81.5, 100.0)
Density/size	135	138	97.8	(93.8, 99.5)
Proportion of subjects with tumor control (CR, PR, or SD) sustained for at least 12 weeks				
Based on best response	139	144	96.5	(92.1, 98.9)
RECIST 1.1	137	141	97.2	(92.9, 99.2)
EORTC	17	17	100.0	(80.5, 100.0)
Density/size	131	135	97.0	(92.6, 99.2)
Proportion of subjects with tumor control (CR, PR, or SD) sustained for at least 24 weeks				
Based on best response	109	111	98.2	(93.6, 99.8)
RECIST 1.1	108	109	99.1	(95.0, 100.0)
EORTC	12	12	100.0	(73.5, 100.0)
Density/size	101	102	99.0	(94.7, 100.0)

Page 1 of 1

n = number of subjects with a response

N1 = number of subjects with at least 2 evaluable time point assessments that were at least 4,8,12, or 24 weeks apart using the respective tumor response criteria.

^a Exact confidence interval.

Source: Tables tia4.9.1, tia4.9.2, tia4.9.3, tia4.9.4; results by individual study are provided in source tables.

表 2.7.3.3-11 腫瘍コントロールが持続した被験者の割合: 試験 AMG162-B-J201
(Efficacy Analysis Set)
(6-month Cutoff Analysis)

	n	N1	Percent	95% CI ^a
Proportion of subjects with tumor control (CR, PR, or SD) sustained for at least 4 weeks				
Based on best response	14	14	100.0	(76.8, 100.0)
RECIST 1.1	14	14	100.0	(76.8, 100.0)
EORTC	14	14	100.0	(76.8, 100.0)
Density/size	14	14	100.0	(76.8, 100.0)
Proportion of subjects with tumor control (CR, PR, or SD) sustained for at least 8 weeks				
Based on best response	14	14	100.0	(76.8, 100.0)
RECIST 1.1	14	14	100.0	(76.8, 100.0)
EORTC	14	14	100.0	(76.8, 100.0)
Density/size	14	14	100.0	(76.8, 100.0)
Proportion of subjects with tumor control (CR, PR, or SD) sustained for at least 12 weeks				
Based on best response	12	12	100.0	(73.5, 100.0)
RECIST 1.1	12	12	100.0	(73.5, 100.0)
EORTC	11	11	100.0	(71.5, 100.0)
Density/size	12	12	100.0	(73.5, 100.0)
Proportion of subjects with tumor control (CR, PR, or SD) sustained for at least 24 weeks				
Based on best response	7	7	100.0	(59.0, 100.0)
RECIST 1.1	7	7	100.0	(59.0, 100.0)
EORTC	7	7	100.0	(59.0, 100.0)
Density/size	7	7	100.0	(59.0, 100.0)

Page 1 of 1

n = number of subjects with a response

N1 = number of subjects with at least 2 evaluable time point assessments that were at least 4,8,12, or 24 weeks apart using the respective tumor response criteria.

^a Exact confidence interval.

Source: Study AMG162-B-J201 CSR Table 15.2-1.6.1

3.2.2.5 客観的奏効の経時推移

外国 2 試験併合解析では、あらゆる評価基準による最良効果に基づく客観的奏効までの期間の中央値が 3.1 ヶ月であり（第 3.2.2.2 項）、評価可能な被験者の多くは（65%超）、第 25 週までに、あらゆる評価基準による最良効果に基づく客観的奏効に達した。最良効果に基づく客観的奏効率の経時推移について Kaplan-Meier 法で推定すると、第 25 週で 68.3%、第 49 週で 75.2%、第 73 週で 77.5%であった（表 2.7.3.3-12）。個々の評価基準ごとの評価でも、客観的奏効率の推定値は同様の推移を示した。

補足的な事後解析において、上記 3 時点における客観的奏効率の推定値を算出したところ、いずれの時点でも Kaplan-Meier 法による推定値と同様であった（Table tie4.4.507）。

試験 AMG162-B-J201 では、あらゆる評価基準による最良効果に基づく客観的奏効率の Kaplan-Meier 推定値は、第 25 週で 87.4%であり、評価可能な被験者の多くが第 25 週までに客観的奏効に達した（表 2.7.3.3-13）。個々の評価基準ごとの評価でも、客観的奏効率の推定値は

同様の推移を示した。これらの結果は、外国 2 試験併合解析の結果と同様の傾向を示した。なお、3 試験併合解析の結果は [Table 1tie4-5.1](#) に示した。

表 2.7.3.3-12 客観的奏効率の経時推移（Kaplan-Meier 推定値）：外国 2 試験併合解析
（1 時点以上で評価可能な画像データが得られた被験者）
（Efficacy Analysis Set）

	Crude Incidence n/N1 (%)	Week 25 % (95% CI)	Week 49 % (95% CI)	Week 73 % (95% CI)
Proportion of subjects with objective tumor response over time				
Based on best response	136/190 (71.6)	68.3 (61.3, 75.2)	75.2 (68.5, 82.0)	77.5 (70.7, 84.3)
RECIST 1.1	47/187 (25.1)	19.7 (13.6, 25.7)	26.9 (19.6, 34.2)	32.8 (24.0, 41.6)
EORTC	25/26 (96.2)	91.2 (79.7, 100.0)	100.0 (91.6, 100.0)	100.0 (91.6, 100.0)
Density/size	134/176 (76.1)	73.0 (66.2, 79.9)	79.6 (73.0, 86.2)	82.2 (75.5, 88.8)

Page 1 of 1

n = number of subjects with a response

N1 = number of subjects with at least one evaluable time point assessment using the respective tumor response criteria.

Modified from Source: Table tie4.2.7; results by individual study are provided in source table.

表 2.7.3.3-13 客観的奏効率の経時推移（Kaplan-Meier 推定値）：試験 AMG162-B-J201
（Efficacy Analysis Set）
（6-month Cutoff Analysis）

	Crude Incidence n/N1 (%)	Week 25 % (95% CI)	Week 49 % (95% CI)	Week 73 % (95% CI)
Proportion of subjects with objective tumor response over time				
Based on best response	14/17 (82.4)	87.4 (71.1, 100.0)	NE (NE, NE)	NE (NE, NE)
RECIST 1.1	5/17 (29.4)	13.1 (0.0, 30.2)	33.6 (4.3, 63.0)	NE (NE, NE)
EORTC	12/17 (70.6)	68.5 (45.7, 91.3)	84.2 (59.6, 100.0)	NE (NE, NE)
Density/size	11/17 (64.7)	56.8 (32.1, 81.5)	74.1 (50.3, 97.8)	NE (NE, NE)

Page 1 of 1

n = number of subjects with a response

N1 = number of subjects with at least one evaluable time point assessment using the respective tumor response criteria.

NE = Not estimable

Source: Study AMG162-B-J201 CSR Table 15.2-1.8.1

3.2.3 有効性探索的評価項目

3.2.3.1 あらゆる評価基準を用いた最良効果別の被験者の割合

外国 2 試験併合解析では、[第 3.2.1 項](#)に要約したように、あらゆる評価基準を用いた最良効果による CR 又は PR と定義した客観的奏効は、評価可能な被験者の 71.6% で認められた。CR、PR、又は SD として定義した腫瘍コントロールが得られた被験者は合計 186 名（97.9%）であり、CR が 4.2%（8 名）、PR が 67.4%（128 名）、SD が 26.3%（50 名）であった（[Table tie4.3.1](#)）。PD は 2.1%（4 名）であった。Modified RECIST 基準では、PR が 25.1%（47 名）、SD が 71.7%

(134 名)であり、CR は認められなかった (Table 4.3.1.2)。Density/size 基準では、PR が 76.1% (134 名)、SD が 22.2% (30 名)であり、CR は認められなかった。Modified EORTC 基準では、CR が 30.8% (8 名)、PR が 65.4% (17 名)であり、SD は認められなかった。

試験 AMG162-B-J201 では、あらゆる評価基準を用いた最良効果に基づく客観的奏効が 17 名中 14 名 (82.4%) で認められた。腫瘍コントロールが得られた被験者は 17 名 (100.0%) であり、CR が 17.6% (3/17)、PR が 64.7% (11/17)、SD が 17.6% (3/17) で、PD は認められなかった (試験 AMG162-B-J201 治験総括報告書 Table 15.2-1.9.1)。Modified RECIST 基準では、PR が 5 名、SD が 12 名であり、CR は認められなかった (試験 AMG162-B-J201 治験総括報告書 Figure 15.2-3)。Density/size 基準では、PR が 11 名、SD が 6 名であり、CR は認められなかった (試験 AMG162-B-J201 治験総括報告書 Figure 15.2-3)。Modified EORTC 基準では、CR が 3 名、PR が 9 名であり、SD が 5 名であった (試験 AMG162-B-J201 治験総括報告書 Figure 15.2-3)。これらの結果は、外国 2 試験併合解析の結果と同様の傾向を示した。なお、3 試験併合解析の結果は Table 4.6.1 に示した。

3.2.3.2 病勢の進行までの期間

外国 2 試験併合解析では、190 名の評価可能な被験者のうち、3 つの評価基準のいずれかを用いて、試験中のいずれかの時点で PD と判定された被験者は 11 名 (5.8%) であり、病勢の進行までの期間は中央値に達しなかった (Table 4.3.2)。これら 11 名におけるデノスマブの投与開始から病勢の進行までの期間は、0.2~21.9 カ月 (中央値: 3.68 カ月) であった (Table 4.3.3)。

試験 AMG162-B-J201 では、17 名の評価可能な被験者のうち、いずれかの評価基準又はいずれかの時点で PD と判定された被験者はいなかった。なお、3 試験併合解析の結果を Table 4.7.1 に示した。

3.2.4 探索的評価項目: 客観的奏効を認めた被験者と認めなかった被験者における痛みの転帰、鎮痛薬の使用、及び臨床的有益性

3.2.4.1 客観的奏効を認めた被験者と認めなかった被験者における BPI-SF の痛みスコア

試験 20062004 で客観的奏効を認めた被験者と認めなかった被験者で、痛みの転帰を評価した。コホート 1 と 2 を併せた結果を以下に要約する。個々のコホートごとの結果は、モジュール 5.3.5.3 (有効性の併合解析) に示す。あらゆる評価基準を用いた最良効果に基づく客観的奏効の有無で痛みの評価を行った。試験 AMG162-B-J201 では被験者ごとに評価した。

3.2.4.1.1 BPI-SF の最もひどい痛み

試験 20062004 では、客観的奏効を認めた被験者と認めなかった被験者のいずれにおいても、ベースライン時の最もひどい痛みのスコアの平均値及び中央値は、試験中のいずれの評価時点よりも高かった。唯一例外として、客観的奏効を認めた被験者における第 8 日のスコアの中央値 (コホート 1 と 2 を併せた結果) が、ベースライン時と同じであった。最もひどい痛みのス

コアの来院日ごとの平均値を [Table tia4.4.1](#)、痛みの重症度のスコアの来院日ごとのベースラインからの変化の平均値を [Table tia4.4.2](#) に、また、痛みによる機能障害のスコアの来院日ごとの平均値を [Table tia4.4.3](#) に要約する。客観的奏効を認めた被験者と認めなかった被験者における最もひどい痛み、痛みの重症度、及び痛みによる機能障害のスコアの来院日ごとのベースラインからの変化の平均値を、それぞれ [Table tia4.5.1](#)、[Table tia4.5.2](#)、及び [Table tia4.5.3](#) に示す。

試験 AMG162-B-J201 では、客観的奏効を認めた被験者 14 名の最もひどい痛みのスコアの推移は、ベースラインと比べ低下傾向又は大きく変化しなかった。客観的奏効を認めなかった被験者 3 名はいずれもベースラインのスコアが 0 又は 1 であり、試験期間中に悪化は認められなかった。最もひどい痛み、痛みの重症度、及び痛みによる機能障害のスコアの来院日ごとのベースラインからの変化の平均値を、それぞれ試験 AMG162-B-J201 治験総括報告書 [Table 15.2-2.17.1](#)、[Table 15.2-2.18.1](#)、及び [Table 15.2-2.19.1](#) に示す。

3.2.4.1.2 痛みの改善

試験 20062004 では、客観的奏効を認めた被験者と認めなかった被験者のいずれにおいても、痛みの改善がみられた。多くの時点で、臨床的に意味のある最もひどい痛みの減少（最もひどい痛みのスコアがベースラインから 2 ポイント以上減少）が認められた被験者の割合は、客観的奏効を認めた被験者と認めなかった被験者の間で同程度であった（[Table tia4.6.1](#)）。いずれの部分集団でも半数以上の被験者で、第 5 週以降の全評価時点において、臨床的に意味のある最もひどい痛みの減少がみられた。また、いずれの部分集団でも、臨床的に意味のある最もひどい痛みが改善されるまでの期間はかなり短く、中央値として 1 ヶ月以内に改善がみられ、客観的奏効を認めなかった被験者では 0.5 ヶ月（85.3% [29/34]）、客観的奏効を認めた被験者では 0.7 ヶ月（84.6% [66/78]）であった（[Table tia4.8.1](#) 及び [Figure taie4.1.4](#)）。

また、試験 20062004 では、中等度又は重度の最もひどい痛み（4 ポイント超）から軽度の痛み（4 ポイント以下）に変化した被験者の割合は、客観的奏効を認めなかった被験者では第 1 週までに 50%を超え（52.9% [9/17]）、客観的奏効を認めた被験者では第 9 週までに 50%を超えた（56.8% [25/44]）（[Table tia4.6.2](#)）。ベースライン時の最もひどい痛みのスコアが 4 ポイントを超えた被験者において、最もひどい痛みのスコアが 4 ポイント以下になるまでの期間は、客観的奏効を認めなかった被験者と認めた被験者で、それぞれ 0.5 ヶ月（90.5% [19/21]）及び 1.0 ヶ月（81.6% [40/49]）であった（[Table tia4.8.2](#) 及び [Figure taie4.1.5](#)）。

試験 AMG162-B-J201 では、臨床的に意味のある最もひどい痛みの減少は、客観的奏効を認めた被験者 8 名（ベースライン時に 2 ポイント以上であった被験者）のうち 5 名に認められた（試験 AMG162-B-J201 治験総括報告書 [Figure 15.2-3](#)）。ベースライン時の最もひどい痛みのスコアが 2 ポイント以上の被験者のうち、客観的奏効を認めた被験者では、半数以上の被験者で、第 5 週以降のほとんどの評価時点において、臨床的に意味のある最もひどい痛みの減少がみられた（試験 AMG162-B-J201 治験総括報告書 [Figure 15.2-3](#)）。なお、客観的奏効を認めなかった被験者はいずれもベースライン時の最もひどい痛みのスコアが 2 ポイント以下であり（0 ポイ

ント2名、1ポイント1名)、臨床的に意味のある最もひどい痛みの減少の評価の対象とはならなかった。また、客観的奏効を認めた被験者では、臨床的に意味のある最もひどい痛みが改善されるまでの期間の中央値は71.0日(62.5% [5/8])であった(試験 AMG162-B-J201 治験総括報告書 Table 15.2-3.10.1)。中等度又は重度の最もひどい痛みから軽度の痛みに変化した被験者は、客観的奏効を認めた被験者では、第1週までに50%に達した(50.0% [2/4]) (試験 AMG162-B-J201 治験総括報告書 Table 15.2-3.7.1)。ベースライン時の最もひどい痛みのスコアが4ポイントを超えた被験者において、最もひどい痛みのスコアが4ポイント以下になるまでの期間は、客観的奏効を認めた被験者では、11.5日(75.0%[3/4])であった(試験 AMG162-B-J201 治験総括報告書 Table 15.2-3.14.1)。

3.2.4.1.3 痛みの悪化

試験 20062004 では、客観的奏効を認めた被験者と認めなかった被験者のいずれにおいても、臨床的に意味のある痛みの悪化(最もひどい痛みのスコアのベースラインから2ポイント以上の増加)がみられた被験者は少なかった。評価可能な被験者が10名を超えていた時点における、最もひどい痛みのスコアが2ポイント以上増加した被験者の割合は、客観的奏効を認めなかった被験者で18.2%以下(第73週、2/11名)、客観的奏効を認めた被験者で16.1%以下(第73週、5/31名)であった(Table tie4.6.4)。同様に、ベースライン時の最もひどい痛みのスコアが4ポイント以下であり、試験中のいずれかの時点(評価可能な被験者が10名を超えていた時点)で最もひどい痛みのスコアが4ポイントを上回った被験者も少なく、客観的奏効を認めた被験者と認めなかった被験者で、それぞれ13%以下及び16.1%未満であった(Table tie4.6.5)。

また、試験 20062004 では、客観的奏効を認めなかった被験者における臨床的に意味のある痛みの悪化までの期間の中央値は23.2ヵ月(37.8% [17/45])であり、客観的奏効を認めた被験者では24.1ヵ月(34.0% [33/97])であった(Table tie4.8.4 及び Figure tie4.1.7)。客観的奏効を認めた被験者と認めなかった被験者のいずれにおいても、中等度又は重度の最もひどい痛みまでの期間の中央値は推定できなかった(Table tie4.8.5 及び Figure tie4.1.8)。

試験 AMG162-B-J201 では、臨床的に意味のある痛みの悪化が認められた被験者は、客観的奏効を認めた被験者14名のうち3名、客観的奏効を認めなかった被験者3名のうち0名であった(試験 AMG162-B-J201 治験総括報告書 Figure 15.2-3)。ベースライン時の最もひどい痛みのスコアが4ポイント以下であり、試験中のいずれかの時点で最もひどい痛みのスコアが4ポイントを上回った被験者は、客観的奏効を認めた被験者10名のうち2名、客観的奏効を認めなかった被験者3名のうち0名であった(試験 AMG162-B-J201 治験総括報告書 Figure 15.2-3)。臨床的に意味のある痛みの悪化までの期間の中央値及び中等度又は重度の最もひどい痛みまでの期間の中央値は推定できなかった(試験 AMG162-B-J201 治験総括報告書 Table 15.2-3.12.1 及び Table 15.2-3.15.1)。

3.2.4.2 客観的奏効を認めた被験者と認めなかった被験者における鎮痛薬使用の変化

試験 20062004 で、10 名を超える被験者からデータが得られた来院日のいずれかにおいて、ベースライン時の鎮痛薬の使用が強力なオピオイド（スコア: 3 ポイント以上）からなし又は少量（スコア: 2 ポイント以下）に変化した被験者における客観的奏効率は、2.9%（第 1 週、1/34 名）～38.9%（第 61 週、7/18 名）であった（Table 4.6.7）。客観的奏効を認めなかった被験者では、データが得られた被験者が少なかったため（11 名以下）、鎮痛薬の使用の減少を評価することができなかった。ベースライン時の鎮痛薬の使用がなし又は少量（スコア: 2 ポイント以下）から、試験中に強力なオピオイドに変化した被験者は少なく、客観的奏効を認めた被験者と認めなかった被験者で、それぞれ 8.6%（第 9 週、3/35 名）以下及び 4.2%（第 25 週、3/71 名）以下であった（Table 4.6.8）。

試験 AMG162-B-J201 では、被験者の来院日のいずれかにおいて、ベースライン時の鎮痛薬の使用が強力なオピオイドからなし又は少量（スコア: 2 ポイント以下）に変化した 1 名（試験 AMG162-B-J201 治験総括報告書 Table 15.2-5.1）は、客観的奏効を認めた被験者であった。鎮痛薬の使用がベースライン時のなし又は少量（スコア: 2 ポイント以下）から、試験中に強力なオピオイドに変化した被験者はいなかった（試験 AMG162-B-J201 治験総括報告書 Table 15.2-5.2）。

3.2.4.3 客観的奏効を認めた被験者と認めなかった被験者における臨床的有益性

試験 20040215 と 20062004 を併合し、客観的抗腫瘍効果の解析対象集団で、治験責任医師が判断した臨床的有益性を評価した。この外国 2 試験併合解析において、治験責任医師が判断した臨床的有益性が得られた被験者の割合は、客観的奏効を認めなかった被験者と認めた被験者で同程度であった（それぞれ 59.3% [32/54] 及び 59.6% [81/136]）（Table 4.7.1）。

試験 AMG162-B-J201 では被験者ごとに、客観的抗腫瘍効果、最もひどい痛みのスコアの推移、及び治験責任医師による評価（疾患の状態、腫瘍効果、臨床的有益性）をまとめた（試験 AMG162-B-J201 治験総括報告書 Figure 15.2-3）。治験責任医師が判断した臨床的有益性は、客観的奏効を認めた被験者 14 名のうち 12 名、客観的奏効を認められなかった被験者 3 名のうち 2 名に認められた（試験 AMG162-B-J201 治験総括報告書 第 11.4.1.1.15 項）。

客観的奏効を認めた被験者と認めなかった被験者における臨床的有益性に関する 3 試験併合解析の結果を Table 4.14.1 に示す。

3.2.4.4 痛みの転帰、鎮痛薬の使用、及び臨床的有益性に関する探索的解析の要約

痛みの改善及び臨床的有益性と、客観的奏効との間に相関性は認められなかった。客観的奏効を認めなかった被験者では、評価可能なデータが得られた被験者数が少なかったことから、相関性の評価には限界があった。

3.2.5 画像コントロール群における客観的抗腫瘍効果の解析結果

外国 2 試験併合解析では、デノスマブ投与開始前の 3 時点以上において中央読影機関で評価

可能な画像データが得られた被験者集団についても客観的抗腫瘍効果を評価した。画像評価におけるバイアスを避けるため、中央読影機関では被験者が特定されないよう盲検下で評価した。投与開始前の3時点以上の画像データが得られた合計26名の被験者を画像コントロール群とした。これらの投与開始前の画像は、試験期間中の客観的抗腫瘍効果の評価に用いた基準と同じ評価基準を用い、盲検下で評価した。画像コントロール群において評価された投与開始前の最初の画像から最後の画像までの期間の中央値は259.5日であり、被験者が臨床試験に組み入れられる65～782日前のものであった（事後解析: [Table tiae4.1.508](#)）。投与開始前の画像の撮影日、投与開始前の画像間の期間、及び中央読影機関による判定結果を [Listing tiae4.1.501.1](#) に示す。画像コントロール群における、あらゆる評価基準を用いた最良効果に基づく客観的奏効率は34.6%（9/26名、95% CI [17.2～55.7]）であった（[Table tiae4.2.8](#)）。個々の評価基準ごとの客観的奏効率は、modified RECIST 基準で8.0%（2/25名、95% CI [1.0～26.0]）、density/size 基準で37.5%（9/24名、95% CI [18.8～59.4]）であった。画像コントロール群に、modified EORTC 基準で評価するための¹⁸FDG-PET 画像又は¹⁸FDG-PET/CT 画像を有する被験者はいなかった。

画像コントロール群で客観的奏効を認めた9名全員では、density/size 基準を用いた最良効果によるPRが認められ、2名ではmodified RECIST 基準によるPRも認められた。9名のうち3名では、投与開始前の1時点の評価で客観的奏効が認められ、その後の投与開始前時点の評価ではSDと判定された。9名のうち7名では、デノスマブ投与開始前にGCTBに対する内科的又は外科的処置や治療が実施されていた。今回の独立した画像評価では、これらの臨床的介入の効果が評価されたものと考えられる。画像コントロール群全体からこれら7名を除外した場合、あらゆる評価基準を用いた最良効果に基づく客観的奏効率は10.5%（2/19）である。この客観的奏効を認めた2名（被験者 006* 及び 007* ）では、投与開始前の1時点でdensity/size 基準によるPRが認められたが、その後の投与開始前時点ではSDであった。[表 2.7.3.3-14](#) に、客観的奏効が認められた画像コントロール群の被験者の情報を要約した。

試験 AMG162-B-J201 では画像コントロール群を設定しなかった。

表 2.7.3.3-14 画像コントロール群において抗腫瘍効果がみられた理由

Subject ID Age Gender	Date of Pretreatment Images	Response ^a / Method	Brief Description of Therapies Received Prior to Denosumab Treatment
008* 3 ■	Image 1: ■ 20 Image 2: ■ 20 Image 3: ■ 20	PR RECIST and density/size	Subject had embolization on ■ 20 and anterior tumor resection on ■ 20.
009* 2 ■	Image 1: ■ 20 Image 2: ■ 20 Image 3: ■ 20	PR density/size	Subject had sacral laminectomy of tumor on ■ 20.
010* 1 ■	Image 1: ■ 20 Image 2: ■ 20 Image 3: ■ 20 Image 4: ■ 20	PR RECIST and density/size	Subject had tumor resection with intraoperative microscopic and dural repair on ■ 20. Subject had another tumor resection on ■ 20 followed by radiation therapy. Subject also received anti-neoplastic chemotherapy prior to study enrollment.
011* 5 ■	Image 1: ■ 20 Image 2: ■ 20 Image 3: ■ 20	PR density/size	Subject had multiple local excisions from ■ 20 to ■ 20.
006* 2 ■	Image 1: ■ 20 Image 2: ■ 20 Image 3: ■ 20	PR density/size	Subject had PR by density/size for pretreatment image 2 and SD for evaluation of pretreatment image 3 by these criteria.
012* 4 ■	Image 1: ■ 20 Image 2: ■ 20 Image 3: ■ 20	PR density/size	Subject received bisphosphonates. Subject also had GCTB surgery and anti-neoplastic chemotherapy prior to study enrollment.
007* 5 ■	Image 1: ■ 20 Image 2: ■ 20 Image 3: ■ 20	PR density/size	Subject had PR by density/size for pretreatment image 2 and SD for evaluation of pretreatment image 3 by these criteria.
013* 3 ■	Image 1: ■ 20 Image 2: ■ 20 Image 3: ■ 20 Image 4: ■ 20 Image 5: ■ 20 Image 6: ■ 20 Image 7: ■ 20	PR density/size	Subject had curettage and cementing on ■ and ■ 20. Subject received bisphosphonates.
014* 5 ■	Image 1: ■ 20 Image 2: ■ 20 Image 3: ■ 20	PR density/size	Subject received bisphosphonates. Subject also had GCTB surgery prior to study enrollment.

Page 1 of 1

PR = partial response; SD = stable disease

^a Best response based on any tumor response criteria.

Source: Listing tiae4-1.501.1, Listing4-1.503

3.3 外科的に切除可能な GCTB 患者に対するデノスマブの有効性

試験 20062004 及び AMG162-B-J201 での外科的に切除可能な GCTB 患者に対するデノスマブの有効性の結果を示す。なお、術後の再発率を低下できるかどうかについては、試験 20062004 及び AMG162-B-J201 とともに、完全切除が確認された後、治験薬を 6 回投与することとしているが、現時点では十分な検証はできていない。

3.3.1 被験者の背景

試験 20062004 では、コホート 2 に登録された被験者は外科的に切除可能な GCTB を有していた。試験 AMG162-B-J201 では、4 名の外科的に切除可能な GCTB を有する被験者が組み入れられた。これらの被験者は、ベースライン時に外科的手術が計画され、治験責任医師又は治験分担医師により、重度の後遺症を伴うと判断された。また、これらの被験者の多くは、登録時に疾患の活動性が確認され、外科的手術が早急に必要であると判断された。

3.3.2 外科的に切除可能な GCTB 患者での手術施行状況

試験 20062004 のコホート 2 では、90.0% (90/100 名) の被験者が第 3 回中間解析のデータカットオフ日時点で、手術が施行されなかったか、又はベースライン時に計画された手術よりも軽度の手術が施行された (試験 20062004 Interim Analysis 3 の治験総括報告書 Table 14-4.2.5)。手術が計画された 100 名のうち、実際に試験中に手術が施行された被験者は 26 名であった。ベースライン時に計画された手術と実際に施行された手術の分布を表 2.7.3.3-15 に、ベースライン時に計画された手術と実際に施行された手術の一覧を表 2.7.3.7-1 示す。試験中に手術が施行された 26 名のうち、16 名ではベースライン時に計画された手術よりも軽度の手術が施行され、9 名ではベースライン時に計画された手術と同じ手術が施行された。残り 1 名

(005*) はベースライン時に骨と軟部組織障害を治療するために骨盤及び腸骨稜で掻爬及び切除 (軟部組織切除) が計画されていたが、切除及び辺縁摘出が実施され、一塊切除となった。なお、第 3 回中間解析のデータカットオフ日時点で手術が施行されなかったのは 74 名であり、計画されていた手術は、掻爬 (9 名)、辺縁切除 (1 名)、一塊切除 (4 名)、一塊摘出 (26 名)、関節切除 (10 名)、人工関節置換術 (5 名)、骨盤半截術 (3 名)、切断術 (15 名)、及びその他骨格手術 (1 名) であった。

表 2.7.3.3-15 ベースライン時に計画された手術と実際に施行された手術: 試験 20062004

Surgical Procedure	Baseline Planned (N = 100)	Actual Total (N = 26)
Total number of surgeries	100	26
Curettage	13	16
Marginal excision, en bloc excision, or en bloc resection	42	6
Major surgeries	44	3
Joint resection	14	2
Joint/prosthesis replacement	9	1
Amputation	17	0
Hemipelvectomy	4	0
Other	1	1

Page 1 of 1

N = number of enrolled Cohort 2 subjects who were eligible for the study and received at least one dose of denosumab

Source: Table 14-4.2.4 of 20062004 Interim Analysis 3 CSR

試験期間は、第 3 回中間解析時点、コホートの 2 の被験者の 68.3% (69/101 名) が 6 ヶ月以上、30.7% (31/101 名) が 12 ヶ月以上であった (試験 20062004 Interim Analysis 3 の治験総

括報告書 Table 14-3.2.1)。ベースライン時に計画された手術と実際に施行された手術について、試験期間別に表 2.7.3.7-2 に示す。手術結果は、26 名中 1 名（ 015* ）で部分切除、その他全被験者で完全切除であった。治験実施計画書に従い、完全切除の後、デノスマブが 6 回投与された。

試験 AMG162-B-J201 では、切除可能な GCTB を有する被験者が 4 名登録されたが、20 年 月 日のデータカットオフ日時点で手術が施行された被験者はいなかった（試験 AMG162-B-J201 治験総括報告書 Table 15.2-2.6.1）。試験期間は、データカットオフ日時点で、2 名の被験者が 6 ヶ月以上、うち 1 名が 12 ヶ月以上であった。

3.3.3 手術までの期間

試験 20062004 では、試験期間中に手術が施行された被験者 26 名の、手術までの期間の中央値の Kaplan-Meier 推定値は、723 日であった（表 2.7.3.7-3）。

試験 AMG162-B-J201 では、切除可能な被験者 4 名のうち、データカットオフ日時点で手術が施行された被験者はいなかったため、手術までの期間の中央値は推定できなかった。

3.3.4 外科的に切除可能な GCTB 患者での臨床的有益性

試験 20062004 では、コホート 2 の被験者で、試験中に手術が施行されたかどうかに関わらず、デノスマブ投与による臨床的有益性（痛みの減少、可動性の改善、機能の改善など）が認められた。また、独立読影機関による客観的抗腫瘍効果の評価には、試験 20062004 のコホート 2 も含めている。これは、レトロスペクティブな評価であるため、独立読影機関による評価に 100 名全員の画像は用いられなかった。コホート 2 では、49 名の被験者でベースラインと 1 時点以上の試験期間中の画像が評価可能であった。コホート 2 の被験者における客観的奏効率は、評価可能であった全被験者における客観的奏効率と同等であった（81.6% [40/49 名] 対 71.6% [136/190 名]）。ベースライン時に手術が計画され、客観的奏効を認めた 40 名のうち、31 名は手術が施行されなかった。9 名は試験中に手術が施行され、5 名では計画された手術よりも軽度の手術、3 名では計画どおりの手術が施行された。残り 1 名（ 005* ）は計画とは異なる外科手術が施行された（表 2.7.3.7-1）。

試験 AMG162-B-J201 では、外科的に手術可能な被験者 4 名全員で、臨床的有益性が認められた。また、外科的に手術可能な被験者の客観的奏効率は、本試験の全被験者における客観的奏効率と同等であった（75.0% [3/4 名] 対 82.4% [14/17 名]）。

3.4 個々の試験ごとの結果

個々の試験の結果として、病理組織学的又は画像評価による抗腫瘍効果、並びに切除不能及び切除可能な GCTB 患者における臨床的有益性が示された。試験 20062004 及び AMG162-B-J201 における痛みの評価結果も、GCTB 患者におけるデノスマブ投与の全体的な臨床的有益性を裏付けるものであった。以下には、表 2.7.3.1-2 に示した評価項目のうち、客観的奏効に関する項目以外の評価項目について、各試験の結果を要約する。なお、外科的に切除

可能な GCTB 患者に対するデノスマブの有効性は、[第 3.3 項](#)に示した。

- 試験 20040215 では、有効性解析対象集団に組み入れられた 35 名の被験者のうち 85.7% (95% CI: 69.7~95.2) において、治験実施計画書で定義された治療効果が認められた。病理組織学的評価を実施した 20 名全員 (100.0%) で治療効果がみられ (すなわち、デノスマブ投与後に 90%以上の巨細胞が消失)、病理組織学的評価を実施できず画像評価を行った 15 名では 10 名 (66.7%) で治療効果が認められた (すなわち、治験責任医師による画像評価において、第 25 週時点で病勢の進行なし [病変の長径の増加率が 20%未満]) (試験 20040215 治験総括報告書、[第 9.2.1 項](#); 第 7 項の表 [2.7.3.7-4](#) 及び表 [2.7.3.7-5](#))。
- 試験 AMG162-B-J201 で、試験 20040215 と同じ定義を用い治療効果を評価した結果、有効性解析対象集団に組み入れられた 17 名の被験者のうち 94.1% (95% CI: 71.3~99.9) に治療効果が認められた (試験 AMG162-B-J201 治験総括報告書 [Table 15.2-2.13.1](#))。なお、病理組織学的評価を実施した被験者はおらず、全被験者において画像評価で治療効果を評価した (試験 AMG162-B-J201 治験総括報告書 [Table 15.2-2.10.1](#))。
- 試験 20062004 では第 3 回中間解析の時点で、コホート 1 (切除不能な GCTB) の 169 名中 6 名 (3.6%) において、治験責任医師の判断による病勢の進行が認められた。病勢の進行までの期間は中央値に達しなかった (試験 20062004 治験総括報告書、[第 9.2.1 項](#))。コホート 1 で投与を受けた被験者における追跡調査期間の中央値は 13 ヶ月であった (第 7 項、表 [2.7.3.7-6](#))。
- 試験 AMG162-B-J201 では、治験責任医師の判断による病勢の進行が認められた被験者はいなかった (試験 AMG162-B-J201 治験総括報告書、[第 11.4.1.1.14 項](#))。投与を受けた被験者における追跡調査期間の中央値は 7.6 ヶ月であった (試験 AMG162-B-J201 治験総括報告書 [Table 15.3-1.1](#))。
- 試験 20062004 ではコホート 1 と 2 を併せ、治験責任医師が報告した最良効果が SD 以上であった被験者は 252 名中 250 名 (99.2%) であった。コホート 3 (試験 20040215 から移行した被験者) では、疾患の状態が評価された 11 名のうち 2 名は治験責任医師が報告した最良効果が PR であり、9 名は SD であった (試験 20062004 治験総括報告書、[第 9.2.3.1 項](#))。
- 試験 AMG162-B-J201 では、治験責任医師が報告した最良効果が SD 以上であった被験者は 17 名中 17 名 (100%) であった。このうち 13 名は治験責任医師が報告した最良効果が PR であり、4 名は SD であった (試験 AMG162-B-J201 治験総括報告書、[Table 15.2-2.4.1](#))。
- 試験 20040215 では、デノスマブを投与後の画像評価で、骨石灰化の増加及び原発腫瘍における骨修復の所見が認められた。また、治験責任医師の評価による臨床的有益性 (31 名中 26 名 [83.9%]、95% CI: 66.3~94.5) も認められた (試験 20040215 治験総括報告書、[第 9.2.3.3 項](#))。
- 試験 20062004 ではコホート 1、2、及び 3 において、それぞれ 39.6%、61.0%、及び 27.3%

の被験者で、被験者の報告による臨床的有益性が認められた。各コホートで最も多く報告された臨床的有益性は痛みの減少であった（試験 20062004 治験総括報告書、[第 9.2.3.4 項](#)）。

- 試験 AMG162-B-J201 では、82.4%の被験者で治験責任医師の評価による臨床的有益性が認められた。多く報告された臨床的有益性は痛みの減少であった（試験 AMG162-B-J201 治験総括報告書、[第 11.4.1.1.15 項](#)）。
- 試験 20062004 では、評価可能な被験者の多くで、デノスマブ投与による速やかな痛みの改善が認められた。ベースライン時に最もひどい痛みのスコアが 2 ポイント以上であった被験者の 31.4%では、投与開始後 1 週間以内に、臨床的に意味のある最もひどい痛みの減少が認められた。また、これらの被験者の 50%以上では、第 5 週以降の全評価時点で臨床的に意味のある最もひどい痛みの減少が認められ、痛みの改善は、試験期間中にわたって持続した（試験 20062004 治験総括報告書、[第 10.2 項](#)）。事後解析の結果、評価可能な被験者の半数以上では、第 9 週以降の全評価時点において、最もひどい痛みのスコアがベースラインから 30%以上減少した。（[第 7 項](#)、[表 2.7.3.7-7](#)）。平均鎮痛薬スコアは、試験期間を通して低かったため（2 点以下）、痛みの改善は鎮痛薬使用の増加に関連したものではないと考えられた。なお、PRO 解析対象集団において、ベースライン時の鎮痛薬使用がなし又は少量であり、いずれかの来院時点で強力なオピオイド使用に移行した被験者の割合は、全被験者の 4.7%以下であった（試験 20062004 治験総括報告書、[第 10.2 項](#)）。
- 試験 AMG162-B-J201 では、評価可能な被験者の多くで、デノスマブ投与による速やかな痛みの改善が認められた。ベースライン時に最もひどい痛みのスコアが 2 ポイント以上であった被験者の 25%では、投与開始後 1 週間以内に、臨床的に意味のある最もひどい痛みの減少が認められた。また、50%以上の被験者では、第 5 週以降に臨床的に意味のある最もひどい痛みの減少が認められ、痛みの改善は、試験期間中にわたって持続した（試験 AMG162-B-J201 治験総括報告書、[第 11.4.1.1.11 項](#)）。PRO 解析対象集団において、ベースライン時の鎮痛薬使用がなし又は少量であり、いずれかの来院時点で強力なオピオイド使用に移行した被験者は 1 名であった（試験 AMG162-B-J201 治験総括報告書、[Table 15.2-5.1](#)）。
- 試験 20040215 では、負荷投与終了時（第 5 週）の血清中デノスマブ濃度は初回投与後の約 2 倍であり、負荷投与を加えた投与スケジュールによって、全身曝露量は予想した目標値まで増加したことが示された。第 9 週から第 49 週までのトラフ濃度の中央値の変動は 9%未満であり、Q4W の投与期間中、曝露量は安定していた。すなわち、デノスマブの薬物動態は反復投与時も経時的に変化しなかった（試験 20040215FA 治験総括報告書、[第 7.1 項](#)）。試験 20040215 で得られた最新の薬物動態に関する結果の概要（幾何平均値及び変動係数を含む）を、治験総括報告書 20040215 の[補遺](#)に示した。
- 試験 AMG162-B-J201 では、負荷投与終了時（第 5 週）の血清中デノスマブ濃度は初回投与後の約 2.5 倍であり、負荷投与を加えた投与スケジュールによって、全身曝露量は

予想した目標値まで増加したことが示された。第9週から第49週までのトラフ濃度の中央値の変動は4.3%未満であり、Q4Wの投与期間中、曝露量は安定していた。すなわち、デノスマブの薬物動態は反復投与時において経時的に変化しなかった（試験 AMG162-B-J201 治験総括報告書、[第11.4.1.2項](#)）。

これらの薬物動態解析及び有効性解析の結果は、GCTB患者に対するデノスマブ投与（120 mg Q4W、第8日及び第15日に120 mgを負荷投与）の臨床的に意味のある治療効果を裏付けるものである。また、これらの結果は、独立した客観的抗腫瘍効果の解析において、デノスマブ投与による腫瘍退縮効果が示されたことから裏付けられた（[第3.2項](#)）。

3.5 部分集団における結果の比較

成人被験者（18歳以上）のみを組み入れた試験 20040215 では、部分集団解析を実施しなかった。成人被験者と骨格が成熟した未成年被験者を組み入れた試験 20062004 では、年齢による部分集団解析を実施した。第3回中間解析のデータカットオフ日の時点で、組み入れられた10名の未成年被験者（12歳以上、18歳未満）における有効性の結果は、試験 20062004 の全被験者の結果と同様であると考えられた（試験 20062004 治験総括報告書、[第9.3項](#)）。

独立した客観的抗腫瘍効果の評価は、試験 20062004 の未成年集団でも実施した。客観的抗腫瘍効果の解析のための評価可能な画像データが1時点以上で得られた6名の未成年被験者を対象とした。第3回中間解析のデータカットオフ日の時点で6名全員が試験を継続し、デノスマブ投与を継続していた（[Table tiae1.50.2](#) 及び [tiae1.50.3](#)）。6名はいずれも切除不能な GCTB であった（[Table tiae2.50.2](#)）。6名のベースライン時の人口統計学的特性を [Table tiae2.50.1](#) に示す。客観的抗腫瘍効果の解析対象集団とした未成年被験者での試験期間の中央値は6.3ヵ月であった（[Table tiae3.50.1](#)）。デノスマブ投与回数の中央値は9.5回（最大値: 21回）であった。あらゆる評価基準を用いた最良効果に基づく客観的奏効率は66.7%（4/6）であった（[Table tiae4.50.1](#)）。個々の効果判定基準ごとの客観的奏効率は、modified RECIST 基準で33.3%（2/6）、density/size 基準で66.7%（4/6）であり、modified EORTC 基準で評価可能な未成年被験者はいなかった（[Table tiae4.50.1](#)）。

試験 AMG162-B-J201 では、12ヵ月カットオフ日までに、1名の日本人未成年（20歳未満）被験者（18歳）が組み入れられた。当該被験者では、あらゆる評価基準を用いた最良効果に基づく評価で客観的奏効が認められた。なお、当該被験者の試験期間は約9ヵ月であった。

3試験併合解析では、20歳未満の被験者（n = 15）を対象として解析を実施した（試験 20040215 及び 20062004 の18又は19歳の被験者8名を含む）。その結果、客観的抗腫瘍効果は外国2試験併合解析の結果と概ね同様であった（[Table 1tiae4-1.2](#)）。

3.6 全試験を通しての結果の概要

試験 20040215、20062004、及び AMG162-B-J201 の対象患者集団は、今回申請する適応症の患者集団全体を代表するものであり、このまれな疾患としては大規模な臨床データセットが得られた。外国2試験併合解析では、独立した画像評価によって、客観的抗腫瘍効果の評価を実

施した。客観的抗腫瘍効果に関する画像評価の解析対象集団には、両試験に参加した被験者又は現在参加中の被験者のうち、評価可能な 190 名を含めた。AMG162-B-J201 においても、同じ評価方法を用いて、客観的抗腫瘍効果を評価した（解析対象 17 名）。

デノスマブは、GCTB 患者の管理において臨床的に意味のある有効性を示した。外国 2 試験併合解析では、あらゆる評価基準を用いた最良効果に基づく客観的奏効率は 71.6%であった。骨腫瘍に対して、腫瘍サイズの縮小に基づく RECIST 基準を適用するには限界があるが、Modified RECIST 基準により 25.1%の奏効率が得られた。RECIST 基準では、軟部組織病変（又は軟部組織成分を含む病変）を有する被験者の 57.1%で客観的奏効が得られた。Modified EORTC 基準及び density/size 基準を用いたときの奏効率は、それぞれ 96.2%及び 76.1%であった。客観的奏効には持続性が認められ、客観的奏効が確認された後に病勢が進行した被験者は 1 名のみであった。客観的奏効は、試験又はベースライン時の疾患の状態（切除不能又は切除可能）にかかわらず、一貫して認められた。

試験 20040215 及び 20062004 の各試験から得られた有効性の結果は、外国 2 試験併合解析における独立した客観的抗腫瘍効果の評価の結果を裏付けるものであった。これらの各試験では評価した多くの被験者において、デノスマブによる明らかな治療効果（病理組織学的評価又は治験責任医師による画像評価による腫瘍の退縮、臨床症状の改善、計画していた外科的手術までの期間の延長、及び計画していた外科的手術の後遺症の低下）が認められた。

以上のように、試験 20040215 及び 20062004 では客観的抗腫瘍効果に関する独立した画像評価の結果、多くの被験者で臨床的に意味のあるベネフィットが認められ、個々の臨床試験においても、治験責任医師の評価によりベネフィットが確認された。骨格が成熟した未成年被験者における有効性の評価結果は、成人被験者での結果と同様であった。デノスマブによる治療効果は、切除不能な GCTB と切除可能な GCTB のいずれの被験者においても認められた。切除可能な GCTB 患者でベースライン時に計画されていた手術の多くは、片側骨盤切断、手足の切断、又は関節切除のような重度の後遺症を伴うものであった。デノスマブの投与により、切除可能な GCTB 患者の多くで、試験期間中に手術が施行されなかったか、ベースライン時に計画していた手術よりも侵襲性の低い手術が施行された。また、切除可能な GCTB 患者においても、臨床的有益性、疼痛の改善、客観的抗腫瘍効果が認められた。このように、両試験のデータから、GCTB 患者の治療におけるデノスマブ（120 mg Q4W に加えて第 8 日及び第 15 日に 120 mg の負荷投与）の良好なベネフィット／リスクプロファイルが示された。

試験 AMG162-B-J201 においても、外国 2 試験併合解析と同様の有効性評価項目について検討した結果、あらゆる評価基準を用いた最良効果に基づく客観的奏効率は 82.4%であり、modified RECIST 基準、modified EORTC 基準、及び density/size 基準によって客観的奏効を認めた被験者の割合は、それぞれ 29.4%、70.6%、及び 64.7%であった。客観的奏効には持続性が認められ、客観的奏効が確認された後に病勢が進行した被験者はいなかった。客観的奏効は、ベースライン時の疾患の状態（切除不能又は切除可能）にかかわらず、一貫して認められた。試験 AMG162-B-J201 での客観的抗腫瘍効果（6 ヶ月カットオフ解析結果）と、外国 2 試験併合解析による客観的抗腫瘍効果は一貫していた。デノスマブの治療効果は切除可能な GCTB

患者においても認められた。6 ヶ月カットオフ時点で手術が施行された患者は認められず、また、切除可能な GCTB 患者においても、臨床的有益性及び客観的抗腫瘍効果が認められた。その他の有効性評価項目においても、デノスマブによる治療効果（病勢の進行なし、臨床症状の改善、計画していた外科的手術までの期間の延長）が認められ、これは客観的抗腫瘍効果の結果を裏付けるものであった。

4. 推奨用法・用量に関する臨床情報の解析

試験 20040215 及び 20062004 では、デノスマブ 120 mg Q4W の皮下投与に加え、第 8 日及び第 15 日の 120 mg 負荷投与を行う用法・用量とした。120 mg Q4W は、癌骨転移患者における骨関連事象 (skeletal-related event(s): SRE) の抑制を適応として承認されたデノスマブの用法・用量と同じである。GCTB 患者においても、定常状態の血清中デノスマブ濃度の目標値を癌骨転移患者と同程度とするのが妥当であると考えた。いずれの患者集団でも目標は、投与した患者の大多数 (95%超) において、投与間隔を通じて RANKL への最大結合及び阻害を得ることである (癌骨転移患者における SRE の抑制に関する承認申請 [ランマーク SRE 申請] 資料のモジュール 5.3.3.5、第 7 項—進行がん患者集団における薬物動態報告書 111914)。また、GCTB 患者における用法・用量は、血清中デノスマブ濃度が速やかに定常状態に到達し、かつ曝露量がこれまでの第 II 相及び第 III 相臨床試験で検討された範囲内にとどまるように設定する必要があると考えた。Q4W の投与スケジュールを裏付ける薬物動態プロファイル及び薬力学プロファイルに関する考察は、ランマーク SRE 申請時の「臨床薬理の概要 (モジュール 2.7.2)」に示しており、以下にはさらなる考察を示した。

癌骨転移患者とは異なり、GCTB 患者における治療目的は SRE を抑制することではなく、直接的な抗腫瘍効果を得ることである。GCTB 患者は痛みを含む重度の合併症を有するため、投与開始後 1 ヶ月以内に目標濃度に達することが望ましいと考えられた。これは、癌骨転移患者とは異なる点である。癌骨転移患者では、120 mg Q4W により約 4～6 ヶ月で血清中デノスマブ濃度が定常状態に達する。試験 20040215 では、120 mg Q4W に第 8 日と第 15 日に 120 mg の負荷投与を加えることによって、投与開始後 1 ヶ月以内に目標濃度に達するという目的が達成され、第 15 日及び第 29 日における血清中デノスマブ濃度トラフ値の中央値は、第 9～49 週の定常状態と同程度 (それぞれ < 16%～35% の差) であった (試験 20040215 治験総括報告書 [主要解析] の第 10.1 項及び 20040215 治験総括報告書補遺)。

国内で実施した試験 AMG162-B-J201 も、120 mg Q4W に第 8 日及び第 15 日の 120 mg 負荷投与を加えた用法・用量とした。試験 AMG162-B-J201 (6 ヶ月カットオフ解析) の血清中デノスマブ濃度のトラフ値に、試験 20040215 と顕著な差は認められなかった (図 2.7.3.4-1)。また、本試験では、骨代謝マーカーである uNTX/Cr の抑制 (約 80%) が第 5 週から認められ、試験 20040215 の結果と類似していた (図 2.7.3.4-2)。

以上のことから、120 mg Q4W に第 8 日及び第 15 日の 120 mg 負荷投与を加えた用法・用量は、国内外の GCTB 患者において適切と考えた。

図 2.7.3.4-1 日本人 GCTB 患者（試験 AMG162-B-J201）及び外国人 GCTB 患者（試験 20040215）における血清中デノスマブトラフ濃度－時間プロファイル

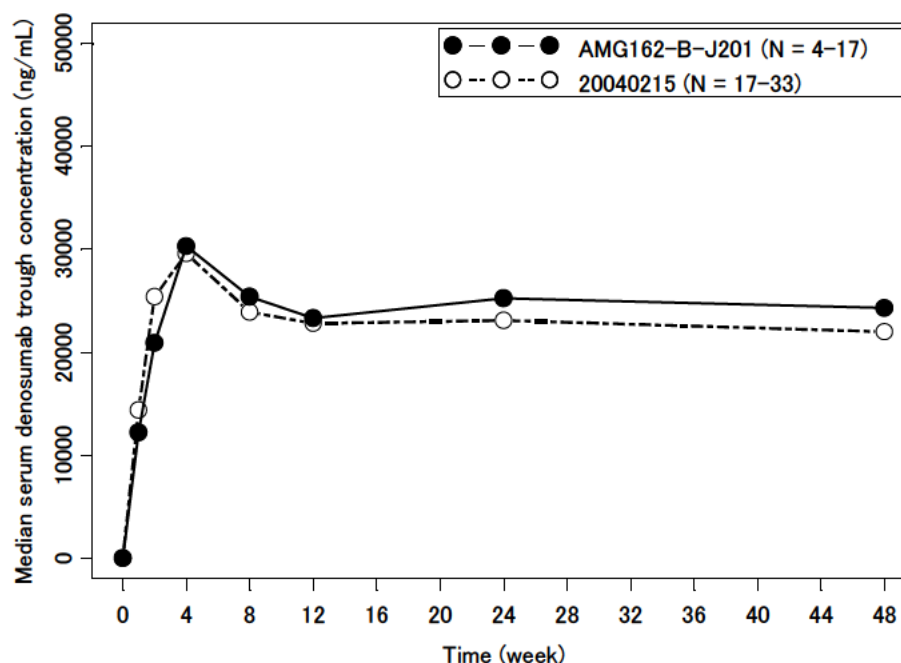
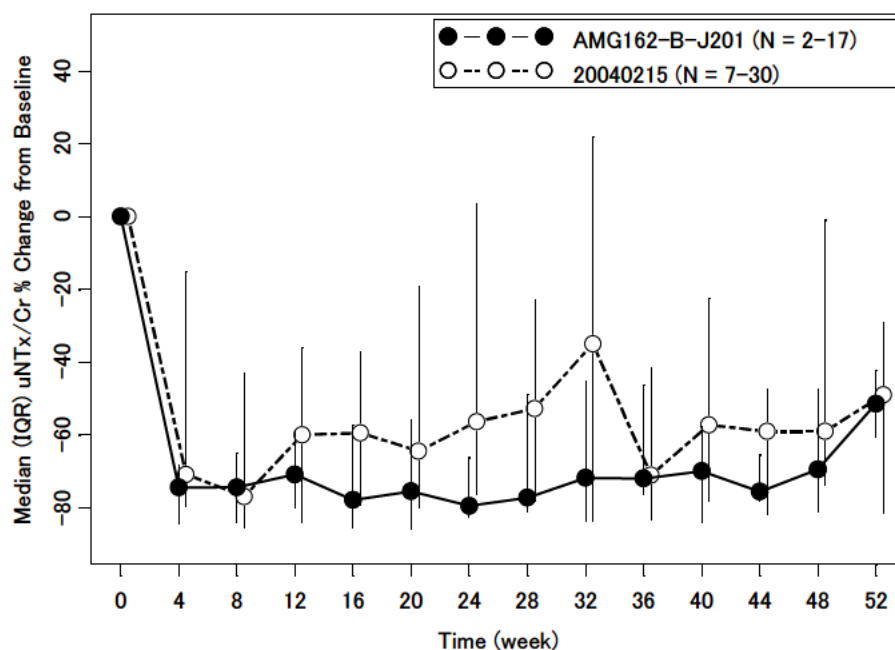


図 2.7.3.4-2 日本人 GCTB 患者（試験 AMG162-B-J201）及び外国人 GCTB 患者（試験 20040215）における uNTX/Cr のベースラインからの変化率の中央値（Q1, Q3）



5. 効果の持続、耐薬性

GCTB 患者において持続的な効果を得るためには長期投与が必要であるため、長期曝露時のデノスマブの有効性を評価することは重要である。

試験 20040215 及び 20062004 に含まれた GCTB 患者のうち、デノスマブ曝露が 2 年以上であった被験者は 46 名、3 年以上であった被験者は 15 名であった (Table 5.1.2)。試験 20040215 及び 20062004 における最大投与回数は 60 回であり、最長試験期間は 54.1 ヶ月であった (Table 5.1.3)。客観的抗腫瘍効果のレトロスペクティブな評価に組み入れられた被験者で、デノスマブ曝露期間が 2 年を超えた被験者は 29 名、3 年を超えた被験者は 7 名であった (ad hoc Table 3.1.500)。

外国 2 試験併合解析として客観的抗腫瘍効果に関する独立した画像評価を実施した結果、デノスマブの効果は長期間持続することが示された。客観的抗腫瘍効果の併合解析対象集団において、あらゆる評価基準を用いていずれかの時点で PD がみられた被験者は少数であった

(n = 11) (Table 4.3.3)。客観的奏効が得られた後に PD が認められた被験者は、試験 20062004 の 1 名のみであった (Table 4.2.1)。第 3.2.2.3 項で示したように、68.5%の被験者では、あらゆる評価基準を用いた最良効果に基づく客観的奏効が 24 週間持続した (表 2.7.3.3-8)。客観的奏効が 4、8、及び 12 週間持続した被験者の割合も同程度であった (それぞれ 66.7%、67.3%、及び 68.1%)。

試験 AMG162-B-J201 では、12 ヶ月カットオフ解析を実施した時点で、デノスマブ曝露が 6 ヶ月以上の被験者は 17 名、1 年以上の被験者は 10 名であり、最大投与回数は 22 回、最長試験期間は 17.9 ヶ月であった (試験 AMG162-B-J201 治験総括報告書 Table 15.3-1.1)。あらゆる評価基準を用いた最良評価に基づく客観的奏効が 4、8、12、又は 24 週間持続した被験者の割合は、それぞれ 88.2% (15/17)、88.2% (15/17)、82.4% (14/17)、又は 86.7% (13/15) であった (Table 15.2-1.4.1)。あらゆる評価基準を用いて客観的奏効が認められた 15 名のうち、1 名の被験者がその後 PD と判定された (Table 15.2-1.2.1)。この PD と判定された被験者では、CT/MRI により新たな病変 (神経膠芽細胞腫 glioblastoma) が認められており、このため治験責任医師は、この被験者を PD でないと判定した。試験 AMG162-B-J201 の 12 ヶ月カットオフ解析の有効性の結果は、6 ヶ月カットオフ解析の結果と大きな違いは認められなかった。なお、本試験の被験者におけるデノスマブの投与は、本剤の GCTB への適用が国内で承認されるまで継続する予定である。また、試験 20040215、20062004、及び AMG162-B-J201 (12 ヶ月カットオフ解析時)における客観的抗腫瘍効果の併合解析の結果は、モジュール 5.3.5.3-10 に示す。

デノスマブを 2～5 年間投与した実績がある他の疾患では、効果が減弱するというエビデンスは得られていない (ランマーク SRE 申請時のモジュール 2.7.3・第 5.1 項及び骨粗鬆症に関する承認申請時のモジュール 2.7.3・第 5.1 項)。

6. 参考文献

- Becker WT, Dohle J, Bernd L, et al. Local recurrence of giant cell tumor of bone after intralesional treatment with and without adjuvant therapy. *J Bone Joint Surg Am*. 2008;90:1060-7.
- Branstetter D, Nelson SD, Manivel JC, et al. Denosumab induces tumor reduction and bone formation in patients with giant cell tumor of bone. *Clin Can Res*. 2012;18:4415-24.
- Caudell JJ, Ballo MT, Zagars GK, et al. Radiotherapy in the management of giant cell tumor of bone. *Int J Radiat Oncol Biol Phys*. 2003;57:158-65.
- Choi H, Charnsangavej C, Faria SC, et al. Correlation of computed tomography and positron emission tomography in patients with metastatic gastrointestinal stromal tumor treated at a single institution with imatinib mesylate: proposal of new computed tomography response criteria. *J Clin Oncol*. 2007;25:1753-9.
- Civista Health Library. Bone Tumors: Benign. Available from: <http://www.health-news-and-information.com/4civista/libv/c33.shtml>. Accessed April 2011.
- Costelloe CM, Chuang HH, Madewell JE, et al. Cancer Response Criteria and Bone Metastases: RECIST 1.1, MDA and PERCIST. *J Cancer*. 2010;1:80-92.
- Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer*. 2009;45:228-47.
- Kostenuik PJ, Nguyen HQ, McCabe J, et al. Denosumab, a fully human monoclonal antibody to RANKL, inhibits bone resorption and increases BMD in knock-in mice that express chimeric (murine/human) RANKL. *J Bone Miner Res*. 2009;24:182-95.
- Leggon RE, Zlotecki R, Reith J, et al. Giant cell tumor of the pelvis and sacrum: 17 cases and analysis of the literature. *Clin Orthop Relat Res*. 2004;423:196-207.
- Malawer M, Helman L, O'Sullivan B. In: *Cancer: Principles and Practice of Oncology*, 9th Edition. Sarcomas of Bone - Chapter 116. Lippincott, Williams, and Wilkins;2011.
- Malawer M, Helman L, O'Sullivan B. In: *Cancer: Principles & Practice of Oncology*, 7th Edition. Sarcomas of the Soft Tissues and Bone - Chapter 35. Lippincott Williams & Wilkins. 2005.
- Singer S, Maki R, O'Sullivan B. In: *Cancer: Principles and Practice of Oncology*, 9th Edition. Soft Tissue Sarcoma – Chapter 115. Lippincott, Williams, and Wilkins;2011.
- Stacchiotti S, Collini P, Messina A, et al. High-grade soft-tissue sarcomas: Tumor response assessment—Pilot study to assess the correlation between radiologic and pathologic response using RECIST and Choi criteria. *Radiology*. 2009;251:447-56.
- Szendrői M. Giant-cell tumour of bone [Review]. *J Bone Joint Surg (British Volume)*. 2004;86B:5-12.
- Szendrői M, Kiss J, Antal I. Surgical treatment and prognostic factors in giant cell tumor of bone. *Acta Chir Orthop Traumatol Cech*. 2003;70:142-50.

Thomas D, Hensaw T, Skubitz K, et al. Denosumab in patients with giant-cell tumour of bone: an open-label, phase 2 study. *Lancet Oncology*. 2010;11:275-80.

Yasko AW. Giant cell tumor of bone. *Curr Oncol Rep*. 2002;4:520-6.

Young H, Baum R, Cremerius U, et al. Measurement of clinical and subclinical tumour response using [^{18}F]-fluorodeoxyglucose and positron emission tomography: review and 1999 EORTC recommendations. European Organization for Research and Treatment of Cancer (EORTC) PET Study Group. *Eur J Cancer*. 1999;35:1773-82.

Zerizer I, Al-Nahhas A, Towey D, et al. The role of early ^{18}F -FDG PET/CT in prediction of progression-free survival after ^{90}Y radioembolization: comparison with RECIST and tumour density criteria. *Eur J Nucl Med Mol Imaging*. 2012;39:1391-9.

日本整形外科学会骨軟部腫瘍委員会, 全国骨腫瘍登録一覧表 (平成 20 年度) ; 2008.

7. 付録

表 2.7.3.7-1 ベースライン時に計画された手術と実際に施行された手術の一覧:

試験 20062004

Subject ID	Baseline Planned Procedure	Actual Procedure
Subjects with Less Debilitating Surgical Procedure (n = 16)		
016*	Amputation	Curettage
017*	Joint prosthesis/replacement	Curettage
018*	Joint prosthesis/replacement	Curettage
019*	Joint prosthesis/replacement	Curettage
020*	En bloc resection	Other ^a
021*	Joint resection	En bloc resection
022*	Amputation	Curettage
023*	En bloc resection	Curettage
024*	Hemipelvectomy	Joint prosthesis/replacement
025*	En bloc resection	Curettage
026*	Joint resection	Curettage
027*	En bloc resection	Curettage
028*	En bloc resection	Curettage
029*	En bloc resection	Curettage
030*	En bloc resection	Curettage
031*	Joint prosthesis/replacement	Curettage
Subjects Receiving Originally Planned Surgical Procedure (n = 9)		
032*	Curettage	Curettage
033*	Curettage	Curettage
034*	Joint resection	Joint resection
035*	En bloc resection	En bloc resection
036*	Joint resection	Joint resection
037*	En bloc resection	En bloc resection
038*	En bloc resection	En bloc resection
015*	En bloc resection	En bloc resection
039*	Curettage	Curettage
Subjects with Different Surgical Procedure from Planned at Baseline (n = 1)		
005*	Curettage	En bloc resection

Page 1 of 1

^a Reported term for “other” surgery was sacral laminectomy and debulking of tumour. Considered a less morbid surgical procedure.

^b Subject 005* had a curettage and exeresis (ie, excision) planned at baseline to treat a bone and soft tissue lesion in the pelvis and iliac crest. On-study, he underwent a procedure of “exeresis plus marginal resection” classified in the database as an “en bloc resection” on study.

Source: Listing 102-2.1

表 2.7.3.7-2 ベースライン時に計画された手術と実際に施行された手術（試験期間別）：試験 20062004

On study	Baseline										Total
	No surgery planned	Curettage	Marginal excision	En bloc excision	En bloc resection	Joint resection	Joint/prosthesis replacement	Amputation	Hemipel-vectomy	Other	
At any time (N=100)	0 (0.0)	13 (13.0)	1 (1.0)	4 (4.0)	37 (37.0)	14 (14.0)	9 (9.0)	17 (17.0)	4 (4.0)	1 (1.0)	100 (100.0)
>0 - ≤6 months											
Cohort 2 (N = 100)											
No surgery	0 (0.0)	12 (12.0)	1 (1.0)	4 (4.0)	34 (34.0)	13 (13.0)	7 (7.0)	16 (16.0)	4 (4.0)	1 (1.0)	92 (92.0)
Curettage	0 (0.0)	1 (1.0)	0 (0.0)	0 (0.0)	1 (1.0)	0 (0.0)	2 (2.0)	1 (1.0)	0 (0.0)	0 (0.0)	5 (5.0)
Marginal excision	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
En bloc excision	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
En bloc resection	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (1.0)	1 (1.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (2.0)
Joint resection	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Joint/prosthesis replacement	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Amputation	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Hemipelvectomy	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Other	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (1.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (1.0)
Total	0 (0.0)	13 (13.0)	1 (1.0)	4 (4.0)	37 (37.0)	14 (14.0)	9 (9.0)	17 (17.0)	4 (4.0)	1 (1.0)	100 (100.0)
>6 - ≤12 months											
Cohort 2 (N = 61)											
No surgery	0 (0.0)	4 (6.6)	1 (1.6)	3 (4.9)	17 (27.9)	4 (6.6)	2 (3.3)	12 (19.7)	1 (1.6)	1 (1.6)	45 (73.8)

Page 1 of 5

n (%)

N = Number of subjects who had the opportunity to be on study for each time interval and did not have on-study surgery before the start of each time interval

Percentages based on N

Modified from Table 102-2.27.1 and Table 14.

表 2.7.3.7-2 ベースライン時に計画された手術と実際に施行された手術（試験期間別）：試験 20062004（続き）

On study	Baseline										Total
	No surgery planned	Curettage	Marginal excision	En bloc excision	En bloc resection	Joint resection	Joint/prosthesis replacement	Amputation	Hemipel-vectomy	Other	
At any time (N=100)	0 (0.0)	13 (13.0)	1 (1.0)	4 (4.0)	37 (37.0)	14 (14.0)	9 (9.0)	17 (17.0)	4 (4.0)	1 (1.0)	100 (100.0)
>6 - ≤12 months											
Cohort 2 (N = 61)											
Curettage	0 (0.0)	2 (3.3)	0 (0.0)	0 (0.0)	4 (6.6)	1 (1.6)	2 (3.3)	0 (0.0)	0 (0.0)	0 (0.0)	9 (14.8)
Marginal excision	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
En bloc excision	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
En bloc resection	0 (0.0)	1 (1.6)	0 (0.0)	0 (0.0)	3 (4.9)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	4 (6.6)
Joint resection	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (3.3)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	2 (3.3)
Joint/prosthesis replacement	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (1.6)	0 (0.0)	1 (1.6)
Amputation	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Hemipelvectomy	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Other	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Total	0 (0.0)	7 (11.5)	1 (1.6)	3 (4.9)	24 (39.3)	7 (11.5)	4 (6.6)	12 (19.7)	2 (3.3)	1 (1.6)	61 (100.0)
>12 - ≤18 months											
Cohort 2 (N = 19)											
No surgery	0 (0.0)	1 (5.3)	1 (5.3)	2 (10.5)	8 (42.1)	1 (5.3)	0 (0.0)	4 (21.1)	0 (0.0)	1 (5.3)	18 (94.7)
Curettage	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (5.3)	0 (0.0)	0 (0.0)	1 (5.3)

Page 2 of 5

n (%)

N = Number of subjects who had the opportunity to be on study for each time interval and did not have on-study surgery before the start of each time interval

Percentages based on N

Modified from Table 102-2.27.1 and Table 14.

表 2.7.3.7-2 ベースライン時に計画された手術と実際に施行された手術（試験期間別）：試験 20062004（続き）

On study	Baseline										Total
	No surgery planned	Curettage	Marginal excision	En bloc excision	En bloc resection	Joint resection	Joint/prosthesis replacement	Amputation	Hemipel-vectomy	Other	
At any time (N=100)	0 (0.0)	13 (13.0)	1 (1.0)	4 (4.0)	37 (37.0)	14 (14.0)	9 (9.0)	17 (17.0)	4 (4.0)	1 (1.0)	100 (100.0)
>12 - ≤18 months											
Cohort 2 (N = 19)											
Marginal excision	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
En bloc excision	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
En bloc resection	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Joint resection	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Joint/prosthesis replacement	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Amputation	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Hemipelvectomy	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Other	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Total	0 (0.0)	1 (5.3)	1 (5.3)	2 (10.5)	8 (42.1)	1 (5.3)	0 (0.0)	5 (26.3)	0 (0.0)	1 (5.3)	19 (100.0)
>18 - ≤24 months											
Cohort 2 (N = 6)											
No surgery	0 (0.0)	1 (16.7)	0 (0.0)	0 (0.0)	3 (50.0)	0 (0.0)	0 (0.0)	1 (16.7)	0 (0.0)	0 (0.0)	5 (83.3)
Curettage	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (16.7)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (16.7)
Marginal excision	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)

Page 3 of 5

n (%)

N = Number of subjects who had the opportunity to be on study for each time interval and did not have on-study surgery before the start of each time interval

Percentages based on N

Modified from Table 102-2.27.1 and Table 14

表 2.7.3.7-2 ベースライン時に計画された手術と実際に施行された手術（試験期間別）：試験 20062004（続き）

On study	Baseline										Total
	No surgery planned	Curettage	Marginal excision	En bloc excision	En bloc resection	Joint resection	Joint/prosthesis replacement	Amputation	Hemipel-vectomy	Other	
At any time (N=100)	0 (0.0)	13 (13.0)	1 (1.0)	4 (4.0)	37 (37.0)	14 (14.0)	9 (9.0)	17 (17.0)	4 (4.0)	1 (1.0)	100 (100.0)
>18 - ≤24 months											
Cohort 2 (N = 6)											
En bloc excision	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
En bloc resection	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Joint resection	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Joint/prosthesis replacement	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Amputation	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Hemipelvectomy	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Other	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Total	0 (0.0)	1 (16.7)	0 (0.0)	0 (0.0)	4 (66.7)	0 (0.0)	0 (0.0)	1 (16.7)	0 (0.0)	0 (0.0)	6 (100.0)
>24 months											
Cohort 2 (N = 1)											
No surgery	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (100.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (100.0)
Curettage	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Marginal excision	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
En bloc excision	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)

Page 4 of 5

n (%)

N = Number of subjects who had the opportunity to be on study for each time interval and did not have on-study surgery before the start of each time interval

Percentages based on N

Modified from Table 102-2.27.1 and Table 14

表 2.7.3.7-2 ベースライン時に計画された手術と実際に施行された手術（試験期間別）：試験 20062004（続き）

On study	Baseline										Total
	No surgery planned	Curettage	Marginal excision	En bloc excision	En bloc resection	Joint resection	Joint/prosthesis replacement	Amputation	Hemipel-vectomy	Other	
At any time (N=100)	0 (0.0)	13 (13.0)	1 (1.0)	4 (4.0)	37 (37.0)	14 (14.0)	9 (9.0)	17 (17.0)	4 (4.0)	1 (1.0)	100 (100.0)
>24 months											
Cohort 2 (N = 1)											
En bloc resection	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Joint resection	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Joint/prosthesis replacement	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Amputation	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Hemipelvectomy	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Other	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Total	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (100.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (100.0)

Page 5 of 5

n (%)

N = Number of subjects who had the opportunity to be on study for each time interval and did not have on-study surgery before the start of each time interval

Percentages based on N

Modified from Table 102-2.27.1 and Table 14

表 2.7.3.7-3 手術までの期間: 試験 20062004

(Cohort 2 Efficacy Analysis Set)

(20062004 Interim Analysis 3)

	Crude Incidence n (%)	KM Estimate of 25%-tile (Days) ^a Pt Est (95% CI)	KM Estimate of Median (Days) ^a Pt Est (95% CI)	KM Estimate of 75%-tile (Days) ^a Pt Est (95% CI)
Cohort 2 (N = 100)	26 (26.0)	253.0 (211.00, 301.00)	723.0 (301.00, NE)	NE (723.00, NE)

Page 1 of 1

N = number of enrolled cohort 2 subjects who were eligible for the study and received at least one dose of denosumab

n = number of subjects who had a surgery

NE = Not estimable

^a Kaplan-Meier estimate

Source: Table 14-4.2.3 of 20062004 Interim Analysis 3 CSR

表 2.7.3.7-4 十分な病理組織学的評価が得られた被験者における GCTB への効果

(Efficacy Analysis Set)

(20040215 Primary Analysis)

	Denosumab 120 mg Q4W (N = 35)		
	n/N1	(%)	(95% CI)
Response ^a	20/20	(100.0)	(83.2, 100.0)

Page 1 of 1

N = Number of subjects who (1) are on study for ≥ 28 days after the first dose of denosumab; and (2) have at least one baseline tissue and at least one post-dose tissue between week 5 and week 25, or have at least one baseline radiograph and at least one post-dose radiograph between week 5 and week 25.N1 = Number of subjects who (1) are on study for ≥ 28 days after the first dose of denosumab; and (2) have at least one baseline tissue and at least one post-dose tissue between week 5 and week 25.

Confidence interval is calculated using exact method.

^a Response is defined per protocol.

表 2.7.3.7-5 十分な病理組織学的評価が得られなかった被験者における GCTB への効果
(Efficacy Analysis Set)
(20040215 Primary Analysis)

	Denosumab 120 mg Q4W (N = 35)		
	n/N1	(%)	(95% CI)
Response ^a	10/15	(66.7)	(38.4, 88.2)

Page 1 of 1

N = Number of subjects who (1) are on study for ≥ 28 days after the first dose of denosumab; and (2) have at least one baseline tissue and at least one post-dose tissue between week 5 and week 25, or have at least one baseline radiograph and at least one post-dose radiograph between week 5 and week 25.

N1 = Number of subjects who (1) are on study for ≥ 28 days after the first dose of denosumab; and (2) do not have at least one baseline tissue or do not have at least one post-dose tissue between week 5 and week 25; (3) have at least one baseline radiograph and at least one post-dose radiograph between week 5 and week 25.

Confidence interval is calculated using exact method.

^a Response is defined per protocol.



表 2.7.3.7-6 コホートごとの治験薬投与の要約
(Efficacy Analysis Set)
(20062004 3rd Interim Analysis)

	Cohort 1	Cohort 2	Cohort 3	All
Number of subjects enrolled	169	100	11	280
Number of months on study ^a				
n	169	100	11	280
Mean	13.29	9.28	5.41	11.55
SD	8.13	5.92	0.44	7.57
Median	13.04	8.95	5.36	10.43
Q1, Q3	5.88, 20.96	4.14, 12.94	5.32, 5.65	5.32, 16.82
Min, Max	0.3, 29.1	0.0, 28.0	4.5, 6.2	0.0, 29.1
Number of subjects receiving ≥ 1 dose of investigational product	169	100	11	280
Number of doses received				
n	169	100	11	280
Mean	16.3	12.0	5.9	14.3
SD	8.4	6.0	0.5	7.9
Median	16.0	12.0	6.0	13.0

Page 1 of 2

^a Defined as the time period from the first dose of investigational product, or enrollment date if subjects did not take any dose, to the end of study date or the analysis cutoff date, whichever comes first.

Denosumab is administered at a dose of 120 mg Q4W with a loading dose of 120 mg on study days 8 and 15.

Four subjects from study 20040215 were directly enrolled into the safety follow up phase in study 20062004 and were excluded from the treatment phase analysis.

[REDACTED]

表 2.7.3.7-6 コホートごとの治験薬投与の要約

(Efficacy Analysis Set)

(20062004 3rd Interim Analysis) (続き)

	Cohort 1	Cohort 2	Cohort 3	All
Q1, Q3	9.0, 24.0	7.0, 16.0	6.0, 6.0	7.0, 20.0
Min, Max	2, 33	1, 33	5, 7	1, 33
Number of doses received - n (%)				
1 - 5	18 (10.7)	13 (13.0)	2 (18.2)	33 (11.8)
6 - 10	29 (17.2)	28 (28.0)	9 (81.8)	66 (23.6)
11 - 15	37 (21.9)	33 (33.0)	0 (0.0)	70 (25.0)
16 - 20	30 (17.8)	17 (17.0)	0 (0.0)	47 (16.8)
21 - 25	24 (14.2)	7 (7.0)	0 (0.0)	31 (11.1)
26 - 30	25 (14.8)	1 (1.0)	0 (0.0)	26 (9.3)
31 - 35	6 (3.6)	1 (1.0)	0 (0.0)	7 (2.5)

Page 2 of 2

^a Defined as the time period from the first dose of investigational product, or enrollment date if subjects did not take any dose, to the end of study date or the analysis cutoff date, whichever comes first.

Denosumab is administered at a dose of 120 mg Q4W with a loading dose of 120 mg on study days 8 and 15.

Four subjects from study 20040215 were directly enrolled into the safety follow up phase in study 20062004 and were excluded from the treatment phase analysis.

[REDACTED]

表 2.7.3.7-7 最もひどい痛みのスコアがベースラインから 30%以上減少した被験者の割合: 試験 20062004

(Subject at Risk in PRO Analysis Set)

(Interim Analysis 3 Adhoc)

	Cohort 1 (N=123) % (n/N1)	Cohort 2 (N=80) % (n/N1)	Cohort 3 (N=4) % (n/N1)	Cohorts 1 and 2 (N=203) % (n/N1)	All (N=207) % (n/N1)
Day 8	26.6 (29/109)	36.4 (28/77)	-	30.6 (57/186)	30.6 (57/186)
Day 15	36.6 (41/112)	41.1 (30/73)	-	38.4 (71/185)	38.4 (71/185)
Week 5	36.9 (41/111)	53.6 (37/69)	25.0 (1/4)	43.3 (78/180)	42.9 (79/184)
Week 9	53.4 (55/103)	66.7 (44/66)	25.0 (1/4)	58.6 (99/169)	57.8 (100/173)
Week 13	50.5 (55/109)	75.4 (49/65)	25.0 (1/4)	59.8 (104/174)	59.0 (105/178)
Week 17	51.0 (50/98)	75.0 (42/56)	33.3 (1/3)	59.7 (92/154)	59.2 (93/157)
Week 21	52.1 (49/94)	69.8 (37/53)	50.0 (1/2)	58.5 (86/147)	58.4 (87/149)
Week 25	57.1 (48/84)	68.0 (34/50)	-	61.2 (82/134)	61.2 (82/134)
Week 37	56.4 (44/78)	78.4 (29/37)	-	63.5 (73/115)	63.5 (73/115)
Week 49	59.0 (36/61)	75.0 (18/24)	-	63.5 (54/85)	63.5 (54/85)
Week 61	52.1 (25/48)	69.2 (9/13)	-	55.7 (34/61)	55.7 (34/61)
Week 73	50.0 (17/34)	100.0 (3/3)	-	54.1 (20/37)	54.1 (20/37)
Week 85	60.0 (18/30)	75.0 (3/4)	-	61.8 (21/34)	61.8 (21/34)
Week 97	52.9 (9/17)	100.0 (1/1)	-	55.6 (10/18)	55.6 (10/18)
Week 109	55.6 (5/9)	-	-	55.6 (5/9)	55.6 (5/9)
Week 121	100.0 (1/1)	-	-	100.0 (1/1)	100.0 (1/1)

Page 1 of 1

N = number of enrolled subjects who were eligible for the study, received at least one dose of denosumab, used consistent recall periods on study, and had baseline worst pain score > 0

N1 = Number of subjects with data at the visit

Percentages are based on N1.

The range of worst pain is 0 - 10; a higher score indicates a less preferred health status.

表 2.7.3.7-8 本評価に用いた基準のオリジナルの基準からの改変箇所 (RECIST 1.1 Criteria)

ORIGINAL RECIST 1.1 CRITERIA	ADJUSTMENT CRITERIA	RATIONALE FOR ADJUSTMENT
Section 3.1.1, Measurable. Tumor lesions must be measured in at least one dimension (longest diameter in the plane of measurement to be recorded) with a minimum size of: • 10 mm by CT scan (CT scan slice thickness no greater than 5 mm) • 10 mm caliper measurement by clinical exam (lesions which cannot be accurately measured by calipers should be recorded as non-measurable) • 20 mm by chest X-ray Malignant lymph nodes: to be considered pathologically enlarged and measurable, a lymph node must be ≥ 15 mm in short axis when assessed by CT scan (CT scan slice thickness recommended to be no greater than 5 mm). At baseline and in follow-up, only the short axis will be measured and followed.	Lesions identified on CT and MRI scans will measure ≥ 10 mm (Reconstruction interval[RI] ≤ 5 mm). To be considered measurable, lymph nodes must measure ≥ 15 mm <i>and</i> 2x the RI in the short axis (when assessed by CTs). Measurement of lesions on chest x-ray will not be performed. Chest CTs will be required for lesion measurement. If applicable, nodal lesions will be measured and recorded bidimensionally, however, at Baseline/Screening and on-study, only the short axis will contribute to the SLD.	Evaluation of lesions on CT/MRI provides more accurate lesion measurement, extent of disease assessment and earlier/more sensitive detection of new lesions. It is CoreLab Partners' convention to require that lymph nodes measure ≥ 15 mm in the short axis and be 2x the RI at baseline in order to be considered a target lesion. Clinical target lesions will not be assessed for this study.
Section 3.1.2, Non-Measurable. All other lesions, including small lesions (LD < 10 mm or pathological lymph nodes ≥ 10 to < 15 mm short axis) as well as truly non-measurable lesions. Lesions considered truly non-measurable include: leptomeningeal disease, ascites, pleural or pericardial effusions, inflammatory breast disease, lymphangitic involvement of the skin or lung, abdominal masses or abdominal organomegaly identified by PE that is not measurable by reproducible imaging techniques.	The following will be considered non-measurable for this study: bone lesions not identified as target lesions lymphangitis of the skin or lung, pathological lymph nodes with short axis measuring ≥ 10 mm and < 15 mm, measurable lesions not chosen as a target lesion, groups of lesions that are small and numerous, and pleural/pericardial effusions and/or ascites.	Clarification regarding the types of lesions/non-measurable sites of disease which CoreLab Partners will consider as non-target.

表 2.7.3.7-8 本評価に用いた基準のオリジナルの基準からの改変箇所 (RECIST 1.1 Criteria) (続き)

ORIGINAL RECIST 1.1 CRITERIA	ADJUSTMENT CRITERIA	RATIONALE FOR ADJUSTMENT
Section 3.1.3, Special Considerations Regarding Lesion Measurability. Bone lesions, cystic lesions and lesions previously treated with local therapy require particular comment: <u>Bone Lesions:</u> • Bone scan, PET scan or plain films are not considered adequate imaging techniques to measure bone lesions. However, these techniques can be used to confirm the presence or disappearance of bone lesions. • Lytic bone lesions or mixed lytic-blastic lesions, with identifiable soft tissue components, that can be evaluated by CT or MRI can be considered as measurable lesions if the soft tissue component meets the definitions of measurable lesions. • Blastic lesions are non-measurable. <u>Cystic Lesions:</u> • Lesions that meet the criteria for radiographically defined simple cysts should not be considered as malignant lesions (neither measurable nor non-measurable) since they are, by definition, simple cysts. • Cystic lesions thought to represent cystic metastases can be considered as measurable lesions if they meet the definition for measurable lesions. However, if non-cystic lesions are present in the same patient, these are preferred for selection as target lesions. <u>Lesions with Prior Local Treatment:</u> Tumors situated in a previously irradiated area, or in an area subjected to other loco-regional therapy, are usually not considered measurable unless there has been demonstrated progression in the lesion. Study protocols should detail the conditions under which such lesions would be considered measurable.	<u>Bone Lesions:</u> For this study, a bone lesion with or without a soft tissue component may be identified as a target lesion. <u>Previously Irradiated Lesions:</u> Target lesions may be chosen from a previously irradiated area.	Standard CoreLab Partners conventions. <u>Bone Lesions:</u> For this study, a bone lesion with or without a soft tissue component may be identified as a target lesion. Bone windows and soft tissue windows maybe used to establish lesion dimensions <u>Previously Irradiated Lesions:</u> Target lesions may be chosen from a previously irradiated area.

表 2.7.3.7-8 本評価に用いた基準のオリジナルの基準からの改変箇所 (RECIST 1.1 Criteria) (続き)

ORIGINAL RECIST 1.1 CRITERIA	ADJUSTMENT CRITERIA	RATIONALE FOR ADJUSTMENT
Section 3.2.2, ¶ 2, Clinical Lesions: Clinical lesions will only be considered measurable when they are superficial and ≥ 10 mm diameter as assessed using calipers (e.g., skin nodules). For the case of skin lesions, documentation by colour photography including a ruler to estimate the size of the lesion is suggested. As noted above, when lesions can be evaluated by both clinical exam and imaging, imaging evaluation should be undertaken since it is more objective and may also be reviewed at the end of the study.	No clinical lesions will be measured for this study. No photographs will be provided for the independent review.	CoreLab Partners will not perform an oncology review for this study.
Section 3.2.2, ¶ 7, Tumor Markers: Tumor markers alone cannot be used to assess objective tumor response. If markers are initially above the upper normal limit, however, they must normalize for a patient to be considered in complete response. Because tumor markers are disease specific, instructions for their measurement should be incorporated into protocols on a disease specific basis. Specific guidelines for both CA-125 response (in recurrent ovarian cancer) and PSA response (in recurrent prostate cancer), have been published. In addition, the Gynecologic Cancer Intergroup has developed CA125 progression criteria which are to be integrated with objective tumor assessment for use in first-line trials in ovarian cancer.	No tumor markers will be provided for the Independent Review.	No tumor markers will be provided for the Independent Review.
Section 3.2.2, ¶ 8, Cytology, Histology: These techniques can be used to differentiate between PR and CR in rare cases if required by protocol (for example, residual lesions in tumor types such as germ cell tumors, where known residual benign tumors can remain). When effusions are known to be a potential adverse effect of treatment (e.g. with certain taxane compounds or angiogenesis inhibitors), the cytological confirmation of the neoplastic origin of any effusion that appears or worsens during treatment can be considered if the measurable tumor has met criteria for response or stable disease in order to differentiate between response (or stable disease) and progressive disease.	CoreLab Partners cannot call progression due to fluid collections without cytological confirmation of malignancy in the absence of a benign etiology.	Non-target progression is rare in a stable or responding subject.

表 2.7.3.7-8 本評価に用いた基準のオリジナルの基準からの改変箇所 (RECIST 1.1 Criteria) (続き)

ORIGINAL RECIST 1.1 CRITERIA	ADJUSTMENT CRITERIA	RATIONALE FOR ADJUSTMENT
Section 4.2, Baseline Documentation of “Target” and “Non-Target” Lesions, ¶ 3 (lymph nodes): Lymph nodes merit special mention since they are normal anatomical structures which may be visible by imaging even if not involved by tumor. As noted in Section 3, pathological nodes which are defined as measurable and may be identified as target lesions must meet the criterion of a short axis of ≥ 15 mm by CT scan. Only the short axis of these nodes will contribute to the baseline sum. The short axis of the node is the diameter normally used by radiologists to judge if a node is involved by solid tumor. Nodal size is normally reported as two dimensions in the plane in which the image is obtained (for CT scan this is almost always the axial plane; for MRI the plane of acquisition may be axial, sagittal or coronal). The smaller of these measures is the short axis. For example, an abdominal node which is reported as being 20 mm $\times$ 30 mm has a short axis of 20 mm and qualifies as a malignant, measurable node. In this example, 20 mm should be recorded as the node measurement. All other pathological nodes (those with short axis ≥ 10 mm but < 15 mm) should be considered non-target lesions. Nodes that have a short axis < 10 mm are considered non-pathological and should not be recorded or followed.	<u>Target Lesions:</u> Pathological lymph nodes may be identified as target lesions if they meet the criteria for measurable (≥ 15 mm in the short axis and 2x the Reconstruction interval[RI]) by CT scan. • Lymph nodes measuring ≥ 10 mm to < 15 mm should be considered non-target lesions. • The lymphatic system will be considered one (1) target organ. A maximum of two (2) lymph nodes may be chosen as target lesions. • Nodes that have a short axis < 10 mm are considered non-pathological and should not be recorded or followed. See notes on Section 4.3.2, Special Notes on the Assessment of Target Lesions, ¶ 1 (Lymph Nodes), in this appendix. <u>Non-Target Lesions:</u> • Pathological lymph nodes with a short axis measuring ≥ 10 mm and < 15 mm will be considered non-target lesions. • Additional criteria is provided for choosing groups of nodes that are small and numerous to qualify as non-target lesions.	Standard CoreLab Partners conventions.

表 2.7.3.7-8 本評価に用いた基準のオリジナルの基準からの改変箇所 (RECIST 1.1 Criteria) (続き)

ORIGINAL RECIST 1.1 CRITERIA	ADJUSTMENT CRITERIA	RATIONALE FOR ADJUSTMENT
Section 4.3.1, Evaluation of Target Lesions: This section provides the definitions of the criteria used to determine objective tumor response for target lesions: • CR: disappearance of all target lesions. Any pathological lymph nodes (whether target or non-target) must have reduction in short axis to < 10 mm. • PR: At least a 30% decrease in the SLD of target lesions, taking as reference the baseline SLD. • PD: at least a 20% increase in the SLD of target lesions, taking as reference the smallest sum on study (this includes the baseline SLD if that is the smallest on study). In addition to the relative increase of 20%, the SLD must also demonstrate an absolute increase of at least 5 mm. Note: the appearance of one or more new lesions is also considered progression. • SD: neither sufficient shrinkage to qualify for PR nor sufficient increase to qualify for PD, taking as reference the smallest SLD while on-study.	In addition, target lesions may also be assessed as UE or TSTM (Too Small To Measure). Assessment of the TPR will be made according to Table 1.	Provides additional clarification regarding how the target lesion may be assessed. Follows standard convention for response assessment of target lesions.
Section 4.3.2, Special Notes on the Assessment of Target Lesions, ¶ 1 (Lymph nodes): Lymph nodes identified as target lesions should always have the actual short axis measurement recorded (measured in the same anatomical plane as the baseline examination), even if the nodes regress to below 10mm on study. This means that when lymph nodes are included as target lesions, the ‘sum’ of lesions may not be zero even if complete response criteria are met, since a normal lymph node is defined as having a short axis of <10mm. Case report forms or other data collection methods may therefore be designed to have target nodal lesions recorded in a separate section where, in order to qualify for CR, each node must achieve a short axis <10mm. For PR, SD and PD, the actual short axis measurement of the nodes is to be included in the sum of target lesions.	When lymph nodes are included as target lesions, the SLD may never be zero (0) even if the CR criteria is met since normal lymph nodes are defined as having a short axis measurement < 10 mm. For lymph nodes, the actual short axis measurement should be recorded even if the nodes regress to < 10 mm on-study. In the event the short axis measurement decreases to < 5 mm, the lymph node will be categorized as TSTM and will follow standard TSTM conventions.	Standard CoreLab Partners convention.

表 2.7.3.7-8 本評価に用いた基準のオリジナルの基準からの改変箇所 (RECIST 1.1 Criteria) (続き)

ORIGINAL RECIST 1.1 CRITERIA	ADJUSTMENT CRITERIA	RATIONALE FOR ADJUSTMENT
Section 4.3.2, Special Notes on the Assessment of Target Lesions, ¶ 2 (TSTM): While on study, all lesions (nodal and non-nodal) recorded at baseline should have their actual measurements recorded at each subsequent evaluation, even when very small (e.g., 2 mm). However, sometimes lesions or lymph nodes which are recorded as target lesions at baseline become so faint on CT scan that the radiologist may not feel comfortable assigning an exact measure and may report them as being TSTM. When this occurs it is important that a value be recorded on the case report form. If it is the opinion of the radiologist that the lesion has likely disappeared, the measurement should be recorded as 0 mm. If the lesion is believed to be present and is faintly seen but too small to measure, a default value of 5 mm should be assigned (Note: It is less likely that this rule will be used for lymph nodes since they usually have a definable size when normal and are frequently surrounded by fat such as in the retroperitoneum; however, if a lymph node is believed to be present and is faintly seen but too small to measure, a default value of 5 mm should be assigned in this circumstance as well). This default value is derived from the 5 mm CT slice thickness (but should not be changed with varying CT slice thickness). The measurement of these lesions is potentially non-reproducible, therefore providing this default value will prevent false responses or progressions based upon measurement error. To reiterate, however, if the radiologist is able to provide an actual measure, that should be recorded, even if it is below 5 mm.	Target lesions that are being followed, which after baseline/screening decreases in size to less than five (5) mm in any dimension, will be categorized as TSTM. A lesion will be assigned a value of five (5) mm for the calculation of the LD. If a lesion subsequently increases in size to five (5) mm or greater, its true size will be recorded. See Section 4.3.2 above, regarding special considerations for lymph nodes that regress to < 10 mm.	Standard CoreLab Partners conventions

表 2.7.3.7-8 本評価に用いた基準のオリジナルの基準からの改変箇所 (RECIST 1.1 Criteria) (続き)

ORIGINAL RECIST 1.1 CRITERIA	ADJUSTMENT CRITERIA	RATIONALE FOR ADJUSTMENT
Section 4.3.2, Special Notes on the Assessment of Target Lesions, ¶ 1 (Lesions that Split or Coalesce on Treatment): When non-nodal lesions ‘fragment’, the longest diameters of the fragmented portions should be added together to calculate the target lesion sum. Similarly, as lesions coalesce, a plane between them may be maintained that would aid in obtaining maximal diameter measurements of each individual lesion. If the lesions have truly coalesced such that they are no longer separable, the vector of the longest diameter in this instance should be the maximal longest diameter for the ‘coalesced lesion’	<u>Split Lesions:</u> in certain scenarios, measuring the long axes of the lesions that split could lead to a % change increase, when in fact, there is a decrease in size. Measuring split lesions along the same vector as the original lesion would provide a more accurate assessment of lesion sizes. <u>Conglomerate Nodal Masses:</u> If a conglomerate nodal mass chosen as a target lesion splits into normal sized nodes on-study, only the longest short axis measurement of all the individual nodes will contribute to the SLD. For example, if a conglomerate nodal mass comprised of four (4) lymph nodes was chosen as a target lesion, and subsequently, the conglomerate nodal mass divided into four (4) distinct normal-sized lymph nodes, only the longest short axis of the four (4) nodes will contribute to the SLD. In this manner, an assessment of CR is still possible.	Standard CoreLab Partners conventions.
Section 4.3.3, Evaluation of Non-Target Lesions: This section provides definitions of the criteria used to determine the tumor response for the group of non-target lesions. While some non-target lesions may actually be measurable, they need not be measured and instead, should be assessed only qualitatively at the time points specified in the protocol. • CR: Disappearance of all non-target lesions and normalization of tumor marker level. All lymph nodes must be non-pathological in size (< 10 mm short axis). • SD (Non-CR/Non-PD): Persistence of one or more non-target lesions and/or maintenance of tumor marker level above the normal limits. • PD: unequivocal progression of existing non-target lesions. Note: the appearance of one or more new lesions is also considered progression.	For the purposes of internal consistency, non-target lesion assessments of “Non-CR/Non-PD” will be as mapped as “SD”. In addition, non-target lesions may also be assessed as UE or NA (no non-target lesions identified at Baseline/Screening).	Internal CoreLab Partners conventions (for SD assessment). UE and NA assessments provide additional clarification regarding how non-target lesions may be assessed.

表 2.7.3.7-8 本評価に用いた基準のオリジナルの基準からの改変箇所 (RECIST 1.1 Criteria) (続き)

ORIGINAL RECIST 1.1 CRITERIA	ADJUSTMENT CRITERIA	RATIONALE FOR ADJUSTMENT
Section 4.3.4, Special Notes on the Assessment of Non-Target Disease: The concept of progression of non-target disease requires additional explanation as follows: ¶ 2 (When the patient also has measurable disease): In this setting, to achieve ‘unequivocal progression’ on the basis of the non-target disease, there must be an overall level of substantial worsening in non-target disease such that, even in presence of SD or PR in target disease, the overall tumor burden has increased sufficiently to merit discontinuation of therapy (see examples in Appendix II and further details below). A modest ‘increase’ in the size of one or more non-target lesions is usually not sufficient to qualify for unequivocal progression status. The designation of overall progression solely on the basis of change in non-target disease in the face of SD or PR of target disease will therefore be extremely rare.	The TPR will be assessed according to Table 1 (see Section 4.4.1, TPR, below).	N/A
Section 4.3.4, Special Notes on the Assessment of Non-Target Disease, ¶ 3 (When the patient has only non-measurable disease): This circumstance arises in some phase III trials when it is not a criterion of study entry to have measurable disease. The same general concepts apply here as noted above, however, in this instance there is no measurable disease assessment to factor into the interpretation of an increase in non-measurable disease burden. Because worsening in non-target disease cannot be easily quantified (by definition: if all lesions are truly non-measurable) a useful test that can be applied when assessing patients for unequivocal progression is to consider if the increase in overall disease burden based on the change in non-measurable disease is comparable in magnitude to the increase that would be required to declare PD for measurable disease: i.e., an increase in tumor burden representing an additional 73% increase in ‘volume’ (which is equivalent to a 20% increase diameter in a measurable lesion). Examples include an increase in a pleural effusion from ‘trace’ to ‘large’, an increase in lymphangitic disease from localized to widespread, or may be described in protocols as ‘sufficient to require a change in therapy’. While it would be ideal to have objective criteria to apply to non-measurable disease, the very nature of that disease makes it impossible to do so, therefore the increase must be substantial.	The TPR will be assessed according to Table 1 (see Section 4.4.1, TPR, below).	N/A

表 2.7.3.7-8 本評価に用いた基準のオリジナルの基準からの改変箇所 (RECIST 1.1 Criteria) (続き)

ORIGINAL RECIST 1.1 CRITERIA	ADJUSTMENT CRITERIA	RATIONALE FOR ADJUSTMENT
Section 4.3.5, New Lesions, ¶ 2 (Lesion Identified on a Follow-Up Study in an Anatomical Location Not Scanned at Baseline): A lesion identified on a follow-up study in an anatomical location that was not scanned at baseline is considered a new lesion and will indicate disease progression. An example of this is the patient who has visceral disease at baseline and while on study has a CT or MRI brain ordered which reveals metastases. The patient's brain metastases are considered to be evidence of PD even if he/she did not have brain imaging at baseline.	In cases where lesions are seen at a subsequent time point in an area that was not scanned at baseline/screening, the rules for progression will apply.	Standard CoreLab Partners convention.
Section 4.3.5, New Lesions, ¶ 4 (Use of FDG-PET): While FDG-PET response assessments need additional study, it is sometimes reasonable to incorporate the use of FDG-PET scanning to complement CT scanning in assessment of progression (particularly possible 'new' disease). New lesions on the basis of FDG-PET imaging can be identified according to the following algorithm: • Negative FDG-PET at baseline, with a positive FDG-PET at follow-up is a sign of PD based on a new lesion. • No FDG-PET at baseline and a positive FDG-PET at follow-up: ○ If the positive FDG-PET at follow-up corresponds to a new site of disease confirmed by CT, this is PD. ○ If the positive FDG-PET at follow-up is not confirmed as a new site of disease on CT, additional follow-up CT scans are needed to determine if there is truly progression occurring at that site (if so, the date of PD will be the date of the initial abnormal FDG-PET scan). ○ If the positive FDG-PET at follow-up corresponds to a pre-existing site of disease on CT that is not progressing on the basis of the anatomic images, this is not PD. Note: a positive FDG-PET scan lesion means one which is FDG-avid with an uptake greater than twice that of the surrounding tissue on the attenuation corrected image.	For the RECIST review, FDG-PET will be considered off-protocol imaging and will not be used for any assessments, including those for assessing new lesions.	FDG-PET will be evaluated by EORTC criteria.

表 2.7.3.7-8 本評価に用いた基準のオリジナルの基準からの改変箇所 (RECIST 1.1 Criteria) (続き)

ORIGINAL RECIST 1.1 CRITERIA	ADJUSTMENT CRITERIA	RATIONALE FOR ADJUSTMENT
Section 4.4.1, TPR: It is assumed that at each protocol specified time point, a response assessment occurs. Table 1 provides a summary of the overall response status calculation at each time point for patients who have measurable disease at Baseline. When patients have non-measurable (there non-target) disease only, Table 2 is to be used.	CoreLab Partners will use the Time Point Response table in the Charter (see Table 1)	Standard CoreLab Partners convention.
Section 4.4.2, Missing Assessments and Inevaluable Designations: When no imaging/measurement is done at all at a particular time point, the patient is not evaluable (NE) at that time point. If only a subset of lesion measurements are made at an assessment, usually the case is also considered NE at that time point, unless a convincing argument can be made that the contribution of the individual missing lesion(s) would not change the assigned time point response. This would be most likely to happen in the case of PD. For example, if a patient had a baseline sum of 50 mm with three measured lesions and at follow-up only two lesions were assessed, but those gave a sum of 80 mm, the patient will have achieved PD status, regardless of the contribution of the missing lesion.	For the term “Inevaluable”, CoreLab Partners will continue to use the term “Unevaluable”, which is designated as “UE”. Standard UE scenarios for target lesions and non-target lesions will be used for TPR assessments. In the event CoreLab Partners does not receive a particular imaging set which precludes independent reviewers from making a measurement, CoreLab Partners will raise an issue for the scans not received and will only read once Amgen or Daiichi Sankyo provides an acknowledgement to proceed with the read (or the issue is deemed irresolvable).	Standard CoreLab Partners convention.
Section 4.4.3, Best Overall Response (all Time Points), ¶ 2 (Confirmation of CR/PR is Required): CRs or PRs may be claimed only if the criteria for each are met at a subsequent time point as specified in the protocol (generally 4 weeks later). In this circumstance, the best overall response can be interpreted as in Table 3.	Confirmation of response will not be performed by CoreLab Partners.	Duration of response will be evaluated according to Statistical Analysis Plan.
Section 4.4.4, Special Notes on Response Assessment, ¶ 1 (When Nodal Disease is Included in the SLD): When nodal disease is included in the sum of target lesions and the nodes decrease to ‘normal’ size (< 10 mm), they may still have a measurement reported on scans. This measurement should be recorded even though the nodes are normal in order not to overstate progression should it be based on increase in size of the nodes. As noted earlier, this means that patients with CR may not have a total sum of ‘zero’ on the case report form (CRF).	When lymph nodes are included as target lesions, the SLD may never be zero (0) even if the CR criteria is met since normal lymph nodes are defined as having a short axis measurement < 10 mm. For TSTM in lymph nodes: the actual short axis measurement should be recorded even if the lymph nodes regress to < 10 mm on-study (see Section 4.3.2 in this Appendix). In the event the short axis measurement decreases to < 5 mm, the lymph node will be categorized as TSTM and will follow standard TSTM conventions.	Standard CoreLab Partners convention.

表 2.7.3.7-8 本評価に用いた基準のオリジナルの基準からの改変箇所 (RECIST 1.1 Criteria) (続き)

ORIGINAL RECIST 1.1 CRITERIA	ADJUSTMENT CRITERIA	RATIONALE FOR ADJUSTMENT
Section 4.4.4, Special Notes on Response Assessment, ¶ 2 (Repeated “NE” Time Point Assessments): In trials where confirmation of response is required, repeated ‘NE’ time point assessments may complicate best response determination. The analysis plan for the trial must address how missing data/assessments will be addressed in determination of response and progression. For example, in most trials it is reasonable to consider a patient with time point responses of PR-NE-PR as a confirmed response.	N/A: Confirmation of response will not be performed by CoreLab Partners.	The response will be evaluated according to Statistical Analysis Plan.
Section 4.4.4, Special Notes on Response Assessment, ¶ 3 (Global Deterioration): Patients with a global deterioration of health status requiring discontinuation of treatment without objective evidence of disease progression at that time should be reported as ‘symptomatic deterioration’. Every effort should be made to document objective progression even after discontinuation of treatment. Symptomatic deterioration is not a descriptor of an objective response: it is a reason for stopping study therapy. The objective response status of such patients is to be determined by evaluation of target and non-target disease as shown in Tables 1–3.	On-study clinical will not be provided for the independent review.	CoreLab Partners will not be performing an oncology review for this study.
Section 4.4.4, Special Notes on Response Assessment, ¶ 4 (Early Progression, Early Death): Conditions that define “early progression, early death and inevaluability” are study specific and should be clearly described in each protocol (depending on treatment duration, treatment periodicity).	For early death and early progression, there is no adjustment. For UE, for examples of when UE may be used and how UEs will affect target and non-target lesion responses.	Standard CoreLab Partners convention.
Section 4.4.4, Special Notes on Response Assessment, ¶ 5 (Biopsies and the Use of FDG-PET): In some circumstances it may be difficult to distinguish residual disease from normal tissue. When the evaluation of complete response depends upon this determination, it is recommended that the residual lesion be investigated (fine needle aspirate/biopsy) before assigning a status of complete response. FDG-PET may be used to upgrade a response to a CR in a manner similar to a biopsy in cases where a residual radiographic abnormality is thought to represent fibrosis or scarring. The use of FDG-PET in this circumstance should be prospectively described in the protocol and supported by disease specific medical literature for the indication. However, it must be acknowledged that both approaches may lead to false positive CR due to limitations of FDG-PET and biopsy resolution/ sensitivity.	On-study clinical data will not contribute to the imaging assessments. For the RECIST evaluation, FDG-PET will be considered off-protocol imaging and will not be used for the assessment of response or progression.	CoreLab Partners will not be performing an oncology review for this study. ¹⁸ FDG-PET will be evaluated by EORTC criteria.

表 2.7.3.7-8 本評価に用いた基準のオリジナルの基準からの改変箇所 (RECIST 1.1 Criteria) (続き)

ORIGINAL RECIST 1.1 CRITERIA	ADJUSTMENT CRITERIA	RATIONALE FOR ADJUSTMENT
Section 4.4.4, Special Notes on Response Assessment, ¶ 6 (Equivocal PD Findings): For equivocal findings of progression (e.g., very small and uncertain new lesions; cystic changes or necrosis in existing lesions), treatment may continue until the next scheduled assessment. If at the next scheduled assessment, progression is confirmed, the date of progression should be the earlier date when progression was suspected.	N/A	A single Date of Progression for each subject will not be transferred to Amgen or Daiichi Sankyo.
Section 4.5, Frequency of Tumor Re-Evaluation: Frequency of tumor re-evaluation while on treatment should be protocol specific and adapted to the type and schedule of treatment. However, in the context of phase II studies where the beneficial effect of therapy is not known, follow-up every 6–8 weeks (timed to coincide with the end of a cycle) is reasonable. Smaller or greater time intervals than these could be justified in specific regimens or circumstances. The protocol should specify which organ sites are to be evaluated at baseline (usually those most likely to be involved with metastatic disease for the tumor type under study) and how often evaluations are repeated. Normally, all target and non-target sites are evaluated at each assessment. In selected circumstances certain non-target organs may be evaluated less frequently. For example, bone scans may need to be repeated only when complete response is identified in target disease or when progression in bone is suspected. After the end of the treatment, the need for repetitive tumor evaluations depends on whether the trial has as a goal the response rate or the time to an event (progression/death). If ‘time to an event’ (e.g., time to progression, disease-free survival, progression-free survival) is the main endpoint of the study, then routine scheduled re-evaluation of protocol specified sites of disease is warranted. In randomized comparative trials in particular, the scheduled assessments should be performed as identified on a calendar schedule (for example: every 6–8 weeks on treatment or every 3–4 months after treatment) and should not be affected by delays in therapy, drug holidays or any other events that might lead to imbalance in a treatment arm in the timing of disease assessment.	For this study, regularly scheduled tumor re-evaluations will not be performed. CoreLab Partners will evaluate all target and non-target lesions on the available imaging exams. The primary endpoint for this study is objective tumor response.	The response will be evaluated according to Statistical Analysis Plan.

表 2.7.3.7-8 本評価に用いた基準のオリジナルの基準からの改変箇所 (RECIST 1.1 Criteria) (続き)

ORIGINAL RECIST 1.1 CRITERIA	ADJUSTMENT CRITERIA	RATIONALE FOR ADJUSTMENT
Section 4.6.1, Confirmation: In non-randomized trials where response is the primary endpoint, confirmation of PR and CR is required to ensure responses identified are not the result of measurement error. This will also permit appropriate interpretation of results in the context of historical data where response has traditionally required confirmation in such trials. However, in all other circumstances i.e., in randomized trials (phase II or III) or studies where stable disease or progression are the primary endpoints, confirmation of response is not required since it will not add value to the interpretation of trial results. However, elimination of the requirement for response confirmation may increase the importance of central review to protect against bias, in particular in studies which are not blinded. In the case of SD, measurements must have met the SD criteria at least once after study entry at a minimum interval (in general not less than 6–8 weeks) that is defined in the study protocol.	N/A: Confirmation of response will not be performed by CoreLab Partners.	The response will be evaluated according to Statistical Analysis Plan.
Section 4.9.1, Reporting Best Response Results (Phase II Trials): When response is the primary endpoint, and thus all patients must have measurable disease to enter the trial, all patients included in the study must be accounted for in the report of the results, even if there are major protocol treatment deviations or if they are not evaluable. Each patient will be assigned one of the following categories: • CR • PR • SD • PD • Inevaluable for response: specify reasons (for example, early death, malignant disease, early death, toxicity, tumor assessments not repeated/incomplete, other). Normally, all eligible patients should be included in the denominator for the calculation of the response rate for phase II trials (in some protocols it will be appropriate to include all treated patients). It is generally preferred that 95% two-sided confidence limits are given for the calculated response rate. Trial conclusions should be based on the response rate for all eligible (or all treated) patients and should not be based on a selected 'evaluable' subset.	Subjects will be assigned one of the following categories: • CR • PR • SD • PD • UE CoreLab Partners assessments will not include information about death.	This is a radiology review. CoreLab Partners will not be provided information about date of death from Amgen and Daiichi Sankyo. Amgen or Daiichi Sankyo, not CoreLab Partners, will report response rate and perform any required statistical analyses.

表 2.7.3.7-8 本評価に用いた基準のオリジナルの基準からの改変箇所 (RECIST 1.1 Criteria) (続き)

ORIGINAL RECIST 1.1 CRITERIA	ADJUSTMENT CRITERIA	RATIONALE FOR ADJUSTMENT
Appendix II, Specifications for Standard Anatomical Radiological Imaging, ¶ 2: Scanner quality control is highly recommended and should follow standard manufacturer and facility maintenance schedules using commercial phantoms. It is likely that for RECIST unidimensional measurements this will be adequate to produce reproducible measurements. Imaging quality control for CT includes an analysis of image noise and uniformity and CT number as well as spatial resolution. The frequency of the quality control analysis is also variable and should focus on the clinically relevant scanning parameters. Dose analysis is also important and the use of imaging should follow the ALARA principle, ‘As Low As Reasonably Achievable’, which refers to making every reasonable effort to maintain radiation exposures as far below the dose limits as possible.	For Studies 20040215 and 20062004, no site qualification will be performed.	This is fully retrospective imaging review.
	For Study AMG162-B-J201, to ensure that CoreLab Partners’ minimum quality requirements are met, all imaging facilities used in this clinical trial must undergo site qualification prior to CoreLab Partners’ receipt of the subject’s first on-study images. Imaging facilities (including satellite facilities) are required to complete an Imaging Capabilities Questionnaire (ICQ), which will provide CoreLab Partners with information on the equipment used (including manufacturer, model number, software version, and in certain situations, how often maintenance is performed), media type(s), and technical capabilities used for image acquisition. It is expected that qualification standards will be maintained by the imaging facilities throughout the course of the study. Any changes to the information provided on the ICQ will require a re-submission of the form with updated information.	Standard CoreLab Partners conventions.
Appendix II, Specific notes, ¶ 3 (Chest X-Ray): Chest X-ray measurement of lesions surrounded by pulmonary parenchyma is feasible, but not preferable as the measurement represents a summation of densities. Furthermore, there is poor identification of new lesions within the chest on X-ray as compared with CT. Therefore, measurements of pulmonary parenchymal lesions as well as mediastinal disease are optimally performed with CT of the chest. MRI of the chest should only be performed in extenuating circumstances. Even if IV contrast cannot be administered (for example, in the situation of allergy to contrast), a non-contrast CT of the chest is still preferred over MRI or chest x-ray.	Measurements of lesions will not be allowed on chest x-rays; CT will be required.	Measurement of lesion size and determination of the extent of disease is more accurate on CT as compared with chest x-rays.

表 2.7.3.7-8 本評価に用いた基準のオリジナルの基準からの改変箇所 (RECIST 1.1 Criteria) (続き)

ORIGINAL RECIST 1.1 CRITERIA	ADJUSTMENT CRITERIA	RATIONALE FOR ADJUSTMENT
Appendix II, Specific notes, ¶ 5 (Image Acquisition Parameters, Subsection D, Alternative Contrast Agents): Alternative contrast agents: There are a number of other, new contrast agents, some organ specific. They may be used as part of patient care for instance, in liver lesion assessment, or lymph node characterization, but should not as yet be used in clinical trials.	N/A (for Studies 20040215 and 20062004)	This is a fully retrospective imaging review.
	For Study AMG162-B-J201, subjects being imaged with contrast-enhanced CT may, if they have developed a medical contraindication to IV contrast at a time during the study, change to contrast-enhanced MRI. At the time of the reading, the radiologist will comment if the imaging techniques are comparable and if there has been any effect on response assessment. If in the medical opinion of the radiologist, there is too much difference in the exams, the response will not be assessable.	While on study, there may be some subjects who develop a medical contraindication to IV contrast for CT. If IV contrast is contraindicated, the chest evaluation should be performed with non-contrast CT. Abdominal and pelvic evaluations should be performed using MRI.
Appendix II, Specific notes, ¶ 7 (PET/CT Scans): Combined modality scanning such as with PET-CT is increasingly used in clinical care, and is a modality/technology that is in rapid evolution; therefore, the recommendations in this paper may change rather quickly with time. At present, low dose or attenuation correction CT portions of a combined PET-CT are of limited use in anatomically based efficacy assessments and it is therefore suggested that they should not be substituted for dedicated diagnostic contrast enhanced CT scans for anatomically based RECIST measurements. However, if a site can document that the CT performed as part of a PET-CT is of identical diagnostic quality to a diagnostic CT (with IV and oral contrast) then the CT portion of the PET-CT can be used for RECIST measurements. Note, however, that the PET portion of the CT introduces additional data which may bias an investigator if it is not routinely or serially performed.	CoreLab Partners' standard convention is to allow the use of the CT portion of a PET/CT in lieu of a dedicated CT provided imaging requirements are met. These parameters distinguish an optimized CT exam from an exam which serves primarily as the attenuation correction map for the PET exam.	Standard CoreLab Partners convention.
Appendix II, Specific notes, ¶ 9 (Physical Examination): While evaluation of lesions by physical examination is also of limited reproducibility, it is permitted when lesions are superficial, at least 10 mm size, and can be assessed using calipers. In general, it is preferred if patients on clinical trials have at least one lesion that is measurable by CT. Other skin or palpable lesions may be measured on physical examination and be considered target lesions.	No clinical lesions (target or non-target) will be identified or followed for this study.	CoreLab Partners will not be performing an oncology review for this study.

表 2.7.3.7-8 本評価に用いた基準のオリジナルの基準からの改変箇所 (RECIST 1.1 Criteria) (続き)

ORIGINAL RECIST 1.1 CRITERIA	ADJUSTMENT CRITERIA	RATIONALE FOR ADJUSTMENT
Appendix II, Specific notes, ¶ 10 (Use of MRI): Use of MRI remains a complex issue. MRI has excellent contrast, spatial and temporal resolution; however, there are many image acquisition variables involved in MRI, which greatly impact image quality, lesion conspicuity and measurement. Furthermore, the availability of MRI is variable globally. As with CT, if an MRI is performed, the technical specifications of the scanning sequences used should be optimized for the evaluation of the type and site of disease. Furthermore, as with CT, the modality used at follow-up should be the same as was used at baseline and the lesions should be measured/assessed on the same pulse sequence. Generally, axial imaging of the abdomen and pelvis with T1 and T2 weighted imaging along with gadolinium enhanced imaging should be performed. The field of view, matrix, number of excitations, phase encode steps, use of fat suppression and fast sequences should be optimized for the specific body part being imaged as well as the scanner utilized. It is beyond the scope of this document or appendix to prescribe specific MRI pulse sequence parameters for all scanners, body parts and diseases. Ideally, the same type of scanner should be used and the image acquisition protocol should be followed as closely as possible to prior scans. Body scans should be performed with breath-hold scanning techniques if possible.	The same method of assessment and the same technique should be used to characterize each identified and reported lesion at baseline/screening and during follow-up. Whenever possible, the same equipment, field strength, sequences, scanning parameters, positioning, angulation, timing, FOV and slice thickness should have been utilized for all MRI exams acquired both pre- and post-contrast for a given subject over the course of the study. Lesions identified on MRI should measure ≥ 10 mm in LD and twice the slice thickness.	Standard CoreLab Partners convention.
Appendix II, Specific notes, ¶ 11 (Selection of Target Lesions): In general, the largest lesions representative of involved organs (up to a maximum of two per organ and five total) are selected to follow as target lesions. However, in some cases, the largest lesions may not be easily measured and are not suitable for follow-up because of their configuration. In these cases, identification of the largest most reproducible lesions is advised.	Up to five (5) radiographic target lesions (two [2] per target organ) will be identified at Baseline/Screening. All target lesions must meet measurable disease criteria. These lesions will generally be the largest tumor masses. They will lend themselves to precise anatomical re-location, tend to be the most reliably measured, and are most representative of the subject's disease.	Standard CoreLab Partners convention.

表 2.7.3.7-8 本評価に用いた基準のオリジナルの基準からの改変箇所 (RECIST 1.1 Criteria) (続き)

ORIGINAL RECIST 1.1 CRITERIA	ADJUSTMENT CRITERIA	RATIONALE FOR ADJUSTMENT
Appendix II, Measurement of Lesions, ¶ 12 (LDs): The longest diameter of selected lesions should be measured in the plane in which the images were acquired. For body CT, this is the axial plane. In the event isotropic reconstructions are performed, measurements can be made on these reconstructed images; however, it should be cautioned that not all radiology sites are capable of producing isotropic reconstructions. This could lead to an undesirable situation of measurements in the axial plane at one assessment point and in a different plane at a subsequent assessment. There are some tumors, for instance paraspinal lesions, which are better measured in the coronal or sagittal plane. It would be acceptable to measure these lesions in these planes if the reconstructions in those planes were isotropic or the images were acquired with MRI in those planes. Using the same plane of evaluation, the maximal diameter of each target lesion should always be measured at subsequent follow-up time points even if this results in measuring the lesion at a different slice level or in a different orientation or vector compared to the baseline study. Software tools that calculate the maximal diameter for a perimeter of a tumor may be employed and may even reduce variability.	The LDs of all non-nodal target lesions will be recorded (in mm) on the Source Documents. LDs (and the short axes of any nodal target lesions, see next section below) will be summed to obtain the SLD at each time point.	Standard CoreLab Partners convention.
Appendix II, Measurement of Lesions, ¶ 13 (Lymph Node Measurement): The only exception to the longest diameter rule is lymph node measurement. Because malignant nodes are identified by the length of their short axis, this is the guide used to determine not only whether they are pathological but is also the dimension measured for adding into the sum of target lesions.	Lymph nodes will be measured bi-dimensionally, but only the short axis will be recorded. Only the short axis measurement of lymph nodes will be added to the SLD.	Standard CoreLab Partners convention.

表 2.7.3.7-9 本評価に用いた基準のオリジナルの基準からの改変箇所 (EORTC Criteria)

ORIGINAL EORTC CRITERIA	MODIFICATION OF THE CRITERIA	RATIONALE FOR THE MODIFICATIONS
Complete response (CR) would be complete resolution of [^{18}F]-FDG uptake within the tumour volume so that it was indistinguishable from surrounding normal tissue.	Complete resolution of abnormal FDG-uptake within the tumor volume of the target lesion to a level which is indistinguishable from surrounding normal tissue.	Language change to be consistent with CoreLab Partners conventions.
Partial response (PR) would be classified as a reduction of a minimum of $15 \pm 25\%$ in tumour [^{18}F]-FDG SUV after one cycle of chemotherapy, and greater than 25% after more than one treatment cycle. Reporting would need to be accompanied by adequate and disclosed reproducibility measurements from each centre. An empirical 25% was found to be a useful cut-off point, but there is a need for a reproducibility analysis to determine the appropriate cut-offs for statistical significance. A reduction in the extent of the tumour FDG uptake is not a requirement for partial metabolic response.	$\% \Delta \Sigma \text{SUV}_{\text{max}}$ decrease of $\geq 25\%$ compared to baseline/screening	Additional detail added in for assessing the response of individual or grouped target lesions.
Stable disease (SD) would be classified as an increase in tumour [^{18}F]-FDG SUV of less than 25% or a decrease of less than 15% and no visible increase in extent of FDG tumour uptake (20% in the longest dimension).	$\% \Delta \Sigma \text{SUV}_{\text{max}}$ increase by $< 25\%$ or decrease by $< 25\%$ compared to baseline/screening.	Additional detail added in for assessing the response of individual or grouped target lesions.
Progressive disease (PD) to be classified as an increase in FDG tumour SUV of greater than 25% within the tumour region defined on the screening scan, visible increase in the extent of FDG tumour uptake (20% in the longest dimension) or the appearance of new FDG uptake in metastatic lesions.	$\% \Delta \Sigma \text{SUV}_{\text{max}}$ increase by $\geq 25\%$ compared to baseline/screening, or the presence of one or more new FDG-positive lesions	Additional detail added in that tumor response (CR, PR or SD) and progression will be compared to baseline/screening.

表 2.7.3.7-10 本評価に用いた基準のオリジナルの基準からの改変箇所 (Choi Criteria)

ORIGINAL Choi CRITERIA	MODIFICATION OF THE CRITERIA	RATIONALE FOR THE MODIFICATIONS
Complete response (CR) Disappearance of all lesions No new lesions	No Modification	N/A
Partial response (PR) A decrease in size of 10% or a decrease in tumor density (HU) 15% on CT No new lesions No obvious progression of nonmeasurable disease	A decrease in the Choi SLD $\geq 10\%$ or, an increase in CT density ($\%\Delta$ Hounsfield Unit [HU] mean) $\geq 15\%$ compared to Baseline/Screening.	Because denosumab inhibits osteoclastic activity, denosumab treatment is expected to result in ossification and calcification of the GCTB lesion (ie, replacement of tumor tissue with trabecular woven calcified bone). Therefore, for this assessment, the Choi criteria designed non calcified GIST were modified to define response based on an increase in lesion density (as measured by a percent change in Hounsfield Units).
Stable disease (SD) Does not meet the criteria for CR, PR, or PD No symptomatic deterioration attributed to tumor progression	Does not meet the criteria for CR, PR or PD.	This is a radiology review. CoreLab Partners will not be provided information about symptomatic deterioration attributed to tumor progression
Progressive disease (PD) An increase in tumor size of 10% and does not meet criteria of PR by tumor density (HU) on CT New lesions New intratumoral nodules or increase in the size of the existing intratumoral nodules	An increase in unidimensional tumor size (Choi SLD) of $> 10\%$ and does not meet the criteria for PR using CT density. Note: The identification of any new lesion(s) identified on CT/MRI will result in a determination of PD.	Intratumoral nodules are not expressed in GCTB.