

Assessing the value of ¹⁸FDG-PET in lung cancer: from theory to practice

CT	computertomografie
CWU	conventional workup
D-RCT	diagnostic randomised controlled trial
FDG	fluorodeoxyglucose
IKA	Integraal Kankercentrum Amsterdam
MRI	magnetic resonance imaging
NSCLC	non-small cell lung cancer
PET	positronemissietomografie
PLUS	PET in LUnG cancer Staging
RCT	randomised controlled trial

Positronemissietomografie (PET) is een betrekkelijk nieuwe nucleair-geneeskundige beeldvormende techniek die het mogelijk maakt biochemische processen in weefsels zichtbaar te maken. Om die reden kan PET een complementaire rol spelen bij traditionele beeldvormende technieken zoals CT en MRI, die anatomische informatie opleveren.

Toen het VU medisch centrum in 1996 overging tot de aanschaf van een PET-scanner, werd besloten de klinisch-wetenschappelijke aandacht te richten op het identificeren van kosteneffectieve toepassingen in de klinische praktijk. Dit proefschrift biedt de theoretische ondersteuning voor dit proces en beschrijft de toepassing bij niet-klein-cellige longkanker.

In *hoofdstuk 1* wordt een inleiding gegeven over de complexiteit van diagnosticeren en stadiëren in het algemeen en bij niet-klein-cellige longkanker (NSCLC) in het bijzonder. In een recent cohortonderzoek, uitgevoerd in twee ziekenhuizen in de relevante geografische regio, werden tekortkomingen gevonden in de routinematige diagnostiek bij longkanker.¹ Bij bijna 50% van de geopereerde patiënten bleek de chirurgische interventie niet geslaagd vanwege niet verwijderbaar tumorweefsel, goedaardige aandoeningen of omdat de ziekte binnen een jaar terug kwam of de patiënt overleed in die periode. Tegelijkertijd lieten PET-studies ¹⁸F-fluorodexoyglucose als 'tracer' (FDG-PET) veelbelovende resultaten zien wat de accuratesse voor verschillende aspecten van NSCLC-stadiëring betreft.² Echter, betere accuratesse betekent niet automatisch dat het klinisch bruikbaar is, dat wil zeggen dat het leidt tot betere diagnostiek en behandeling van patiënten en uiteindelijk tot betere uitkomsten.³

Het evalueren van de waarde van een diagnostische test kan over het algemeen het best geschieden volgens een gefaseerd hiërarchisch proces, zoals in de literatuur beschreven door Fryback en anderen.⁴ Een dergelijk proces omvat een vijftal niveaus: technische aspecten, zoals kwaliteit van de beelden en reproduceerbaarheid, meerwaarde op het diagnostisch inzicht, invloed op behandelingsbeslissingen, uitkomsten voor patiënten, en als hoogste niveau geldt een kosten-batenanalyse bij introductie van de nieuwe test. Het aantonen van effectiviteit op een bepaald niveau is een voorwaarde, maar is op zichzelf geen garantie voor effectiviteit op een hoger niveau. Gegeven de veelbelovende resultaten van de accuratessestudies en geleid door het resultaat van een modelmatige studie naar de kosteneffectieve toepassing van de nieuwe test,⁵ waren wij van mening dat de toegevoegde waarde van FDG-PET het best kon worden bestudeerd door gelijktijdig het huidige diagnostische beleid te vergelijken met hetzelfde beleid, maar waar dan FDG-PET aan toegevoegd zou zijn. Om mogelijke verstoringen door verschillen in patiënt- of tumorkarakteristieken zo klein mogelijk te maken, leek een gerandomiseerde studie (RCT) de meest geschikte studieopzet.

Om het begrip ten aanzien van de potentiële invloed van PET op diagnostisch handelen en op behandelingsbeleid in de preoperatieve situatie van NSCLC te vergroten, hebben we eerst een 'clinical value'- of 'voor-na'-studie uitgevoerd (*hoofdstuk 2*).⁶ Met behulp van vragenlijsten werden artsen – gaande het diagnostisch proces – verzocht aan te geven wat hun diagnostische begrip was en welke behandelingsstrategie ze zouden volgen net voor, meteen na en drie maanden na het verkrijgen van de resultaten van de PET-scan. Patiënten, verdacht van NSCLC, werden naar het PET-centrum verwezen indien niet-invasieve testen een bepaald diagnostisch probleem, meestal van radiologische aard, niet hadden kunnen oplossen. Op basis van 164 patiënten liet de studie zien dat PET een gunstige invloed had op het diagnostische begrip bij 84% van de patiënten en dat, volgens de clinici, PET het behandelingsbeleid in 50% in positieve zin had bepaald. Afstel van chirurgie op basis van PET werd het meest frequent gerapporteerd. PET bleek een sleutelrol te hebben gespeeld bij één op de vier patiënten in de studie.

Ten behoeve van het ontwerp van de diagnostische RCT voerden we een literatuuranalyse uit (in 1996 en herhaald in 2005) om andere diagnostische RCT's te vinden. In *hoofdstuk 3* worden de resultaten van die zoekactie gepresenteerd. We ontwikkelden een taxonomie van diagnostische RCT's op basis van de vergelijking binnen studies en de uitkomstmaten. De taxonomie maakte het mogelijk om eventuele veranderingen in de tijd, in uitkomstmaten in relatie tot de beeldvormende testen, te onderzoeken. De eerste zoekactie omvatte drie verschillende jaren over een tijdspanne van 1990 tot 2002. De tweede zoekactie was volledig gericht op studies waarbij MRI als diagnostisch instrument onderwerp van studie was, en omvatte de gehele periode van 1990 tot en met 2005. MRI kan worden beschouwd als de meest recente tomografische beeldvormende techniek vóór de komst van PET, en we verwachtten overeenkomsten in de evaluatiemethoden. Uit de specifieke zoekactie bleek dat het grootste deel van de gevonden studies betrekking had op het vergelijken van 'tracers' of technische uitvoeringskenmerken behorende bij een enkel diagnostisch instrument. Uiteindelijk konden we slechts 15 RCT's vinden waarin patiëntuitkomsten werden bestudeerd. De ziekte-indicaties varieerden van cardiovasculaire en gewrichts- en spierproblemen tot kanker. In de tweede zoekactie, die in totaal 16 jaar omvatte, vonden we 13 unieke RCT's waarin de waarde van MRI werd onderzocht. Vier van deze studies werden gepubliceerd in 2005 en vier in de periode van 2000 tot 2005. In drie van de 13 studies werden twee onderzoekstechnieken bij dezelfde patiënt toegepast. In de overige tien studies werden de patiënten willekeurig verdeeld over de beide strategieën. Op basis van de resultaten concluderen we onder andere dat, hoewel in veel literatuur aangehaald als optimale methode voor de evaluatie van diagnostische technieken, RCT's niet met regelmaat worden uitgevoerd. Daardoor ontbreekt in het algemeen het bewijs voor een (kosten)effectieve toepassing van beeldvormende technieken.

Omdat het literatuuronderzoek liet zien dat D-RCT's schaars zijn en zeker op het gebied van PET nog niet waren uitgevoerd, achtten we het van belang om het ontwerp en de logistiek van de studie te publiceren (*hoofdstuk 4*).⁷ De studie kreeg als acroniem de naam PLUS mee (PET in LUnG cancer Staging). Om aan de studie deel te nemen moesten de patiënten aan een aantal criteria voldoen: zo moest onder andere sprake zijn van een verdenking op, of bewezen NSCLC op basis van klinische stadiëring (klinisch stadium I-III), en de patiënten moesten medisch operabel en potentieel resectabel zijn naar het oordeel van de plaatselijk behandelend longarts. Net voordat een invasieve ingreep (diagnostisch of therapeutisch) aan de orde kwam, werden de patiënten willekeurig ingedeeld in de groep 'conventionele strategie' of 'conventionele strategie met PET-scan'. Veel aandacht werd besteed aan de definitie van het eindpunt. Als primair eindpunt werd gekozen voor het verschil in aantal 'onnodige of onterechte' chirurgische ingrepen (thoracotomieën) tussen beide groepen. Een thoracotomie werd onnodig of onterecht geacht indien er sprake was van: goedaardige aandoeningen, pathologisch bewezen mediastinale lymfeklieren (stadium IIIA-N2), stadium IIIB of hoger, exploratieve chirurgie of terugkerende ziekte of sterfte binnen één jaar na randomisatie. Het protocol vereiste pathologische bevestiging van beleidsbepalende PET-positieve verdachte laesies. Naast secundaire eindpunten als morbiditeit en duur van het diagnostisch proces werd een kostenanalyse gepland. De studie werd zodanig ontworpen dat met 95% zekerheid een absoluut verschil van 25% in het aantal thoracotomieën zou kunnen worden vastgesteld (met onderscheidend vermogen van 90%), uitgaande van 45% onterechte thoracotomieën in de huidige situatie (zoals uit het vooronderzoek was gebleken). Analyses zouden worden uitgevoerd volgens het 'intention-to-diagnose' principe (patiënten worden in de analyses meegenomen bij de groep waar ze volgens randomisatie aan toebehooren).

In één enkel jaar werden 188 patiënten uit negen verschillende ziekenhuizen in de studie opgenomen.⁸ Gegevens uit de kankerregistratie van de IKA-regio suggereren dat dit aantal overeenkomt met 65% van de potentieel beschikbare populatie in dat jaar. 96 patiënten werden ingedeeld in de 'conventionele strategie'-groep (CWU) en 92 in de 'conventionele + PET'-groep (*hoofdstuk 5*). In beide groepen had 70% van de patiënten een stadium I- of II-ziekte. Een significant kleiner deel van de patiënten in de CWU+PET-groep onderging een onterechte thoracotomie in vergelijking met de CWU-groep. De relatieve afname bedroeg 51% ($p=0,003$) en de absolute afname 20% (41% van de thoracotomieën in de CWU-groep bleek onterecht). Dit resultaat betekent dat op elke vijf patiënten die een PET-scan ondergaan één onterechte thoracotomie wordt voorkomen. In 81% van de gevallen suggereerde PET terecht het doen van een thoracotomie en in 71% om er terecht van af te zien. Het belangrijkste effect van PET was het tonen van meer ziektelast, ofwel een hoger stadium.

Tegelijkertijd met de medische gegevens werden gegevens over kosten verzameld. Onder kosten werden gerekend alle invasieve en niet-invasieve testen (inclusief PET), longoperaties en het aantal dagen ziekenhuisverblijf. De kosten werden berekend op individuele basis en op ziekenhuisniveau. De kosten met betrekking tot PET omvatten personele kosten, afschrijvings- en onderhoudskosten, materiaal (FDG) en overhead. De kostprijs van een PET-scan varieerde van 736 tot 1588 euro, afhankelijk van het type ziekenhuis en de FDG-beschikbaarheid (*hoofdstuk 6*).⁹ In de 'CWU'-groep bedroegen de gemiddelde kosten per patiënt 9573 euro en in de CWU+PET-groep 8284 euro. De belangrijkste kostenpost bleek het aantal hersteldagen in het ziekenhuis na chirurgie. Uit sensitiviteitsanalyses, op basis van variaties in de effectiviteit van PET (betrouwbaarheidsintervallen uit de PLUS-studie), bleek dat de resultaten van de kostenanalyse robuust zijn in het voordeel van PET. In aanvulling op de kostenanalyse van de PLUS-studie onderzochten we drie scenario's waarbij verschillende parameters van PET en de ziekenhuissituatie werden gevarieerd. In een 'dure' variant werd uitgegaan van PET in een universitaire setting, met FDG-productie ter plekke en PET-scans voor zowel klinische als onderzoeksdoeleinden. Als een 'goedkope' variant werd de situatie gekozen van een algemeen ziekenhuis waar twaalf klinische PET-scans per dag kunnen worden gemaakt en de FDG van elders wordt betrokken. De 'tussen-in'-variant vertegenwoordigde een groot regionaal ziekenhuis waar FDG ter plekke beschikbaar is, acht PET-scans per dag worden gemaakt en slechts beperkt onderzoek wordt verricht. Alle scenario's, uitgaande van de puntschatting van de effectiviteit van PET, vielen gunstig uit voor PET. Alleen bij de minste effectiviteit van PET (één onterechte operatie voorkomen tegen 14 PET-scans) in de duurste variant (universiteit) bleek de conventionele aanpak goedkoper (namelijk 542 euro).

Niet alle experts onderschrijven de noodzaak voor randomisatie bij diagnostische studies.¹⁰ Onnodig gebruik van middelen en ethische bezwaren worden zelfs gebruikt als argumenten tegen het doen van RCT's. Dit zette ons ertoe aan om onze ideeën over het nut en de noodzaak van gerandomiseerd diagnostisch onderzoek uiteen te zetten (*hoofdstuk 7*).¹¹ Samenvattend stellen wij dat de mate waarin een patiënt uiteindelijk voordeel heeft bij het toepassen van een diagnostisch instrument (in termen van iatrogene toxiciteit of verbetering in overleving) het best kan worden onderzocht door gelijktijdige vergelijking met de bestaande klinische praktijk zonder dat daar onrealistische of niet-verifieerbare aannames voor hoeven te worden gemaakt. De meest efficiënte strategie daartoe is, om de patiënten willekeurig te verdelen over de conventionele strategie met of zonder de nieuwe diagnostiek, waardoor bekende en deels onbekende maar bepalende factoren evenwichtig worden verdeeld en de test als enige variabele overblijft. Daarnaast is het van belang om voldoende patiënten in de studie op te nemen. Het opzetten van gerandomiseerde studies brengt bovendien andere voordelen met zich mee die niet zozeer met het randomiseren zelf te maken hebben, maar met intrinsieke factoren als transparantie, doordat een duidelijk protocol aan de studie ten grondslag ligt, en het opstellen van hypothesen met bijbehorende berekeningen voor de steekproefgrootte. De twee gerandomiseerde studies met FDG-PET bij longkanker laten zien dat dergelijke pragmatische RCT's uitvoerbaar zijn. In de PLUS-studie werden 188 patiënten, met stadium I-III, in één jaar geïncludeerd. In de POORT-studie, waarbij de rol van FDG-PET vroeg in het diagnostische traject werd gepositioneerd ter mogelijke vervanging van andere testen, werden in minder dan twee jaar tijd 465 patiënten aangemeld.¹² Het feit dat FDG-PET in die tijd nog weinig beschikbaar was heeft mogelijk wel bijgedragen aan dit succes.

Hoofdstuk 8 is een totaaloverzicht van onze ervaring ten aanzien van de evaluatie van de kosteneffectiviteit van FDG-PET voor verschillende indicaties. Het vormt een raamwerk dat in grote lijnen de hiërarchische aanpak volgt voor de waardebeoordeling van diagnostische studies, zoals eerder hierboven beschreven. Het startpunt van onze studies was de geconstateerde inefficiëntie bij conventionele diagnostiek, waarbij de betere accuratesse van FDG-PET, zoals gesuggereerd, winst zou kunnen opleveren. Het retrospectieve cohortonderzoek toonde de aard van de problematiek en betekende een belangrijke motivering voor de klinici om aan verder onderzoek mee te werken. De 'aanloop'-studie naar de klinische waarde van PET vormde de leercurve voor diagnostici en behandelaars. Uiteindelijk droegen de opeenvolging en samenhang van de studies bij aan een vruchtbare samenwerking, wat naar onze mening de RCT's en het ontwikkelen en implementeren van richtlijnen in de regio ten goede kwam. Uit gegevens van de IKA-regio blijkt dat met de komst van de PET-scanner in de regio het absolute aantal longresecties met 20% is afgenomen (*Figuur 1*)²⁸. Dit is geheel in overeenkomst met de resultaten uit de PLUS-studie.

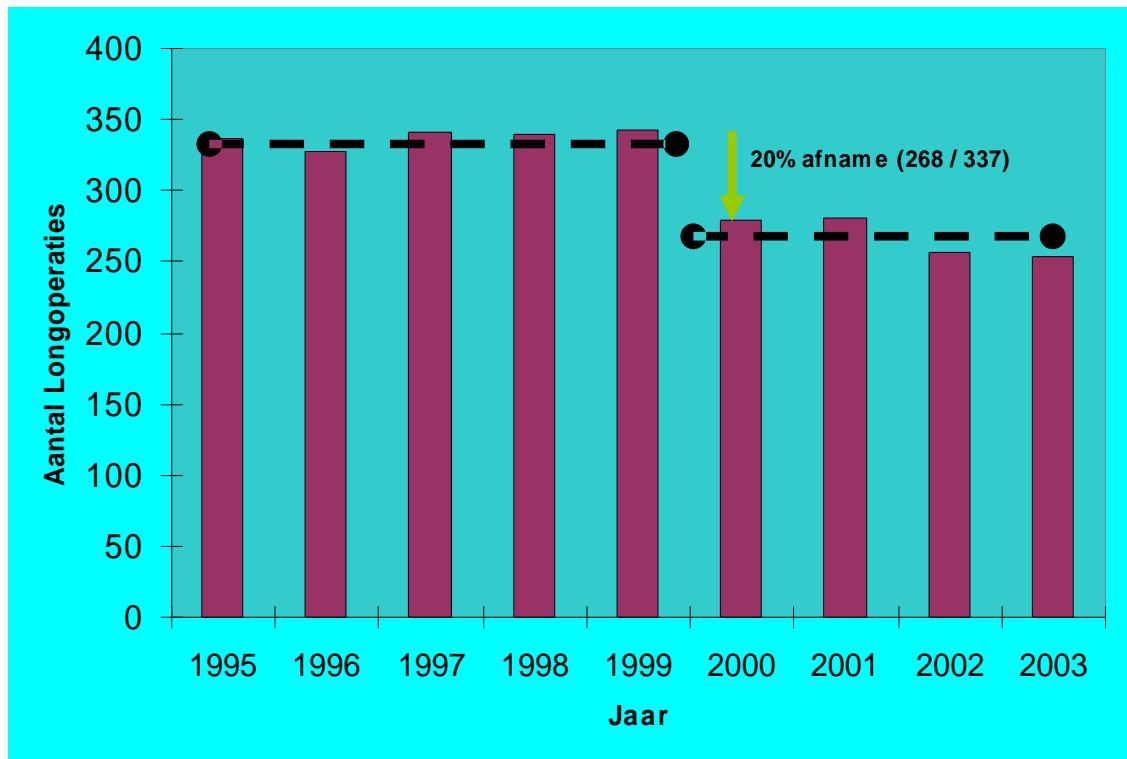
Amsterdam, 13 december 2006

Dr. H. van Tinteren

Promotoren: prof.dr. M. Boers, klinisch epidemioloog
prof.dr. G.J.J. Teule, nucleair-geneeskundige
Copromotor: prof.dr. O.S. Hoekstra, nucleair-geneeskundige

Vrije Universiteit medisch centrum

Figuur 1: Aantal of longresecties voor en na implementatie van richtlijnen met betrekking tot het gebruik van PET bij NSCLC in 2000. De IKA-regio, waar de richtlijn werd geïntroduceerd, bestrijkt 2,6 miljoen inwoners (Bron: Kanker registratie IKA, Visser O. Aantal longresecties in de IKA regio, maart 2005).



Literatuur

1. Herder GJ, Verboom P, Smit EF, et al. Practice, efficacy and cost of staging suspected non-small cell lung cancer: a retrospective study in two Dutch hospitals. *Thorax* 2002;57:11-4.
2. Flynn K, Adams E. Technology assessment: positron emission tomography. Original Report, 1996.
3. Freedman LS. Evaluating and comparing imaging techniques: a review and classification of study designs. *Br J Radiol* 1987;60:1071-81.
4. Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. *Med Decis Making* 1991;11:88-94.
5. Verboom P, Herder GJ, Hoekstra OS, et al. Staging of non-small-cell lung cancer and application of FDG-PET. A cost modeling approach. *Int J Technol Assess Health Care* 2002;18:576-85.
6. Herder GJ, Tinteren H van, Comans EF, et al. Prospective use of serial questionnaires to evaluate the therapeutic efficacy of 18F-fluorodeoxyglucose (FDG) positron emission tomography (PET) in suspected lung cancer. *Thorax* 2003;58:47-51.
7. Tinteren H van, Hoekstra OS, Smit EF, et al. Toward less futile surgery in non-small cell lung cancer? A randomized clinical trial to evaluate the cost-effectiveness of positron emission tomography. *Control Clin Trials* 2001;22:89-98.
8. Tinteren H van, Hoekstra OS, Smit EF, et al. Effectiveness of positron emission tomography in the preoperative assessment of patients with suspected non-small-cell lung cancer: the PLUS multicentre randomised trial. *Lancet* 2002;359:1388-93.
9. Verboom P, Tinteren H van, Hoekstra OS, et al. Cost-effectiveness of FDG-PET in staging non-small cell lung cancer: the PLUS study. *Eur J Nucl Med Mol Imaging* 2003;30:1444-9.
10. Valk PE. Do we need randomised trials to evaluate diagnostic procedures? Against. *Eur J Nucl Med Mol Imaging* 2004;31:132-5.
11. Tinteren H van, Hoekstra OS, Boers M. Do we need randomised trials to evaluate diagnostic procedures? For. *Eur J Nucl Med Mol Imaging* 2004;31:129-31.
12. Herder GJ, Kramer H, Hoekstra OS, et al. Traditional versus up-front [18F] fluorodeoxyglucose-positron emission tomography staging of non-small-cell lung cancer: a Dutch cooperative randomized study. *J Clin Oncol* 2006;24:1800-6.
13. Meerbeek JP van, Koning CC, Tjan-Heijnen VC, et al. [Guideline on 'non-small cell lung carcinoma; staging and treatment']. *Ned Tijdschr Geneesk* 2005;149:72-7.
14. Tinteren H van, Smit EF, Hoekstra OS. FDG-PET in addition to conventional work-up in non-small-cell lung cancer. *J Clin Oncol* 2005;23:1591-2.
15. Beyer T, Townsend DW, Brun T, et al. A combined PET/CT scanner for clinical oncology. *J Nucl Med* 2000;41:1369-79.
16. Vogel WV, Oyen WJ, Barentsz JO, et al. PET/CT: panacea, redundancy, or something in between? *J Nucl Med* 2004;45(Suppl 1):15S-24S.
17. Townsend DW, Carney JP, Yap JT, et al. PET/CT today and tomorrow. *J Nucl Med* 2004;45(Suppl 1):4S-14S.
18. Halpern BS, Schiepers C, Weber WA, et al. Presurgical staging of non-small cell lung cancer: positron emission tomography, integrated positron emission tomography/CT, and software image fusion. *Chest* 2005;128:2289-97.
19. Cerfolio RJ, Ojha B, Bryant AS, et al. The accuracy of integrated PET-CT compared with dedicated PET alone for the staging of patients with nonsmall cell lung cancer. *Ann Thorac Surg* 2004;78:1017-23.
20. Lardinois D, Weder W, Hany TF, et al. Staging of non-small-cell lung cancer with integrated positron-emission tomography and computed tomography. *N Engl J Med* 2003;348:2500-7.
21. Antoch G, Stattaus J, Nemat AT, et al. Non-small cell lung cancer: dual-modality PET/CT in preoperative staging. *Radiology* 2003;229:526-33.
22. Keidar Z, Haim N, Guralnik L, et al. PET/CT using 18F-FDG in suspected lung cancer recurrence: diagnostic value and impact on patient management. *J Nucl Med* 2004;45:1640-6.
23. Allen-Auerbach M, Yeom K, Park J, et al. Standard PET/CT of the chest during shallow breathing is inadequate for comprehensive staging of lung cancer. *J Nucl Med* 2006;47:298-301.
24. Annema JT, Hoekstra OS, Smit EF, et al. Towards a minimally invasive staging strategy in NSCLC: analysis of PET positive mediastinal lesions by EUS-FNA. *Lung Cancer* 2004;44:53-60.
25. Vansteenkiste J, Fischer BM, Dooms C, et al. Positron-emission tomography in prognostic and therapeutic assessment of lung cancer: systematic review. *Lancet Oncol* 2004;5:531-40.
26. Hoekstra CJ, Stroobants SG, Smit EF, et al. Prognostic relevance of response evaluation using [18F]-2-fluoro-2-deoxy-D-glucose positron emission tomography in patients with locally advanced non-small-cell lung cancer. *J Clin Oncol* 2005;23:8362-70.
27. Cerfolio RJ, Bryant AS, Winokur TS, et al. Repeat FDG-PET after neoadjuvant therapy is a predictor of pathologic response in patients with non-small cell lung cancer. *Ann Thorac Surg* 2004;78:1903-9.
28. Tinteren H van, Hoekstra OS, Smit EF, Boers M. The implementation of PET in non-small-cell lung cancer in the Netherlands. *Clin Oncol* 2006;18:156-7.

This document was created with Win2PDF available at <http://www.daneprairie.com>.
The unregistered version of Win2PDF is for evaluation or non-commercial use only.