

GERANDOMISEERDE KLINISCHE VELDSTUDIES IN DE VETERINAIRE EPIDEMIOLOGIE

D. Maes¹, K. Mintiens², H. Laevens², F. Boelaert³, D. Verloo², J. Dewulf¹

¹ Vakgroep Voortplanting, Verloskunde en Bedrijfsdiergeneeskunde, Afdeling voor Veterinaire Epidemiologie, Faculteit Diergeneeskunde, Universiteit Gent, Salisburylaan 133, B-9820 Merelbeke

² Coördinatiecentrum voor Diergeneeskundige Diagnostiek, Centrum voor Onderzoek in de Diergeneeskunde en Agrochemie, Groeselenberg 99, B-1180 Brussel

³ Europese Autoriteit voor Voedselveiligheid (European Food Safety Authority), Palazzo Ducale, Parco Ducale 3, I-43100 Parma, Italië
Dominiek.maes@UGent.be

SAMENVATTING

In tegenstelling tot observationeel onderzoek gaat men bij klinische veldstudies interveniëren en dieren behandelen volgens verschillende protocollen om de effecten van een bepaalde profylactische of therapeutische ingreep te onderzoeken. Omdat aan een ganse reeks voorwaarden moet voldaan worden om uit een klinische veldstudie betrouwbare conclusies te trekken, is het essentieel dat men een studieprotocol opstelt. De voornaamste aspecten hierbij zijn het formuleren van specifieke doelstellingen en daaraan gekoppeld de hypothese die men wenst te toetsen, het formuleren van selectiecriteria, het gebruik van een geschikte controlegroep en een goede randomisatie- en behandelingsprocedure. Tevens zijn het vastleggen van de parameters die worden gebruikt ter vergelijking van de groepen, de opvolgingsprocedure van de studieobjecten met aandacht voor blinding van de behandeling en het beperken van het aantal dieren dat uitvalt, de statistische analyse en de interpretatie van de resultaten van zeer groot belang. De dubbelblinde, (placebo)gecontroleerde en gerandomiseerde studie is de meest geschikte studievorm om duidelijke uitspraken te kunnen doen omtrent het causaal effect van diergeneeskundige ingrepen.

INLEIDING

Epidemiologische studies kunnen op verschillende manieren ingedeeld worden. Een belangrijke stap bij de indeling is onderscheid maken tussen observationele studies en klinische veldstudies. In observationeel onderzoek beperkt men zich tot het beschrijven en analyseren van de vastgestelde frequenties van het voorkomen van bepaalde kenmerken of ziekten in de bestudeerde populatie en hun eventuele associatie met bepaalde omgevingsfactoren (Dewulf *et al.*, 2005). In klinische veldstudies daarentegen komt men tussen in het gebeuren. De waarnemer heeft zelf de omstandigheden waarin de studie verloopt deels of geheel onder controle en hij kan bijvoorbeeld de dieren die in de studie opgenomen worden, indelen in een behandelings- en een controlegroep. Een klinische veldstudie heeft tot doel de positieve en de negatieve effecten van diergeneeskundig handelen te bestuderen. Het lijkt eenvoudig om tussen een diergeneeskundige behandeling en de daaropvolgende genezing een oorzakelijk verband te veronderstellen. Dit hoeft niet noodzakelijk het geval

te zijn. De hamvraag is dan ook hoe dergelijke oorzakelijke verbanden kunnen worden aangetoond. Daarbij speelt de klinische veldstudie waarbij dieren (of groepen dieren) op een toevallige wijze (randomisatie) worden toegekend aan de verschillende behandelingen onder studie, een cruciale rol.

De klinische studie

In een klinische veldstudie worden twee of meerdere behandelingen 'at random' aan dieren of groepen dieren toegekend. Het is een prospectieve studie om de effecten van een bepaalde profylactische of therapeutische ingreep te onderzoeken (Petrie en Watson, 1999). Dikwijls gaat het om het uittesten van het effect van nieuwe geneesmiddelen (bijvoorbeeld vaccins, antimicrobiële middelen), nieuwe behandelingsvormen (bijvoorbeeld chirurgie versus conservatieve behandeling) of van bepaalde veranderingen in bedrijfsvoering (bijvoorbeeld hokbezetting, voeding) op de ontwikkeling van het ziekteproces zelf (ontstaan van de ziekte of letsels), de evolutie van de

Tabel 1. Voornaamste onderscheid tussen een klinische veldstudie en een laboratoriumexperiment.

	Klinische veldstudie	Laboratoriumexperiment
Variabiliteit	hoog	laag
Ziekte	echt	gecontroleerd (model)
Dieren	heterogeen	homogeen
Bedrijfsvoering	echt	gecontroleerd

ziekte (snelheid van genezen, ergheid van de symptomen) of de productieresultaten (dagelijkse groei, melkproductie).

In een klinische veldstudie heeft men geen invloed op het tijdstip van de infectie en de infectiedosis. Men is immers afhankelijk van het natuurlijk verloop van de infectie op het bedrijf of bij het dier. Onder praktijkomstandigheden heeft men meestal ook te maken met een heterogene groep dieren (Tabel 1). Vaak zal men daarom in een eerste fase een nieuwe therapie onder meer gecontroleerde omstandigheden uittesten, waarbij de dieren meer homogeen zijn en men bijvoorbeeld het infectietijdstip en de infectiedosis kan controleren. In dergelijke laboratoriumexperimenten (*laboratory experiment*) stelt men behandelingen bij de dieren in én bepaalt men bijvoorbeeld het tijdstip van infectie, het soort infectie en de infectiedosis (*challenge*). Tevens heeft men de karakteristieken van de proefdieren (leeftijd, ras, gewicht, geslacht, enz.) onder controle, waardoor de dieren in het experiment meestal vrij homogeen zijn.

Het grote voordeel van klinische veldstudies in vergelijking met laboratoriumexperimenten ligt in het feit dat men minder problemen heeft met de extrapolatie van de resultaten naar de praktijk toe. De externe validiteit van een klinische veldstudie vormt minder een probleem dan bij laboratoriumexperimenten. De studie wordt immers onder praktijkomstandigheden uitgevoerd. Klinische veldstudies kunnen gebruikt worden om een specifieke hypothese te toetsen, maar ook om bevindingen van observationale studies of laboratoriumexperimenten te bevestigen.

Protocol van een klinische veldstudie

Essentieel bij een klinische veldstudie is dat men vóór de start van het onderzoek een gedetailleerd protocol opstelt en dat men de proef uitvoert volgens

Tabel 2. Algemene inhoud van een protocol van een klinische studie.

1.	Achtergrond en 'het waarom' van de studie
2.	Specifieke doelstellingen
3.	Criteria voor het selecteren van dieren
4.	Beschrijving van de studieopzet
5.	Beschrijving van de behandelingsprocedure
6.	Beschrijving van alle klinische, biologische e.a. methoden
7.	Kwaliteitscontrole
8.	Parameters ter vergelijking
9.	Opvolging (<i>follow-up</i>)
10.	Definitie van bijwerkingen
11.	(Statistische) analyse

de richtlijnen van het protocol. Men spreekt van "good clinical practice" (GCP) als een studie op die manier wordt uitgevoerd. De algemene inhoud van een protocol van een klinische veldstudie wordt in Tabel 2 samengevat.

Achtergrond en 'het waarom' van de studie

Vóór men begint met een klinische veldstudie moet men er zich van vergewissen of de vraag die men wenst te beantwoorden niet reeds het onderwerp heeft uitgemaakt van een vorig onderzoek. Tevens dient men na te gaan of een klinische veldstudie wel de meest geschikte vorm is om de onderzoeksvraag op te lossen. Het is bijvoorbeeld onverantwoord, zowel

ethisch als economisch, om een grote klinische veldstudie uit te voeren met een vaccin waarvan de werkzaamheid nog niet bewezen is in meer gecontroleerde omstandigheden.

Specifieke doelstellingen

Het is belangrijk om na te gaan of de doelstellingen van het onderzoek wetenschappelijk en/of diergeneeskundig gezien relevant zijn en voldoende opwegen tegenover de kosten van de studie en het eventuele dierlijke leed. De objectieven van de studie moeten nauwkeurig worden vastgelegd en de onderzoeker moet er zich aan houden eens de studie is aangevat. Op basis van de doelstellingen zal een primaire hypothese geformuleerd worden die achteraf al dan niet zal weerlegd worden op basis van de verzamelde gegevens.

Criteria voor het selecteren van dieren

De onderzoeker dient een willekeurige steekproef uit de populatie te nemen waarop dan de studie zal uitgevoerd worden. Indien een behandeling aan een individueel dier wordt toegekend, zal de onderzoeker dieren selecteren, maar soms bestaat de eenheid waaraan de behandeling wordt toegekend uit een groep dieren (hok, toom, kennel, afdeling, stal, bedrijf) en dan zal de onderzoeker groepen dieren selecteren. Een overzicht van de verschillende methoden die gebruikt kunnen worden om onderzoekseenheden te selecteren, wordt gegeven door Boelaert *et al.* (2004).

Het is belangrijk om bij een klinische veldstudie het selectieproces duidelijk weer te geven. Dit is een zeer delicaat punt. Meestal wordt gebruik gemaakt van in- en exclusiecriteria die bepalen welke dieren of bedrijven in de studie opgenomen worden en welke niet. Wanneer men teveel criteria invoert, zullen de conclusies enkel gelden voor een beperkt deel van de populatie aangezien de resultaten dan enkel representatief zijn voor die dieren of bedrijven die aan de criteria voldoen (Martin *et al.*, 1994; Lesaffre, 1998). Wanneer men te weinig of geen criteria hanteert, dan kunnen de verschillende bedrijven die aan het onderzoek deelnemen zo divers zijn dat de resultaten dermate variabel zijn dat geen éénduidige conclusies kunnen getrokken worden. Alvorens dieren of bedrijven in een studie worden opgenomen, moeten ze dus door een filteringsproces. Om een idee te krijgen van de algemene toepasbaarheid van de klinische veldstudie, noteert men in sommige studies elk dier, of elk bedrijf dat gecontacteerd werd of dat zich

aangemeld heeft voor de studie. Hierdoor krijgt men een idee van het percentage van het totaal aantal gecontacteerde dieren (bedrijven) dat in de studie opgenomen is.

Tenslotte moet de eigenaar of de veehouder, nadat hij/zij volledig werd ingelicht, nog de toestemming geven om aan de studie deel te nemen. Hij/zij moet met andere woorden zijn/haar "informed consent" geven.

Beschrijving van de studieopzet

De controlegroep

In een goed opgezette klinische veldstudie moet steeds een controlegroep worden ingesloten. Dit is noodzakelijk omdat men anders nooit kan concluderen dat het dankzij de behandeling is dat de behandelde dieren gezond zijn. Veronderstel bijvoorbeeld dat een vaccin dat in een veldsituatie geen bescherming biedt, getest wordt en er zich op het moment van testen geen infecties voordoen. Dan zou men onterecht concluderen dat het vaccin werkt. In het geval ook een controlegroep in de studie zou opgenomen zijn en ook al zouden deze dieren in het studieverloop gezond zijn, zou terecht besloten worden dat geen conclusies kunnen getrokken worden betreffende de werkzaamheid van het vaccin. De controlegroep moet simultaan opgevolgd worden met de behandelingsgroep. Een historische controlegroep, dit wil zeggen een groep dieren waarvan de resultaten bekomen werden vóór de start van de studie en waarmee de resultaten van de behandelingsgroep vergeleken worden, is in dit opzicht zeker af te raden omdat tijdgebonden evoluties kunnen optreden. *Op varkensbedrijven zijn de foktechnische resultaten de laatste decennia gemiddeld gezien duidelijk verbeterd (Maes et al., 2003). De huidige foktechnische resultaten, bijvoorbeeld na vaccinatie vergelijken met deze van één jaar geleden, zal dus meestal leiden tot een overschatting van de effecten (positieve bias).* Historische controles worden soms gekozen omdat men dan het aantal dieren dat in het onderzoek wordt opgenomen kan beperken of omdat men (uit ethische overwegingen) een nieuwe behandeling aan alle dieren wenst toe te dienen. Bij de selectie van een controlegroep kan men zich bovendien de vraag stellen of men beter een negatieve controlegroep, dit wil zeggen een groep die blootgesteld wordt aan de aan-doening en niet behandeld wordt, dan wel een positieve controlegroep, dit wil zeggen een groep die behandeld wordt met een standaardtherapie, verkiest.

Indien het een ernstige aandoening betreft waarvoor er een goede standaardtherapie bekend is, dan zal men meestal deze therapie als vergelijkingsbasis nemen. Als de aandoening minder ernstig is en/of er geen goede therapie voorhanden is, dan zal men meestal een negatieve controlegroep kiezen. Vanuit statistisch oogpunt zijn er minder dieren nodig om superioriteit aan te tonen van een nieuwe behandeling tegenover een negatieve controlegroep, dan om equivalentie aan te tonen van een nieuwe behandeling tegenover een bestaande standaardtherapie (Rosner, 1995) (zie verder).

Blokken en randomisatie garanderen vergelijkbaarheid tussen behandelings- en controlegroepen

Essentieel voor een klinische veldstudie is dat de behandelings- en de controlegroepen vergelijkbaar zijn voor wat betreft alle factoren uitgezonderd de behandelingsfactor. In het bijzonder moet de vergelijkbaarheid verzekerd zijn voor alle factoren die zelf invloed kunnen hebben op de bestudeerde ziekte (Dewulf *et al.*, 2005).

Voor factoren waarvan geweten is dat ze een effect hebben op de bestudeerde ziekte kan 'geblokt' worden. Een blok bestaat uit een homogene groep dieren (of groepen dieren) die eenzelfde waarde hebben voor de factor. Binnen deze homogene groep worden dan de behandelingen gerandomiseerd ("restricted randomisation") zodat het effect van de behandeling kan bestudeerd worden binnen de verschillende blokken. Het blokken van dergelijke factoren is niet altijd mogelijk en bovendien kan men nooit uitsluiten dat er een aantal onbekende factoren is dat ook een invloed uitoefent.

Daarom is het nodig de behandelingen volgens toeval toe te kennen aan de onderzoekseenheden (dieren of groepen dieren), zodat enkel het toeval beslist welke behandeling een dier krijgt (bijvoorbeeld al dan niet vaccinatie). Op die manier kan een vertekening door selectie ("selection bias") worden tegengegaan: een onderzoeker kent op een subjectieve manier een behandeling toe aan specifieke dieren zodat er een verschil ontstaat, zelfs indien de behandeling geen effect teweegbrengt. *Een dergelijke vertekening ontstaat bijvoorbeeld wanneer het effect van een vaccin tegen Infectieuze Boviene Rhinotracheïtis (IBR) wordt onderzocht op 20 rundveebedrijven (10 gevaccineerde en 10 niet-gevaccineerde). Als de behandeling niet op een aselektieve wijze (at random) toegekend wordt aan de bedrijven maar overgelaten wordt aan de keuze van de vee-*

houder, kan het zijn dat er op de 10 bedrijven met het beste management wordt gevaccineerd omdat dit de veehouders zijn die zich bewust zijn van het risico en daarom opteren voor vaccinatie. Op de 10 andere bedrijven, met een minder goed management, wordt niet gevaccineerd. Wanneer dan een verschil in voorkomen van IBR wordt gevonden tussen de twee groepen, kan niet met zekerheid worden uitgemaakt of dit het gevolg is van het vaccin dan wel van het management.

In veel gevallen vormt het individuele dier de experimentele eenheid. Dit betekent dat het individuele dier de kleinste eenheid is die een behandeling krijgt en waarbij het effect onafhankelijk is van de andere dieren in de studie. Echter, in sommige gevallen en vooral bij landbouwhuisdieren komt het frequent voor dat niet het individuele dier maar een groep dieren een eenheid vormt (Maes *et al.*, 1999). Bepaalde voeders, behandelingen of vaccinaties worden immers soms aan een volledige groep dieren toegediend en niet aan de individuele dieren afzonderlijk. Dit wordt gedaan omdat verschillende dieren van eenzelfde afdeling vaak niet onafhankelijk zijn van elkaar, zoals bij infectieziekten. De kans dat een tweede dier van eenzelfde afdeling wordt geïnfecteerd is immers groter dan de kans dat een dier van een andere afdeling waar de ziekte nog niet aanwezig is, geïnfecteerd wordt. In deze gevallen is een groepsrandomisatie te verkiezen.

Tenslotte kan de vergelijkbaarheid tussen groepen ook door biologische factoren beïnvloed worden. Veronderstel dat men een vaccinatie-experiment wenst uit te voeren tegen een infectieuze aandoening bij varkens. Vermits we vooral geïnteresseerd zijn in de effecten op bedrijfsniveau en vermits vaccinatie meestal op het volledige bedrijf wordt toegepast om zodoende een goede bedrijfsimmunitet te bekomen, dienen gevaccineerde met niet-gevaccineerde bedrijven vergeleken te worden. Omwille van logistieke en financiële redenen is dit echter niet steeds mogelijk en wordt dikwijls binnen een bedrijf slechts een deel van de dieren gevaccineerd en vergeleken met een andere groep dieren die niet gevaccineerd worden. Gevaccineerde dieren hebben een betere immuniteit en zorgen voor minder ziekteverspreiding op het bedrijf. Hierdoor wordt de infectiedruk bij de niet-gevaccineerde dieren lager en de niet-gevaccineerde dieren worden aldus bevoordeeld. Anderzijds fungeren de niet-gevaccineerde dieren als een infectiebron voor de gevaccineerde waardoor de gevaccineerde dieren benadeeld worden. In beide gevallen wordt het ver-

schil tussen gevaccineerde en niet-gevaccineerde dieren kleiner en heeft men dus minder kans om significante verschillen te kunnen aantonen. Dit is vooral het geval indien gevaccineerde en niet-gevaccineerde dieren rechtstreeks contact hebben met elkaar, bijvoorbeeld wanneer ze gemengd worden in eenzelfde hok. Indien het niet mogelijk is om gevaccineerde met niet-gevaccineerde bedrijven te vergelijken, kan het probleem van transmissie van pathogenen tussen de betrokken groepen beperkt worden door de dieren van beide groepen te huisvesten in aparte hokken, of beter nog in aparte afdelingen of stallen. Een gelijkaardig fenomeen doet zich voor in het geval van parasitaire behandelingen bij runderen, schapen en paarden op de weide. Als men slechts enkele dieren ontwormt, dan kunnen niet-behandelde dieren ‘profiteren’ van de gedaalde infectiedruk, of ‘lijden’ de behandelde dieren onder de hoge infectiedruk vanwege de niet-behandelde dieren. Het effect van de behandeling wordt dus onderschat.

Het aantal dieren, hokken of bedrijven

Het aantal dieren, hokken of bedrijven dat in de studie wordt opgenomen, moet vooraf vastgelegd worden. Dit aantal hangt nauw samen met het beoogde doel. Men moet er voor zorgen dat met de beschikbare middelen de resultaten zo betrouwbaar mogelijk zijn. Hierbij kunnen fouten van het type I en van het type II een rol spelen. Een type I-fout () treedt op wanneer de resultaten van het onderzoek wel een behandelingseffect aantonen terwijl er in werkelijkheid geen is. Een type II-fout () treedt op wanneer de resultaten van het onderzoek geen effect aantonen terwijl er in werkelijkheid wel een effect is. Wanneer men de kans op een type II-fout aftrekt van 1, dan bekomt men het onderscheidingsvermogen (*power*) van het experiment: dit is de kans dat in de studie terecht een significant verschil wordt aangetoond. De kans op een fout van het type II wordt meestal op 20% gebracht, het onderscheidingsvermogen van de studie is dan 80%. Analoog, wanneer men de kans van een type I-fout aftrekt van 1, dan bekomt men het betrouwbaarheidsniveau. De type I-fout wordt meestal op 5% gezet, en dus het betrouwbaarheidsniveau op 95%. Traditioneel heeft men steeds meer aandacht besteed aan het beperken van een type I-fout dan van een type II-fout. Een samenvatting van de verschillende mogelijkheden bij het testen van hypothesen wordt weergegeven in Tabel 3.

Om de kans op type I- en type II-fouten te reduceren, is het noodzakelijk het juiste aantal dieren op te

Tabel 3. Overzicht van de verschillende mogelijkheden bij het testen van hypothesen.

	Werkelijkheid	
	Niet-effectief	Effectief
Waargenomen in experiment Niet-effectief	Correct besluit: $1-\alpha$ betrouwbaarheidsniveau (vb. 95 %)	Foutief besluit β type II-fout (vb. 20 %)
Effectief	Foutief besluit: α type I-fout (vb. 5 %)	Correct besluit: $1-\beta$ power (vb. 80 %)

nemen in de studie. Te veel dieren in een onderzoek opnemen, is duur en kan leiden tot het significant aantonen van biologisch irrelevante factoren. Te weinig dieren opnemen, vermindert het onderscheidingsvermogen van de studie en aldus de kans om klinisch relevante effecten aan te tonen.

Het aantal dieren dat voor de studie nodig is, kan bepaald worden door *a priori* een “berekening van de benodigde steekproefgrootte” (*power calculation*) uit te voeren zodat met voldoende kans verschillen in behandelingen (indien aanwezig) kunnen aangetoond worden (Thrusfield, 1997). Het vereiste aantal dieren neemt toe als een kleine en/of vereist word(t)(en). Het noodzakelijke aantal dieren zal ook toenemen als er veel variatie is in de resultaten en als het effect van een behandeling klein is, m.a.w. als we een klein verschil wensen aan te tonen. Tenslotte hangt het aantal dieren ook af van het soort hypothese, meer bepaald of er eenzijdig of tweezijdig getest wordt. Bij eenzijdig testen wordt onder andere onderzocht of een nieuwe behandeling superieur is tegenover een bestaande behandeling. In dit geval zijn we enkel geïnteresseerd in een resultaat waarbij de nieuwe behandeling beter presteert dan de gebruikelijke behandeling. Bij tweezijdig testen daarentegen wordt onderzocht of twee behandelingsgroepen significant verschillend zijn van elkaar, zonder aan te geven van welke groep verwacht wordt om beter te presteren. Om met eenzelfde betrouwbaarheid een zelfde onderscheidingsvermogen te bereiken, zijn er bij tweezijdig testen meer dieren nodig dan bij eenzijdig testen.

Beschrijving van de behandelingsprocedure

Men dient een nauwkeurige en gedetailleerde beschrijving te geven van de behandelingsprocedure die men wenst te testen: toedieningswijze, dosering, tijdstip van toedienen, duur van behandeling, enz. Tevens worden de meeste klinische studies waarbij er invasieve ingrepen gebeuren of waarbij de dieren op het einde van de proef geëuthanaseerd worden, slechts uitgevoerd na goedkeuring door een ethische commissie voor dierproeven. De leden van een dergelijke commissie zien erop toe dat alle richtlijnen betreffende dierenwelzijn gerespecteerd worden.

Beschrijving van alle klinische, biologische en andere methoden

Het protocol bevat verder een beschrijving van alle methoden (bloednamen, autopsie, enz.) die worden aangewend om de bestudeerde factor en andere variabelen op te volgen.

Kwaliteitscontrole

De methodologie van alle aangewende technieken wordt gevolgd door een programma van kwaliteitscontrole. Voorbeelden zijn onder andere het bepalen van de concentratie van het antibioticum in het voeder, het bepalen van de hoeveelheid antigeen in het vaccin, het inbouwen van negatieve en positieve controles bij diagnostica, enz.

Parameters ter vergelijking

Essentieel voor een klinische veldstudie is dat de parameters die men zal gebruiken om de nieuwe behandelingsvorm te beoordelen, op voorhand worden vastgelegd. Zo weet men reeds op voorhand welke vorm van analyse zal aangewend worden en tot welke vorm van resultaatbeschrijving men zal komen. Men dient een onderscheid te maken tussen primordiale (*major variables*) en bijkomende parameters (*minor variables*). Bij voorkeur wordt een beperkt aantal primordiale parameters, die objectief meetbaar zijn, gebruikt. *Belangrijke foktechnische parameters bij vleesvarkens zijn bijvoorbeeld dagelijkse groei, voederconversie, sterftepercentage en kosten voor medicijnen; bijkomende parameters zijn bijvoorbeeld seroconversie, hoestindex, fecesscores, longletsels.*

Opvolging (follow-up)

Blindering

De beoordeling van de invloed van de behandeling op de gemeten parameter kan subjectief beïnvloed worden indien de eigenaar van het dier of de onderzoeker op de hoogte is van de behandeling die het dier kreeg. Om dit te vermijden, maakt men in klinische veldstudies vaak gebruik van *blinderingsprocedures* (Toma *et al.*, 1999). Bij een 'single-blind' onderzoek is ofwel de eigenaar ofwel de onderzoeker onwetend over de behandelingsvorm. Bij een 'double-blind' onderzoek is zowel de eigenaar als de onderzoeker onwetend over de behandelingsvorm. 'Double-blind' onderzoek is noodzakelijk wanneer de parameters niet op een objectieve manier kunnen gemeten worden, bijvoorbeeld fecesscores, de algemene klinische toestand, enz. Soms gebruikt men 'triple-blinding', waarbij zowel de eigenaar, de onderzoeker als de verantwoordelijke voor de analyse van de resultaten onwetend is over de behandelingsvorm. Om de blindering te kunnen garanderen, is het dikwijls nodig placebo's te gebruiken bij de controledieren. Soms wordt er geen placebo gebruikt omdat men het effect van een behandeling wil vergelijken met de praktijksituatie. Onder praktijkomstandigheden zullen dieren immers ofwel gevaccineerd worden, ofwel niets toegediend krijgen, ook geen injectie met een placebo (Maes *et al.*, 1999; Mateusen *et al.*, 2002).

Loss to follow-up

Wanneer de incidentie van een ziekte de primordiale variabele is, is het van kapitaal belang dat alle dieren in de proef tot het einde van het onderzoek kunnen worden gevolgd. Bij studies van lange duur of in zeer grote groepen dieren is dit echter niet steeds mogelijk. Alleen van de dieren of de bedrijven die de volledige studieduur aanwezig blijven, is het uiteindelijke effect bekend. Men dient het aantal dieren dat men uit het oog verliest, zoveel mogelijk te beperken en er tevens voor te zorgen dat het aantal dieren in de behandelingsgroep en het aantal in de controlegroep niet verschillend zijn. *Bij het overplaatsen van het kraamhok naar de batterij of van de batterij naar de vleesvarkensafdeling worden de kleinere biggen er soms uitgenomen. Op deze manier probeert men homogene groepen te krijgen. Wanneer de behandeling reeds op jonge leeftijd heeft plaatsgevonden, dient het wegnemen van biggen zoveel mogelijk vermeden te worden en het percentage weggenomen biggen gelijk te zijn voor de twee groepen.*

Tijdens het verloop van de proef zijn er dikwijls dieren die achterblijven of die (om ethische redenen) een bijkomende behandeling vereisen. Het is goed om vooraf een procedure klaar te hebben voor de opvolging van dergelijke probleemdieren en deze procedure duidelijk af te spreken met de eigenaar. *Wanneer in een vaccinatiestudie corticosteroiden worden gebruikt omdat de dieren erg ziek zijn, dan dienen deze dieren uit de studie genomen te worden. Corticosteroiden kunnen immers interfereren met de immuniteit en aldus de vaccinatie-resultaten beïnvloeden.*

'Compliance'

'Compliance' betekent het uitvoeren en opvolgen van de proef volgens de instructies van het protocol. Het gebeurt echter frequent dat veehouders en ook dierenartsen het protocol niet strikt volgen. 'Poor compliance' treedt nogal eens op in geval van een te vaag studieprotocol, onduidelijke instructies naar de veehouder toe, ingewikkelde en langdurige proeven, zwakke tussentijdse resultaten, enz. Een duidelijk protocol en eenduidige afspraken met veehouders zijn van cruciaal belang om het probleem van *poor compliance* te voorkomen.

Definitie van bijwerkingen

De meeste nieuwe behandelingsvormen kunnen ook bijwerkingen veroorzaken. Daarom moet er een methode worden ontwikkeld die mogelijke bijwerkingen (lokale of algemene, acute of meer chronische) kan opsporen. In bepaalde gevallen zal men de procedure die gevolgd wordt, wanneer er bijwerkingen optreden, op voorhand vastleggen.

(Statistische) analyse

Statistische analyse en interpretatie van de resultaten

Het protocol van de proef beschrijft tenslotte in detail de analysetechniek, met inbegrip van de statistische methoden. Vaak wordt gebruik gemaakt van de *P*-waarde om te beslissen of een geobserveerd verschil tussen behandelingen aan het toeval te wijten is of kan beschouwd worden als een echt bestaand verschil in de populatie. De *P*-waarde geeft de kans weer dat het geobserveerde verschil of een extremer verschil opgemeten wordt in de studie als er geen verschil zou zijn tussen de behandelingen. Bijgevolg zal bij een kleine *P*-waarde een verschil tussen de

behandelingen geclaimd worden. In het klassieke geval claimt men een significant resultaat als de *P*-waarde kleiner is dan 0,05 (5%), dit om het risico te minimaliseren op het foutief besluit dat de twee behandelingen verschillend zouden zijn wat de werking betreft. De waarde waaronder de *P*-waarde moet liggen om een significant resultaat te claimen, wordt het significantieniveau genoemd (type I-fout zie hoger).

Het berekenen van een (95%) betrouwbaarheidsinterval (BI) is een andere manier om na te gaan of er een significant verschil is tussen de groepen. Stel dat men 2 verschillende behandelingen van koliek bij paarden wil vergelijken op het vlak van overleving. Voor de vergelijking van de twee behandelingen is men geïnteresseerd in het verschil van de echte mortaliteitscijfers voor de twee behandelingen, bekomen voor de ganse populatie paarden met koliek. Dit verschil is niet bekend, maar op basis van de experimentele studie kan wel een BI voor het verschil bepaald worden. *Stel dat we op basis van een experiment met 50 paarden in elke groep ($n_A = n_B = 50$) een mortaliteit bekomen van $\hat{\pi}_A = 20\%$ in de behandelingsgroep A en van $\hat{\pi}_B = 30\%$ in behandelingsgroep B. Het 95% BI voor het verschil tussen de twee behandelingsgroepen, $\hat{\pi}_A - \hat{\pi}_B$, wordt weergegeven door*

$$\hat{\pi}_A - \hat{\pi}_B \pm 1,96 \sqrt{\frac{\hat{\pi}_A (1 - \hat{\pi}_A)}{n_A} + \frac{\hat{\pi}_B (1 - \hat{\pi}_B)}{n_B}} = [-0,269; 0,069]$$

*Uit dit BI blijkt aldus dat behandeling A zowel slechter als beter kan zijn dan behandeling B. De studie heeft helemaal geen uitsluitel gegeven over de relatieve waarde van behandeling A. Besluiten als zou behandeling A gelijkwaardig zijn aan of beter zijn dan B, zijn dus verkeerd. Indien men dit resultaat had bekomen met 2 maal 500 paarden, dan zou het 95% BI verkleinen tot $[-0,153; -0,047]$. Alhoewel het verschil in mortaliteit tussen beide behandelingsgroepen gelijk is in beide studies, heeft de laatste studie wel degelijk aangetoond (met 95% betrouwbaarheid) dat behandeling A een significant lagere mortaliteit met zich meebrengt dan behandeling B, omdat het BI voor het verschil niet de waarde 0 bevat. Het BI biedt in essentie dezelfde informatie als de *P*-waarde, maar het laat bovendien ook toe het verschil in werking van de behandelingen kwantitatief in te schatten.*

Statistische significantie versus klinische relevantie

Een statistisch significant resultaat is daarom nog niet klinisch relevant. *Neem het voorbeeld van een*

mortaliteitsstudie bij vleesvarkens. Bij een groep van 100 vleesvarkens met behandeling B wordt een sterftepercentage van 10% geobserveerd terwijl met behandeling A in een groep van 100 varkens dit cijfer teruggebracht wordt tot 2%, een resultaat dat zowel klinisch relevant (verschil in mortaliteit is dus $10\% - 2\% = 8\%$ en behandeling A resulteert dus in een duidelijke winst) als statistisch significant is ($P = 0,008$). Veronderstel nu dat eenzelfde studie werd opgezet met in elke groep 100 000 varkens. De nieuwe studie resulteert in een geobserveerd sterftepercentage van 10% in groep A en 9,68% in groep B. Dit resultaat heeft dezelfde P -waarde van 0,008 als voor het vorige experiment; met die nuance dat het verschil in sterftepercentage nu klein is en klinisch minder relevant. Bij de interpretatie van de resultaten is het dus belangrijk niet louter de P -waarden te beoordelen, maar ook rekening te houden met de grootte van het effect van de behandeling. Een bepaalde behandeling kan een groot effect hebben dat om een of andere reden (bijvoorbeeld een te gering aantal dieren) net niet significant is, terwijl men met een zeer groot aantal dieren een miniem effect significant kan maken.

BESLUIT

Uit het bovenstaande blijkt dat aan heel wat voorwaarden moet voldaan worden om uit een klinische veldstudie betrouwbare conclusies te kunnen trekken. Een (placebo-) gecontroleerde en gerandomiseerde klinische veldstudie is essentieel om een causaal verband aan te tonen en vormt daarmee de hoeksteen van wat sinds kort "evidence based veterinary medicine" genoemd wordt (Cockcroft en Holmes, 2003). Deze laatste term geeft aan dat het diergeneeskundig handelen niet uitsluitend uitgaat van de klinische ervaring van de dierenarts en van algemene fysiologische principes, maar dat diergeneeskundige handelingen pas ondernomen worden wanneer wetenschappelijk onderzoek evidentie gebracht heeft over het feit dat beslissingen ter zake goed onderbouwd zijn.

REFERENTIES

- Boelaert F., Verloo D., Dewulf J., Maes D., Mintiens K., Laevens H. (2005). Beschrijvende veterinaire epidemiologie. *Vlaams Diergeneeskundig Tijdschrift* 74, 14-26.
- Cockcroft P., Holmes M. (2003). *Handbook of evidence-based veterinary medicine*. Blackwell Publishing, Oxford UK, pp 210.
- Dewulf J., Maes D., Mintiens K., Laevens H., Boelaert F., Verloo D. (2005). Observationeel epidemiologisch onderzoek. *Vlaams Diergeneeskundig Tijdschrift* 74, 35-45.
- Lesaffre E. (1998). De statistische fundamente van een geneeskunde gebaseerd op evidentie. *Tijdschrift voor Geneeskunde* 54, 75-82.
- Maes D., Deluyker H., Verdonck M., Castryck F., Miry C., Vrijens B., Verbeke W., Viaene J., de Kruif A. (1999). Effect of vaccination against *Mycoplasma hyopneumoniae* in pig herds with an all-in/all-out production system. *Vaccine* 17, 1024-1034.
- Maes D., Verbeke W., Vicca J., Verdonck M., de Kruif A. (2003). Benefit to cost of vaccination against *Mycoplasma hyopneumoniae* in pig herds under Belgian market conditions from 1996 to 2000. *Livestock Production Science* 83, 85-93.
- Martin S., Meek A., Willeberg, P. (1994). In: Martin S., Meek A., Willeberg P. (eds). *Veterinary Epidemiology. Principles and Methods*. Iowa State University Press, Ames, Iowa, USA, pp 343.
- Mateusen B., Maes D., Van Goubergen M., Verdonck M., de Kruif A. (2002). Effectiveness of treatment with lincomycin hydrochloride and/or vaccination against *Mycoplasma hyopneumoniae* for controlling chronic respiratory disease in a herd of pigs. *The Veterinary Record* 151 (5), 135-140.
- Petrie A., Watson P. (1999). Experimental design and clinical trials. In: Petrie A., Watson P. (Editors). *Statistics for Veterinary and Animal Science* (Chapter 5), Blackwell Science Ltd., p 53-68.
- Rosner B. (1995). Hypothesis testing: two-sample inference. In: Rosner B. (Editor), *Fundamentals of Biostatistics*, 4th Edition, Wadsworth Publishing Company (ITP), California US, p 251-298.
- Thrusfield M. (1997). The design and conduct of clinical trials. In: Noordhuizen J., Frankena K., van der Hoofd C., Graat E. (editors). *Application of quantitative methods in veterinary epidemiology*, Wageningen Press, p 223-246.
- Toma B., Vaillancourt J., Dufour B., Eloit M., Moutou F., Marsh W., Bénét J., Sanaa M., Michel P. (1999). In: Toma B., Vaillancourt J., Dufour B., Eloit M., Moutou F., Marsh W., Bénét J., Sanaa M., Michel P. (Editors). *Dictionary of veterinary epidemiology*. Iowa State University Press, Ames Iowa US, p 25.