PROJECTE FINAL DE CARRERA

LLICENCIATURA EN CIÈNCIES I TÈCNIQUES ESTADÍSTIQUES

EFICIÈNCIA RELATIVA DEL "*CROSS-OVER*" RESPECTE DEL DISSENY DE GRUPS PARAL·LELS

Autores: Marta Benet i Mora
Raquel Garcia Esteban
Director: Erik Cobo Valeri

Facultat de Matemàtiques
POLITÈCNICA DE CATALUNYA

Facultat de Matemàtiques
i Estadística

UNIVERSITAT POLITÈCNICA DE CATALUNYA

DADES DEL PROJECTE:

Nom de l'estudiant: RAQUEL GARCIA ESTEBAN MARTA BENET i MORA

DNI: 38847573 36530990

Títol del Projecte: EFICIÈNCIA RELATIVA DEL "CROSS-OVER" RESPECTE DEL DISSENY DE GRUPS PARAL·LELS

Director del Projecte: ERIK COBO VALERI

Tutor del Projecte:

QUALIFICACIÓ

Excel·lent (9)

MEMBRES DEL TRIBUNAL (nom i signatura)

President: José Antonio González Alastrue

Vocal: Joan Carles Lario Lago

Secretari: En/N Cobo Valeri

Data: 27/05/03

SÍNTESI

La principal raó per la què s'escull un disseny "*cross-over*", enlloc d'un disseny de grups paral·lels, és que són necessaris un menor nombre de participants per obtenir la mateixa precisió en els estimadors. D'altra banda, cada individu actua com el seu propi control (eliminant la font de variabilitat introduïda per l'individu) i, en teoria, això fa possible comparar tractaments amb molta més precisió.

La grandària de la mostra en un disseny de grups paral·lels és una funció de la variància de la resposta, mentre que en un disseny amb intercanvi del tractament és funció de la variància intra-pacients ("*within*"). Per tant, l'eficiència estadística del "*cross-over*" respecte del paral·lel depèn del control de la variància entre-individus ("*between*"). Però els assaigs clínics poden també fer ús d'altres mètodes per controlar la variància entre- (com per exemple, restricció via criteris d'inclusió, l'estudi del canvi respecte del valor inicial, etc.).

L'estudi de l'eficiència del "*cross-over*" es duu a terme a partir de l'anàlisi de les variàncies de les respostes d'articles cercats a MEDLINE entre els anys 2000-2002 a les revistes mèdiques *Am J Med, Ann Intern Med, BMJ, JAMA, Lancet, N Engl J Med.* Altres criteris de selecció han estat: resposta principal contínua i mètode d'anàlisi basat en la distribució Normal (t-test, ANOVA i modelització).

D'aquesta cerca resulten 50 articles, 21 compleixen tots els criteris d'inclusió dels quals 12 s'han beneficiat del disseny "*cross-over*" i permeten efectuar els càlculs dels paràmetres d'interès: fiabilitat (R; quocient entre la variància entre i la variància total), Relació de tamanys mostrals (o raó de mostres, '*rn*', (1-R)/2), Reducció en la grandària de la mostra (SSR, complementari de '*rn*') i, finalment, l'eficiència relativa (que és el quocient entre els costos del disseny paral·lel i "*cross-over*", assumint com a unitari el cost del període i, resultant la inversa del complementari de la fiabilitat, $1/(1-R)$). Els 9 restants no proporcionen la informació adient perquè no analitzen les dades d'acord amb el model complert eliminant la variabilitat entre-individus i per tant, s'estableix que la seva eficiència és nul·la en quant a presentació dels resultats.

Dels articles que han permès efectuar els càlculs, un 66.63% (8 de 12 articles) posen de manifest una important reducció del cost del disseny "*cross-over*" respecte del disseny de grups paral·lels (la resta presenta valors d'eficiència inferiors a 3). Valors de l'eficiència propers a la unitat serien indicadors de que la planificació d'un disseny amb intercanvi no suposa un guany substancial i l'elecció d'un disseny o altre hauria de dependre d'altres factors aliens a les seves propietats.

ÍNDEX

1.- INTRODUCCIÓ

En les diferents fases del seu desenvolupament, medicaments i d'altres tractaments per a malalties són comparats via els "*randomized controlled clinical trials*"; assaigs en què s'estudia, sobre unes variables o respostes d'interès, l'efecte de dues o més intervencions (tractaments) assignades segons un ordre aleatori a una sèrie d'individus humans. Als "*Placebo-controlled trials*", una de les intervencions és una "*dummy*" o tractament inactiu per proporcionar un emmascarament al tractament actiu en la seva comparació amb la referència, que hauria de ser el procediment estàndard fins al moment. Normalment, es refereix l'efecte al tractament actiu quan en realitat el que s'estudia és la diferència entre els efectes del tractament i de la referència (els efectes purs del tractament actiu i del placebo no són identificables per separat). Pertanyen a aquest tipus de dissenys el "*paral·lel group trial*" (cada grup d'individus s'exposa únicament a una intervenció) i el "*cross-over trial*" (cada individu rep totes les intervencions d'estudi en successius períodes); que són els dos disseny que es pretén examinar en aquest estudi.

L'objectiu del projecte és estudiar el benefici en la planificació d'un disseny "*cross-over*" comparat amb el disseny de grups paral·lels en quant a eficiència relativa, que es defineix com el quocient entre la grandària de mostra d'un paral·lel i la d'un disseny "*cross-over*" o, equivalentment, el quocient de les variàncies de les respostes respectives.

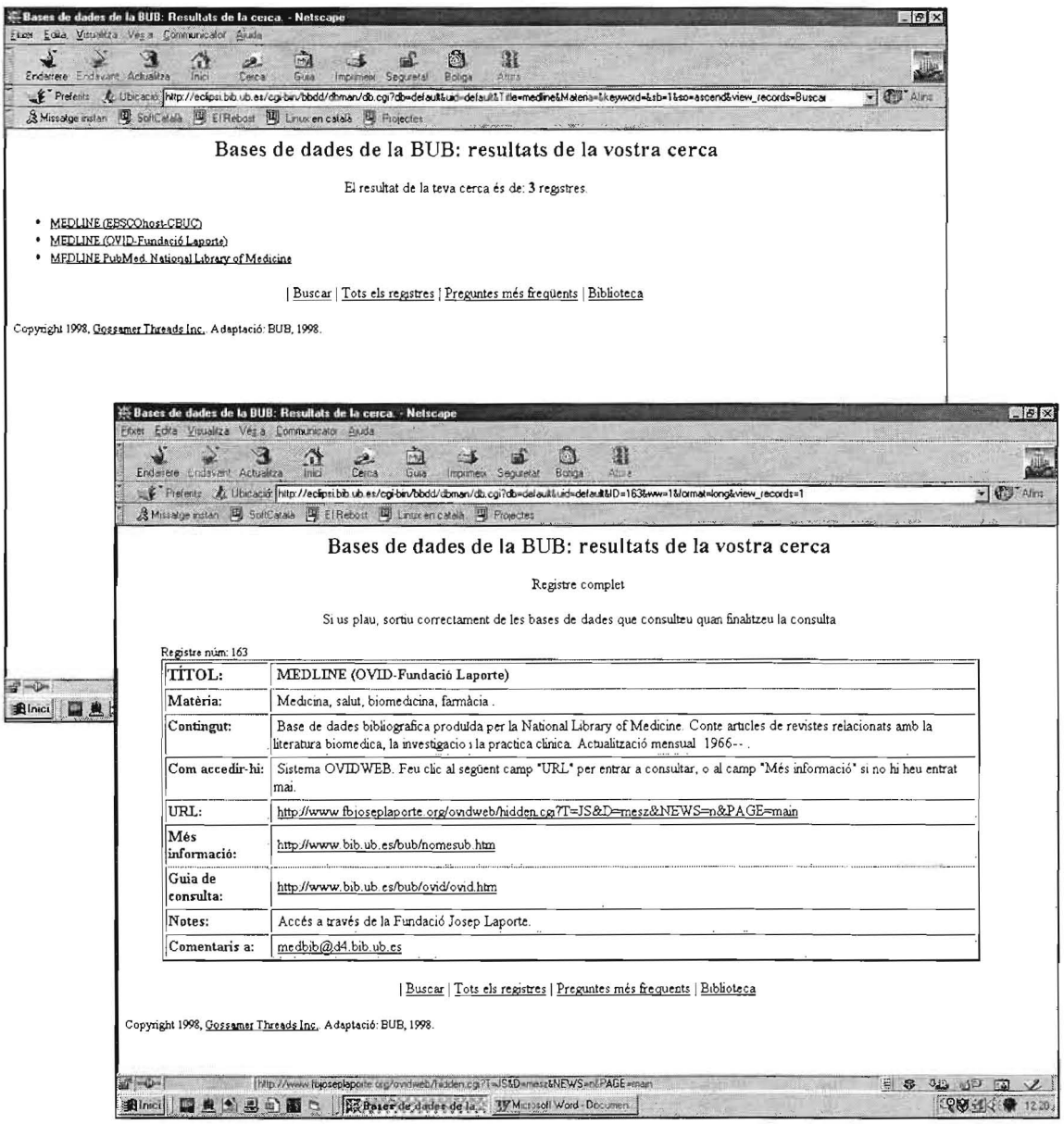
La informació s'extreu a partir d'articles publicats entre els anys 2000 i 2002 a revistes mèdiques amb alt factor d'impacte, concretament a "*The American Journal of Medicine*", "*Annals of Internal Medicine*", "*British Medical Journal*", "*The Lancet*", "*The Journal of the American Medical Association*" i "*The New England Journal of Medicine*" amb la intenció de participar al "*3rd JOINT MEETING International Society for Clinical Biostatistics*", que es portarà a terme del 20 al 24 de juliol de 2003 a Londres i on l'han acceptat per exposar-ne un pòster.

El cos del projecte consta d'una introducció teòrica a l'anàlisi dels dos tipus de dissenys, de l'obtenció de la mostra de casos i dels resultats cercats. Per poder realitzar l'estudi, és necessari disposar de certa informació que no es proporciona en tots els articles; de manera que ha calgut mantenir contacte (via mail o telefònic) amb els seus autors i, a hores d'ara, no s'ha rebut gran part de la informació mancant.

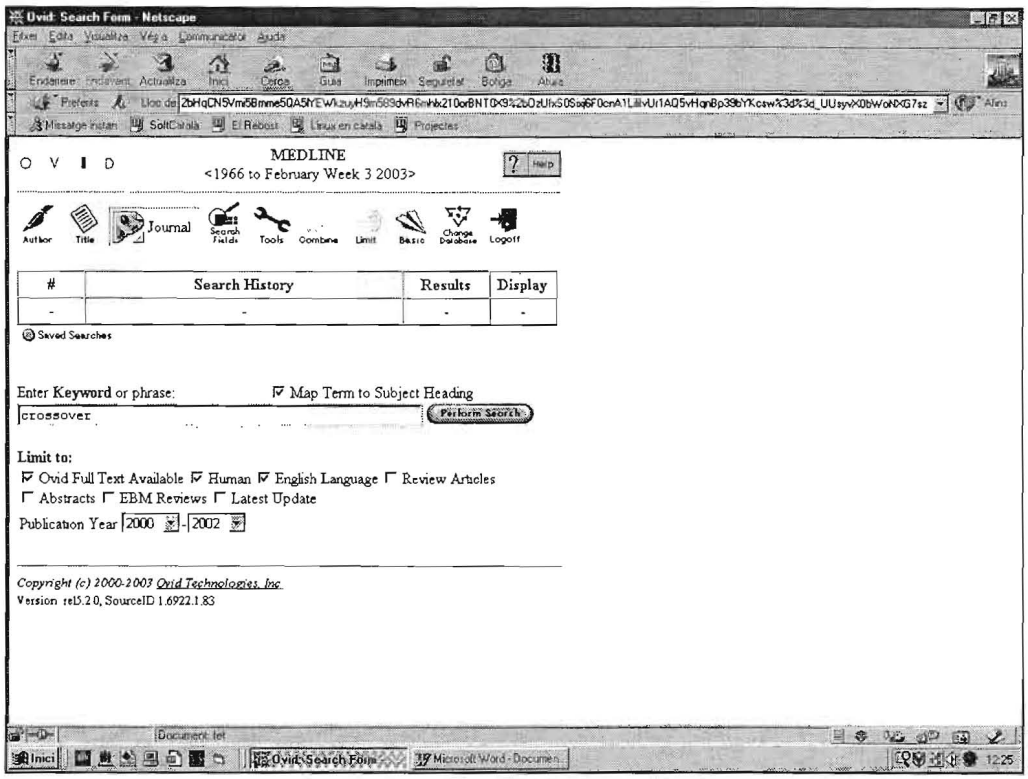
2.- MATERIAL I MÈTODES

2.1.- OBTENCIÓ DE LES DADES

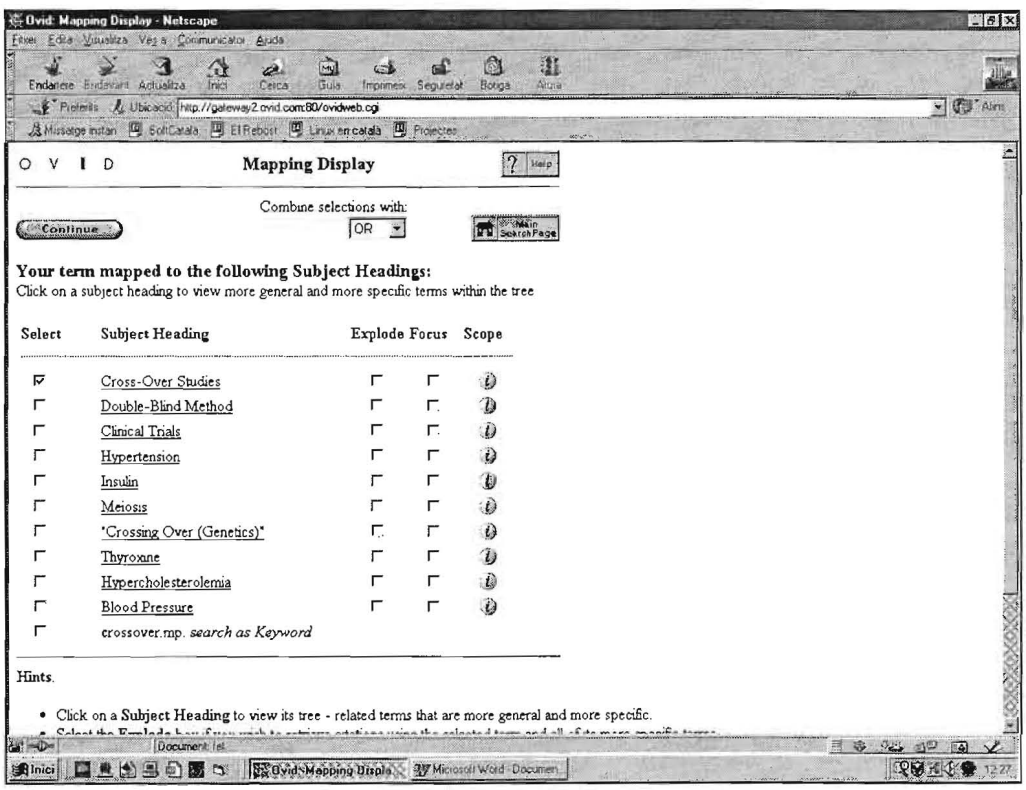
La recerca d'articles s'ha dut a terme consultant la base de dades mèdiques *MEDLINE*, en la segona opció *OVID.- Fundació Laporte*; accedint des de l'adreça *URL*:

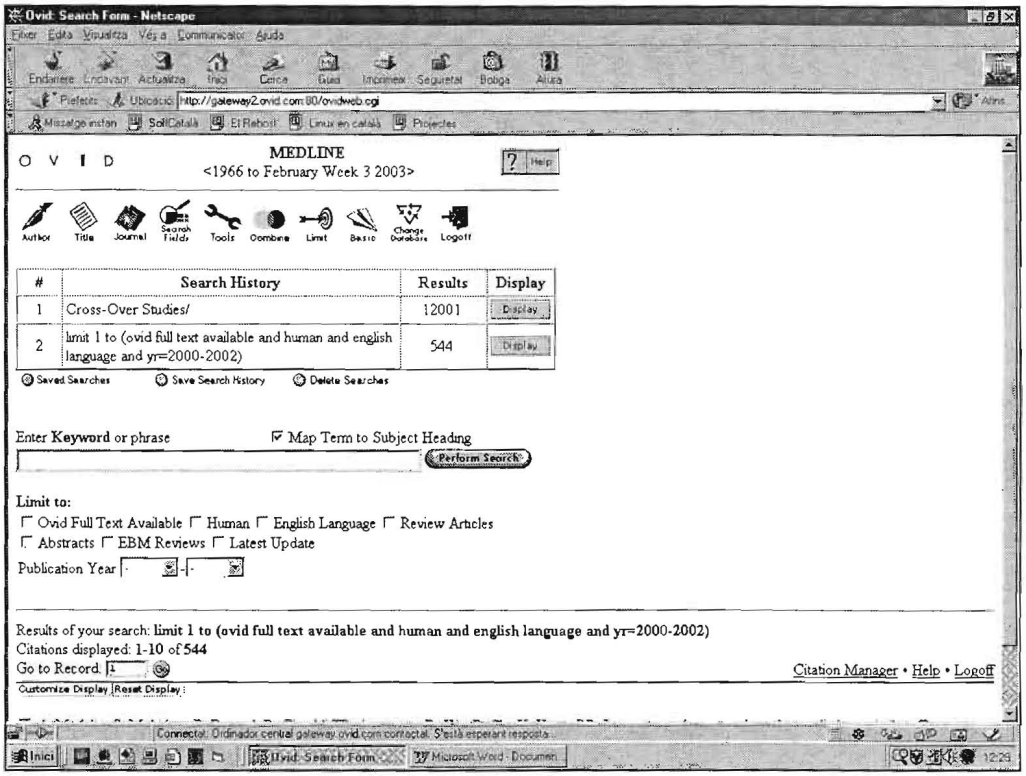


Seguidament, es limita la cerca als articles que tractin d'assaigs clínics amb humans publicats entre el 2000 i 2002 en anglès i amb accés a tot l'article:



Es selecciona el *Subject Heading Cross-Over Studies* i es prem *Continue*, obtenint finalment 544 articles.





La inclusió dels articles que formen part de l'estudi es realitza seguint els següent criteri:

- Es limita la cerca a les revistes The American Journal of Medicine (Am J Med), Annals of Internal Medicine (Ann Intern Med), British Medical Journal (BMJ), The Lancet (Lancet), The Journal of the American Medical Association (JAMA) i The New England Journal of Medicine (N Engl J Med) per ser revistes prestigioses amb elevat factor d'impacte.
- Variable principal ("*main outcome*") contínua.
- Mètode d'anàlisi ("*Statistical analysis*") basat en la teoria de la Normalitat (t-test, ANOVA, General Lineal Model...)

Les pàgines web de les revistes mèdiques citades són www.amjmed.com (Am J Med), www.annals.org (Ann Intern Med), www.bmj.com (BMJ), www.thelancet.org (Lancet), www.jama.com (JAMA) i www.nejm.org (N Engl J Med).

2.2.- DISSENY DE GRUPS PARAL·LELS

Els "*parallel group trials*" són dissenys que es planifiquen amb l'objectiu d'estudiar l'efecte de dos o més tractaments (intervencions) assignats aleatòriament a una sèrie d'individus amb la particularitat que cada grup d'individus s'exposa únicament a una intervenció.

Els individus són aleatòriament dividits en grups i a cada grup s'assigna a l'atzar un dels tractaments a comparar. Les estimacions de les diferències entre tractaments s'obtenen de les comparacions entre grups d'individus, és a dir, es basen en la informació entre-individus.

Assumint com a premissa que la variable d'interès es distribueix Normalment, el model per a l'observació de l'individu *i-èssim* sota el tractament *j* y_{ij}, es modelarà segons:

$$y_{ij} = E[y_{ij}] + \tau_i + \varepsilon_{ij} = \mu + \alpha_j + \tau_i + \varepsilon_{ij}$$

on

μ és la mitjana general,

α_j és l'efecte fix del tractament, $_{j=1,2,...k}$

τ_i és l'efecte aleatori degut a les diferències entre els individus, $_{i=1,2...n^*}$

$\tau \sim N(0, \sigma_\tau^2)$; σ_τ^2 és la variabilitat inter-individual (σ_B^2 ; *Between Patient Variation*)

ε_{ij} és l'efecte aleatori degut a l'individu

$\varepsilon \sim N(0, \sigma_e^2)$; σ_e^2 és la variabilitat intra-individual (σ_W^2 ; *Within Patient Variation*)

i, per tant, la variància de la resposta serà: $Var(Y_P) = Var(y_{ij}) = \sigma_\tau^2 + \sigma_e^2$

2.3.- DISSENY "CROSS-OVER"

El disseny "*cross-over*" és adequat per investigar tractaments per a enfermetats contínues o cròniques (asma, reumatisme, migranya, hipertensió, epilèpsia...); és a dir, per a condicions en què no es qüestiona la cura del problema subjacent que causa l'enfermetat sinó que s'espera moderar el seu efecte amb el tractament. També son útils per a quantificar les concentracions de fàrmac en els diferents compartiments del cos humà (sang, líquid cefaloraquídic,...), estudis coneguts com a farmacocinètica.

Comparat amb el disseny de grups paral·lels, el disseny amb intercanvi del tractament presenta una sèrie de desavantatges: el problema de les pèrdues ("drop-outs") dificulta l'anàlisi i la interpretació i, si bé en el paral·lel el temps fins que l'individu abandona l'estudi pot proporcionar informació que es pot recuperar, en un "*cross-over*" és bastant més difícil sobretot si la pèrdua es produeix en el primer període. D'altra banda, el "*cross-over*" no és aconsellable quan el pacient pot experimentar alguna millora o empitjorament durant el curs del tractament, a més, l'individu s'ha de sotmetre a varis tractaments i aleshores requereix un període d'observació que deu ser més llarg.

Un inconvenient afegit del disseny amb intercanvi respecte del disseny de grups paral·lels és la possible existència del fenòmen conegut com "*carry-over*" (efecte residual o arrossegat del tractament), que es defineix com la persistència o bé física o bé dels seus efectes d'un tractament aplicat en un període previ. La principal conseqüència d'aquest fenòmen seria el biaix en l'estimació de l'efecte del tractament: si, per exemple, persisteix l'efecte curatiu del primer tractament subministrat, al final del segon tractament pot semblar que és el darrer el que beneficia el pacient quan en realitat és el primer el responsable. A més pot passar que, per alguna raó, les condicions presents en els diferents períodes siguin tals que les diferències entre tractaments depenguin del període en que són subministrats (l'efecte del tractament no és constant al llarg del temps). Ens referirem a la interacció entre període i tractament.

Algunes aproximacions per al problema del "*carry-over*" són:

1) "*two-stage procedure*" que consisteix en una sèrie de tests estadístics per examinar la possibilitat de que el fenòmen de l'efecte arrossegat del tractament hagi ocorregut
2) incloure paràmetres pel "*carry-over*"[1] i estimar el tractament i el "*carry-over*" simultàniament, o

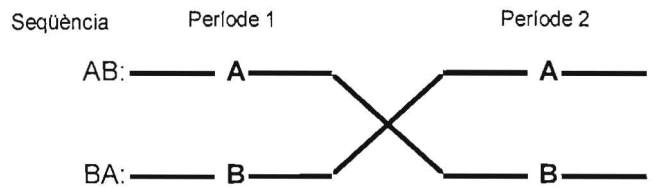
1 Veure apartat 2.3.1

3) deixar transcórrer un interval de temps entre els períodes de tractaments, conegut com a "*wash-out*", durant el qual es pressuposa que l'efecte del tractament subministrat abans desapareix [si no es proporciona cap tractament, el "*wash-out*" és passiu i s'assumeix que el pacient retorna a l'estat natural abans que el següent tractament s'inicií; sinó, el "*wash-out*" és actiu].

Per tal de clarificar a quin tipus d'efecte del tractament es fa referència en cada cas, '**l'efecte directe del tractament**' serà l'efecte que té el tractament durant el període que és subministrat; de manera que l'efecte del tractament que persisteix després del període de tractament serà '**l'efecte arrossegat del tractament**'.

En un "*cross-over*" els casos són assignats a seqüències de tractaments amb l'objectiu d'estudiar les diferències individuals entre tractaments (o subseqüències d'aquests). Cada individu experimental rep dos o més tractaments en un ordre que dependrà del disseny particular escollit.

El disseny més senzill és aquell en què dos tractaments (etiquetats convencionalment com A i B) són comparats i cada individu rep ambdós tractaments en una de les seqüències A,B o B,A en dos períodes successius.



El patró del disseny seria:

Seqüència ò ordre	Individu	Període 1		Període 2	
		A	B	A	B
1: AB	1 2 3 ... n	1 2 3 ... n			1 2 3 ... n
2:BA	n+1 n+2 n+3 ... n+m		n+1 n+2 n+3 ... n+m	n+1 n+2 n+3 ... n+m	

A l'Annex 6.3, es reprodueix el patró de disseny anterior aplicat a les dades.

2.3.1.- NOTACIÓ I MODELS

S'assumirà que hi ha *s* grups diferents d'individus; cada grup rep *t* tractaments en diferent ordre durant el transcurs de *p* períodes de tractament. Per exemple, per a un disseny que compari tres tractaments usant tres períodes hauria d'haver sis grups d'individus: ABC, ACB, BAC, BCA, CAB, CBA de manera que els individus de cada grup rebrien els tractaments en l'ordre especificat pel grup. Sigui y_{ijk} la resposta observada en l'individu *k-èssim* del grup *i* en el període *j*, suposarem y_{ijk} l'observació de la variable aleatòria Y_{ijk} que pot ser modelada per:

$$Y_{ijk}=\mu+\tau_{ik}+\pi_j+\alpha_{d[i,j]}+\lambda_{d[i,j-1]}+\varepsilon_{ijk}$$

Els termes que apareixen al model corresponen a:

μ: Mitjana general

τ_{ik}: efecte aleatori de l'individu *k* en el grup *i*, i=1, 2, ..., *s*; k=1, 2, ..., n_i. S'assumeix independent i idènticament distribuit (i.i.d.) segons $N(0,\sigma_\tau^2)$; σ_τ^2 és la variabilitat inter-individual (σ_B^2 ; "*Between patient variation*")

π_j: Efecte fix del període j, j=1, 2, ..., *p*

$\alpha_{d[i,j]}$: efecte fix directe del tractament administrat al grup *i* al període *j*

$\lambda_{d[i,j-1]}$: efecte fix del "carry-over" del tractament rebut al període *j*-1 del grup *i*, on $\lambda_{d[i,0]}=0$

ε_{ijk}: efecte aleatori de l'individu *k* en el període *j* en el grup *i*. Els errors e_{ijk} s'assumeixen independents, identicament distribuits segons $\varepsilon \sim N(0,\sigma_e^2)$; σ_e^2 és la variabilitat intra-individual (σ_W^2 ; "*Within patient variation*").

En el disseny "*2x2 cross-over*", el model lineal inclou els paràmetres per ajustar la possible diferència entre els efectes del "*carry-over*", si existeix, però no el paràmetre per a la possible interacció entre període i tractament (que pot ocórrer si la diferència entre els efectes directes del tractament no és la mateixa en ambdós períodes). La raó per la qual s'omet el paràmetre de la interacció entre el període i el tractament és que a l'haver-hi només quatre mitjanes mostrals, $\bar{y}_{AB,1}$, $\bar{y}_{AB,2}$, $\bar{y}_{BA,1}$ i $\bar{y}_{BA,2}$, només es poden incloure com a màxim tres paràmetres (graus de llibertat) per a respondre sobre les diferències entre les mitjanes: dos graus de llibertat s'associen a les diferències entre els períodes i l'efecte directe dels tractaments, deixant només un associat al "carry-over" i a la interacció període-tractament, que tenen el mateix àlias i, per tant, resten confosos[2].

2 veure Annex 6.4

La principal característica que diferencia els "*cross-over*" d'altres dissenys que comparen tractaments és que aquesta comparació s'efectua a nivell 'intra-individual', és a dir, cada individu proporciona una comparació directa entre els tractaments que ha rebut. Per a contrastar l'efecte dels tractaments, les dades es poden analitzar en forma de diferència versus el tractament control ("*cross-over differences*" i/o "*period differences*"; cada individu actúa com el seu propi control) i aquesta diferència elimina la font de variabilitat deguda a les diferències entre-individus.

Tenint present el model anunciat per a una observació individual i assumint que no es manifesta el fenòmen del "*carry-over*" ($\lambda_A=\lambda_B=0$):

$$Y_{\text{cross-over differences}} = \begin{cases} Y_{AB1k} - Y_{AB2k} = (\alpha_A - \alpha_B) + (\pi_1 - \pi_2) + (\varepsilon_{AB1k} - \varepsilon_{AB2k}) \\ Y_{BA2k} - Y_{BA1k} = (\alpha_A - \alpha_B) + (\pi_2 - \pi_1) + (\varepsilon_{BA2k} - \varepsilon_{BA1k}) \end{cases}$$

$$Y_{\text{period differences}} = \begin{cases} Y_{AB1k} - Y_{AB2k} = (\alpha_A - \alpha_B) + (\pi_1 - \pi_2) + (\varepsilon_{AB1k} - \varepsilon_{AB2k}) \\ Y_{BA1k} - Y_{BA2k} = (\alpha_B - \alpha_A) + (\pi_1 - \pi_2) + (\varepsilon_{BA1k} - \varepsilon_{BA2k}) \end{cases}$$

amb esperances:

$$E[Y_{\text{cross-over differences}}] = \begin{cases} E[Y_{AB1k} - Y_{AB2k}] = (\alpha_A - \alpha_B) + (\pi_1 - \pi_2) \\ E[Y_{BA2k} - Y_{BA1k}] = (\alpha_A - \alpha_B) + (\pi_2 - \pi_1) \end{cases}$$

$$E[Y_{\text{period differences}}] = \begin{cases} E[Y_{AB1k} - Y_{AB2k}] = (\alpha_A - \alpha_B) + (\pi_1 - \pi_2) \\ E[Y_{BA1k} - Y_{BA2k}] = (\alpha_B - \alpha_A) + (\pi_1 - \pi_2) \end{cases}$$

i, per tant, la variància d'aquestes variables serà:

$$Var(Y_{C-O}) = 2\sigma_e^2$$

2.4.- ALTRES MÈTODES QUE CONTROLEN σ^2_B

L'avantatge teòric del disseny "*cross-over*" és la seva major eficiència estadística comparat amb els dissenys de grups paral·lels, degut al control de la variància entre-individus. Si es disposa de mesures prèvies de l'individu que estiguin relacionades amb la resposta en estudi, altres eines metodològiques com, per exemple, la restricció via criteris d'inclusió, l'anàlisi del canvi respecte dels valors inicials, l'estratificació o el modelatge, entre d'altres, també permeten controlar aquesta font de variabilitat.

L'ajust pels valors basals pretén baixar la variància residual i amb això assumir els objectius: a) augmentar la potència i b) disminuir els biaixos potencials. Quan, per exemple, s'analitza el canvi respecte dels valors a l'inici de l'estudi (Z) abans de prendre el tractament, la resposta és també, com pel disseny "*cross-over*", una diferència però en aquest cas:

$$Y_{CANVI} = Y_{ij} - Z_{ij}$$

I la seva variància correspondrà a:

$$Var(Y_{CANVI}) = \sigma^2_Y + \sigma^2_Z - 2\sigma_{YZ} = \sigma^2_Y + \sigma^2_Z - 2\rho_{YZ}\sigma_Y\sigma_Z$$

Si es fa servir el mètode del modelatge,

$$Y_{ij} = \mu + \phi X_{ij} + \alpha_i + e_{ij} \ ,$$

la variabilitat del terme d'error serà $V(e) = \sigma^2_Y(1-R^2)$, on R^2 és el coeficient de determinació (que per a la regressió lineal simple coincideix amb la correlació entre les dues variables, Y i X) i $(1-R^2)$ és la part de la variabilitat de la resposta no explicada pels termes inclosos al model. L'eficiència d'aquest mètode és, per tant, com a mínim la mateixa que la del dissenys de grups paral·lels. Notar, a més, que l'estudi del canvi és un cas particular del modelatge si la covariable és la mesura basal i $\mu=0$ i $\phi=1$.

En l'apartat 6.4 es parla de l'ajust per la mesura basal i/o per covariables quan el disseny que es duu a terme per contrastar l'efecte dels tractaments és el "*cross-over*".

2.5.- EFICIÈNCIA D'UN DISSENY

Generalment, la grandària de la mostra en un estudi comparatiu de dos grups de pacients, assumint que la variància de la resposta en els dos grups és la mateixa, σ^2, es calcularia fent ús de la següent fòrmula:

$$n^* = \frac{2\sigma^2(z_{\alpha/2} + z_\beta)^2}{\delta^2}$$

suposant δ el veritable valor de la diferència de mitjanes de la resposta ($\delta=\mu_1-\mu_2$), tenint en compte que α és el nivell de significació del test que compara les dues mitjanes (la diferència és significativa si el valor absolut de l'estadístic *t* supera el valor de $z_{\alpha/2}$), la potència del test és $1-\beta$ (z_β és el valor de la distribució normal que talla una proporció β en la cua superior).

En un model lineal, $X=T+e$, X és el valor observat d'una variable per a un individu, T representa la mitjana de vàries hipotètiques repeticions de la mesura sobre aquest individu i es distribueix al voltant d'una mitjana μ amb variància σ_τ^2 i, e és l'error comès o diferència entre el valors X i T; amb mitjana zero. Assumint que la distribució dels errors és independent del valor de T, e té variància σ_e^2 qualsevol que sigui el valor de T i, per tant, la variància de la variable X es pot descomposar en:

$$\sigma_X^2 = \sigma_\tau^2 + \sigma_e^2$$

Quan es realitzen una sèrie de mesures a diferents individus, una mesura que expressa la magnitud relativa de les dues components de la variància d'X és el "*intraclass correlation coefficient of reliability*" (*"reliability"*):

$$R = \frac{\sigma_\tau^2}{\sigma_\tau^2 + \sigma_e^2}$$

com s'ha vist, σ_τ^2 és la variància entre pacients i σ_e^2 és la variància intra pacients.

El paràmetre R augmenta quan el quocient σ_e^2/σ_τ^2 disminueix, és a dir, quan decreix l'error en proporció respecte del que s'observa i viceversa, mantenint-se en un rang de variación entre 0 i 1. R és directament interpretable com la proporció de variància deguda a la variabilitat entre pacients i la majoria d'inconvenients que

presenta la falta de fiabilitat d'un estudi són funció d'R, com és el cas de l'increment de la grandària de la mostra 'n' que s'ha vist que depèn de la variància de la resposta.

La principal raó per què el disseny "*cross-over*" es prefereix al disseny de grups paral·lels és perquè per obtenir la mateixa precisió en les estimacions són necessaris un menor nombre de pacients. Si el veritable valor de la diferència entre les mitjanes de dos tractaments, A i B, fos de δ unitats, establerts els valors d'α i β, la grandària de la mostra pel cas del disseny de grups paral·lels es calcularia:

$$n^* = \frac{2\left(\sigma_W^2 + \sigma_B^2\right)\left(z_{\alpha/2} + z_\beta\right)^2}{\delta^2}$$

mentre que planificant l'estudi amb un "*cross-over*":

$$n = \frac{\sigma_W^2 \left(z_{\alpha/2} + z_\beta\right)^2}{\delta^2}$$

pel que la relació resultant entre ambdós tamanys de mostra és: $n^* = \frac{2n}{(1-R)}$.

Assumint com a unitari el cost del període, és a dir, té el mateix cost mesurar un individu dues vegades que mesurar dos individus diferents i, per tant, el cost del disseny "*cross-over*" seria dues vegades el del paral·lel, definim:

- la 'Relació de les Grandàries Mostrals' o 'Raó de Mostres' ('rn') com el quocient entre n i n^*:

$$rn = \frac{(1-R)}{2}$$

- el "*Sample Size Reduction*" (SSR) com:

$$\mathrm{SSR} = \frac{\mathrm{n}^* - \mathrm{n}}{\mathrm{n}^*} = 1 - rn$$

- finalment, l'eficiència relativa del "*cross-over*" respecte al disseny de grup paral·lels es calcularia com:

$$\textit{Eficiència} = \frac{\cos t_{paral\cdot lel}}{\cos t_{cross-over}} = \frac{F(n^*)}{F(n = 2n^*)} = \frac{1}{2 \cdot rn} = \frac{\sigma_\tau^2 + \sigma_e^2}{\sigma_e^2} = \frac{1}{(1-R)}$$

2.6.- EXTRACCIÓ DE LA INFORMACIÓ

Per tal de contrastar les hipòtesis dels tests a efectuar, caldrà assumir addicionalment que les dades observades són una mostra aleatòria d'una variable amb distribució normal (Gaussiana). A continuació, ens referirem a la forma de procedir a analitzar les dades en un disseny "*2x2 cross-over*" via t-test i anàlisi de la variància.

2.6.1.- ANÀLISI VIA t-TEST

El mode de proporcionar els resultats que posen a prova l'efecte tractament és via mitjana, desviació estàndard i el corresponent interval de confiança de les variables "*period differences*" i/o "*cross-over differences*".

El procediment emprat quan es compleix l'assumpció de que la variable "*cross-over differences*" es distribueixen aleatòriament al voltant del veritable efecte del tractament és el "**INDEPENDENT t-TEST BASED ON MATCHED-PAIRS DIFFERENCES**"[3].

Els factors que poden causar que la variable "*cross-over differences*" no estigui distribuida aleatòriament al voltant del veritable efecte són:

1) L'efecte PERÍODE: existeix una tendència general afectant a tot l'experiment que faci, per exemple, que els valors observats en el segon període de tractament siguin superiors als del primer.
2) La INTERACCIÓ entre el PERIODE i el TRACTAMENT i/o CARRY-OVER.
3) La INTERACCIÓ entre el PACIENT i el TRACTAMENT: l'efecte del tractament és diferent per a cada individu.
4) La INTERACCIÓ entre el PACIENT i el PERÍODE es dóna quan els individus estan subjectes a una tendència que no és igual per a cadascun.

És necessari assumir que no existeixen les interaccions entre pacient-tractament i pacient-període perquè sinó el tractament no tindria interès pràctic; els models que les ajusten haurien de ser "*cross-over*" en què cada pacient a cada període rebés el mateix tractament vàries vegades (rèpliques). Si existeix algun dels efectes període o efecte arrossegat del tractament, el model lineal que s'assumeix és l'exposat en l'apartat 2.3.1 i els efectes fixos del model complert del "*2x2 cross-over*" associats a cada individu en cada període i en cada seqüència són:

3 Senn l'anomena "*matched-pairs t-test*" o "*correlated t-test*". Veure Annex 6.5

Seqüència	Període 1	Període 2
1 (AB)	$\mu+\alpha_1+\pi_1$	$\mu+\alpha_2+\pi_2+\lambda_1$
2 (BA)	$\mu+\alpha_2+\pi_1$	$\mu+\alpha_1+\pi_2+\lambda_2$

La premissa que s'efectua és la d'igualtat dels efectes del "*carry-over*"; és a dir, **que** $\lambda_1=\lambda_2$ i que es pot validar via:

Es defineix la variable 'suma' com:

$t_{1k} = y_{11k} + y_{12k}$ per l'individu k - èssim del grup AB (1)

$t_{2k} = y_{21k} + y_{22k}$ per l'individu k - èssim del grup BA (2)

amb esperances:

$$E[t_{1k}] = 2\mu + \pi_1 + \pi_2 + \alpha_1 + \alpha_2 + \lambda_1$$
$$E[t_{2k}] = 2\mu + \pi_1 + \pi_2 + \alpha_1 + \alpha_2 + \lambda_2$$

Si $\lambda_1=\lambda_2$, les esperances per als dos grups són iguals. Conseqüentment, per contrastar $\lambda_1=\lambda_2$ es pot aplicar el two-sample t-test a la 'suma'.

Si definim $\lambda_d=\lambda_1-\lambda_2$ i $\hat{\lambda}_d = \bar{t}_{1.} - \bar{t}_{2.}$, aleshores:

$$E[\hat{\lambda}_d] = \lambda_d$$
$$V[\hat{\lambda}_d] = 2\left(2\sigma_\tau^2 + \sigma_e^2\right)\left(\frac{1}{n_1}+\frac{1}{n_2}\right) = \sigma_T^2 m$$
$$\text{on} \quad \sigma_T^2 = 2\left(2\sigma_\tau^2 + \sigma_e^2\right) \qquad m = \frac{n_1+n_2}{n_1 n_2}$$

Per estimar σ_T^2, es calcula:

$$S_T^2 = \sum_{i=1}^{2}\sum_{k=1}^{n_i}\left(t_{ik} - \bar{t}_{i.}\right)^2 / (n_1+n_2-2)$$

amb n_1+n_2-2 graus de llibertat. Sota H_0, l'estadístic

$$t_\lambda = \frac{\hat{\lambda}_d}{\sqrt{S_T^2 m}}$$

es distribueix segons una t-Student de (n_1+n_2-2) graus de llibertat.

Com es tracta d'un test preliminar, anterior al test per a la diferència d'efectes dels tractaments, és usual testejar la hipòtesi nul·la amb un nivell de significació del 10% enlloc de l'habitual 5%. (També es baixa el 1-α per compensar el descens en 1-β degut a la major variància de la variable 'suma').

Assumint $\lambda_1=\lambda_2$, es procedeeix a contrastar la hipòtesi nu·la d'igualtat d'efectes **di**rectes, $H_0:\alpha_1=\alpha_2$, via el procediment "**TWO-SAMPLE t-TEST APPROACH**".

Si s'assumeix que $\lambda_1=\lambda_2$, llavors la variable "*period difference*":

$$d_{1k} = y_{11k} - y_{12k} \qquad \text{per l'individu k - èssim del grup AB (1)}$$
$$d_{2k} = y_{21k} - y_{22k} \qquad \text{per l'individu k - èssim del grup BA (2)}$$

té esperança:

$$E[d_{1k}] = \pi_1 - \pi_2 + \alpha_1 - \alpha_2$$
$$E[d_{2k}] = \pi_1 - \pi_2 + \alpha_2 - \alpha_1$$

Sota la hipòtesi nul·la de $\alpha_1=\alpha_2$, les esperances per als dos grups són iguals i, per tant, es pot aplicar el two-sample t-test al "*period differences*".

Si definim $\alpha_d=\alpha_1-\alpha_2$ i $\hat{\alpha}_d = \frac{1}{2}\left[\bar{d}_{1.} - \bar{d}_{2.}\right]$, aleshores:

$$E[\hat{\alpha}_d] = \alpha_d$$
$$V[\hat{\alpha}_d] = \frac{2\sigma_e^2}{4}\left(\frac{1}{n_1}+\frac{1}{n_2}\right) = \frac{\sigma_D^2}{4}m$$
$$\text{on} \qquad \sigma_D^2 = 2\sigma_e^2$$

L'estimació ponderada de σ_D^2 és:

$$S_D^2 = \sum_{i=1}^{2}\sum_{k=1}^{n_i}\left(d_{ik} - \bar{d}_{i.}\right)^2 / (n_1 + n_2 - 2)$$

Sota H_0, l'estadístic:

$$t_\alpha = \frac{\hat{\alpha}_d}{\sqrt{S_D^2 m/4}}$$

es distribueix segons una t-Student de (n_1+n_2-2) graus de llibertat.

L'interval de confiança al 95% per al veritable efecte del tractament, $\alpha=\alpha_1-\alpha_2$ és:

$$\hat{\alpha}_d \pm t_{n_1+n_2-2,\alpha/2}\left(\frac{mS_D^2}{4}\right)^{1/2}$$
$$\frac{1}{2}\left[\bar{d}_{1.} - \bar{d}_{2.}\right] \pm t_{n_1+n_2-2,\alpha/2}\left(\frac{1}{2}S_D\right)\sqrt{\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

I *S* serà la mitjana de:

$$S_{UCI} = \frac{2\left(UCI - \frac{1}{2}\left[\bar{d}_{1.} - \bar{d}_{2.}\right]\right)}{t_{n_1+n_2-2,\alpha/2}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \qquad S_{LCI} = \frac{2\left(\frac{1}{2}\left[\bar{d}_{1.} - \bar{d}_{2.}\right] - LCI\right)}{t_{n_1+n_2-2,\alpha/2}\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

La justificació del factor $\frac{1}{2}$ es troba en el següent raonament: si ens fixem en l'esperança matemàtica de la variable "*period differences*"

$$E(d_{1i}) = E(Y_{A1i} - Y_{B2i}) = [\mu + \alpha_A + \pi_1] - [\mu + \alpha_B + \pi_2 + \lambda_A] = (\alpha_A - \alpha_B) + (\pi_1 - \pi_2) - \lambda_A$$
$$E(d_{2i}) = E(Y_{B1i} - Y_{A2i}) = [\mu + \alpha_B + \pi_1] - [\mu + \alpha_A + \pi_2 + \lambda_B] = (\alpha_B - \alpha_A) + (\pi_1 - \pi_2) - \lambda_B$$

L'esperança de la seva diferència, assumint que l'efecte residual dels tractaments no existeix ($\lambda_A=\lambda_B=0$) o són iguals ($\lambda_A=\lambda_B$), serà:

$$E(d_{1i} - d_{2i}) = E(d_{1i}) - E(d_{2i}) = 2(\alpha_A - \alpha_B)$$

És a dir, degut a que cada individu aporta informació sobre els dos tractaments, l'esperança de la diferència de les d_i és el doble dels efectes directes d'ambdós tractaments i, per tant, per estimar l'efecte diferencial del tractament cal dividir per dos els resultats. Es tracta d'una prova de dades independents, compara d_i en ambdues seqüències que estan formades per individus diferents.

Si interessa contrastar la diferència entre els efectes dels períodes, és a dir, H_0:$\pi_1=\pi_2$ (assumint $\lambda_1=\lambda_2$) llavors:

Per tal de contrastar la igualtat dels efectes dels períodes, la variable que s'utilitza és la "*cross-over differences*":

$c_{1k} = d_{1k} = y_{11k} - y_{12k}$ per l'individu k - èssim del grup AB (1)

$c_{2k} = -d_{2k} = y_{22k} - y_{21k}$ per l'individu k - èssim del grup BA (2)

amb esperances:

$$E[c_{1k}] = \pi_1 - \pi_2 + \alpha_1 - \alpha_2$$
$$E[c_{2k}] = \pi_1 - \pi_2 + \alpha_2 - \alpha_1$$

Sota la hipòtesi nul·la de $\pi_1=\pi_2$, les esperances per als dos grups són iguals i, per tant, es pot aplicar el two-sample t-test a la variable "*cross-over differences*".

Si $\pi_d=\pi_1-\pi_2$ i $\hat{\pi}_d = \frac{1}{2}[\bar{c}_{1.} - \bar{c}_{2.}]$, sota la hipòtesi nul·la:

$$t_\pi = \frac{\hat{\pi}_d}{\sqrt{S_D^2 m/4}}$$

es distribueix segons una t-Student de (n_1+n_2-2) graus de llibertat.

Si hi ha evidència de que no es verifica la premissa $\lambda_1=\lambda_2$, no es pot procedir a contrastar $\alpha_1=\alpha_2$ ni $\pi_1=\pi_2$, donat que si $\lambda_d=\lambda_1-\lambda_2\neq 0$ llavors:

$$E[\hat{\alpha}_d] = E\left[\frac{1}{2}(\bar{d}_{1.} - \bar{d}_{2.})\right] = \alpha_d - \frac{\lambda_d}{2}$$

i, $\hat{\alpha}_d$ és un estimador esbiaixat de α_d si $\lambda_d\neq 0$.

Si notem: $\hat{\lambda}_d = \bar{y}_{11.} + \bar{y}_{12.} - \bar{y}_{21.} - \bar{y}_{22.}$ i $\hat{\alpha}_d = \frac{1}{2}[\bar{y}_{11.} - \bar{y}_{12.} - \bar{y}_{21.} + \bar{y}_{22.}]$ es pot veure que un estimador d'α_d, quan $\lambda_d\neq 0$ és:

$$\hat{\alpha}_d | \lambda_d = \frac{1}{2}[\bar{y}_{11.} - \bar{y}_{12.} - \bar{y}_{21.} + \bar{y}_{22.}] + \frac{1}{2}[\bar{y}_{11.} + \bar{y}_{12.} - \bar{y}_{21.} - \bar{y}_{22.}] = \bar{y}_{11.} - \bar{y}_{21.}$$

que és la diferència entre els grups en termes de les mitjanes del primer període; amb variància $V(\hat{\alpha}_d|\lambda_d)=m(\sigma_\tau^2+\sigma_e^2)$. És a dir, si $\lambda_d\neq 0$ l'estimador de α_d està basat en la informació entre-individus i és l'estimador que s'hagués obtingut si l'experiment s'hagués dissenyat com un "*parallel group trial*" amb només el primer període. Així doncs, l'existència d'un efecte arrossegat anul·la la informació aportada pel segon període i fa ineficient el disseny amb intercanvi de tractaments.

2.6.2.- ANÀLISI VIA ANOVA

El principal objectiu és ajustar models en ordre seqüencial fent ús de mínims quadrats ordinaris o, quan sigui apropiat, mínims quadrats generalitzats. Així, per exemple, partint del model citat en l'apartat 2.3.1 (Model 4) es podrien ajustar els següents:

Model 1: $Y_{ijk}=\mu+\tau_{ik}+e_{ijk}$
Model 2: $Y_{ijk}=\mu+\tau_{ik}+\pi_j+e_{ijk}$
Model 3: $Y_{ijk}=\mu+\tau_{ik}+\pi_j+\alpha_{d[i,j]}+e_{ijk}$
Model 4: $Y_{ijk}=\mu+\tau_{ik}+\pi_j+\alpha_{d[i,j]}+\lambda_{d[i,j-1]}+e_{ijk}$

Comparant la suma de quadrats de la regressió d'aquests models s'obtenen les sumes de quadrats de cada nou conjunt de paràmetres inclòs al model després d'haver ajustat la contribució de la resta de paràmetres al model. És a dir, del model 1 s'obté la Suma de Quadrats (SQ) entre-individus, del 2 l'addicional SQ del període (després d'ajustar μ i τ_{ik}), del 3 l'addicional SQ de l'efecte directe del tractament (després d'ajustar μ, τ_{ik} i π_j) i del 4 l'addicional SQ de l'efecte del carry-over (després d'ajustar μ, τ_{ik}, π_j i $\alpha_{d[i,j]}$). La SQ residual és al resultant d'ajustar el més complert dels models.

Tot i que es poden contrastar les hipòtesis d'interès via t-test, també es poden realitzar usant *F*-test obtinguts a partir de les taules d'anàlisi de la variància. La taula d'anàlisi de la variància[4] pel disseny "*2x2 cross-over*" és:

FONT DE VARIACIÓ:	g. ll.	SUMA DE QUADRATS
Entre-individus:		
"Carry-over"	1	$\frac{2n_1n_2}{(n_1+n_2)}(\bar{y}_{1..}-\bar{y}_{2..})^2$
Residual Entre-Individus	(n_1+n_2-2)	$\sum_{i=1}^{2}\sum_{k=1}^{n_i}\frac{y_{i.k}^2}{2}-\sum_{i=1}^{2}\frac{y_{i..}^2}{2n_i}$

4 Referència bibliogràfica (6)

FONT DE VARIACIÓ:	g. ll.	SUMA DE QUADRATS
Intra-individus:		
Directe dels Tractaments (ajustada pels Períodes)	1	$\frac{n_1 n_2}{2(n_1+n_2)}(\bar{y}_{11.}-\bar{y}_{12.}-\bar{y}_{21.}+\bar{y}_{22.})^2$
Períodes (ajustats pels Tractaments)	1	$\frac{n_1 n_2}{2(n_1+n_2)}(\bar{y}_{11.}-\bar{y}_{12.}+\bar{y}_{21.}-\bar{y}_{22.})^2$
Residual Intra-Individus	(n_1+n_2-2)	$\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{n_i} y_{ijk}^2-\sum_{i=1}^{2}\sum_{k=1}^{n_i}\frac{y_{i.k}^2}{2}-\sum_{i=1}^{2}\sum_{j=1}^{2}\frac{y_{ij.}^2}{n_i}+\sum_{i=1}^{2}\frac{y_{i..}^2}{2n_i}$
Total	$2(n_1+n_2)-1$	$\sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{n_i} y_{ijk}^2-\frac{y_{...}^2}{2(n_1+n_2)}$

essent les Esperances de les Quadrats Mitjos (EQM):

$$\text{EQM(Carry-over)}=\frac{2n_1n_2}{(n_1+n_2)}(\lambda_1-\lambda_2)^2+2\sigma_\tau^2+\sigma_e^2$$

$$\text{EQM(Residual Entre-Individus)}=2\sigma_\tau^2+\sigma_e^2$$

$$\text{EQM(Directe dels Tractaments)}=\frac{2n_1n_2}{(n_1+n_2)}\left[(\alpha_1-\alpha_2)-\frac{(\lambda_1-\lambda_2)}{2}\right]^2+\sigma_e^2$$

$$\text{EQM(Períodes)}=\frac{2n_1n_2}{(n_1+n_2)}\left(\pi_1-\pi_2-\frac{(\lambda_1-\lambda_2)}{2}\right)^2+\sigma_e^2$$

$$\text{EQM(Residual Intra-Individus)}=\sigma_e^2$$

A les sortides dels paquets estadístics convencionals apareix, amb les sumes de quadrats, els Quadrats Mitjos (QM) i els valors dels *F*-ratis, que són útils a l'hora de contrastar les hipòtesis $\alpha_1=\alpha_2$ i $\pi_1=\pi_2$ un cop s'hagi assumit que $\lambda_1=\lambda_2$.

Per validar la premissa $\lambda_1=\lambda_2$, es calcula l'*F*-rati:

$$FC=\frac{QM\_Carry-over}{QM\_Residual\ Entre-Individus}$$

i si es verifica que $\lambda_1=\lambda_2$, FC és un valor observat de la distribució $F_{1,(n1+n2-2)}$.

Assumint igualtat d'efectes arrossegats dels tractaments o, simplement que no existeixen, per contrastar la hipòtesi nul·la $\alpha_1=\alpha_2$ es calcula l'*F*-rati:

$$FT=\frac{QM\_Directe\ dels\ Tractaments}{QM\_Residual\ Intra-Individus}$$

Sota la hipòtesi nul·la, FT és un valor observat de la distribució $F_{1,(n1+n2-2)}$.

Per contrastar la hipòtesi nul·la $\pi_1=\pi_2$ es calcularà:

$$FP=\frac{QM\_Períodes}{QM\_Residual\ Intra-Individus}$$

que sota H_0, és un valor observat de la distribució $F_{1,(n1+n2-2)}$.

Per mostrar-ne una aplicació pràctica de l'esmentat, a continuació es mostra la sortida proporcionada pel paquet estadístic SAS de l'anàlisi de les dades facilitades pel Dr. Craig Crandall, autor de l'article "***Mechanism of Cocaine Induced Hyperthermia in Humans***" publicat a Annals of Internal Medicine, 136(11):785-91, 2002 Jun.

El model ajustat és el model complert:

$$Y_{ijk}=\mu+\tau_{ik}+\pi_j+\alpha_{d[i,j]}+\lambda_{d[i,j-1]}+\varepsilon_{ijk}$$

on

y_{ijk} ésl'observació de l'"*esophageal temperature*", TES, en l'individu *k*, període *j* i grup *i*

τ_{ik} és l'efecte de l'individu *k* en el grup *i*, *i*=1≡"*lidocaine-cocaine*"; i=2≡"*cocaine-lidocaine*";

π_j és l'efecte del període *j*, *j*=1,2

$\alpha_{d[i,j]}$ és l'efecte directe del tractament subministrat al grup *i* en el període *j*; *d[1,1]=d[2,2]=*"*cocaine*"; *d[1,2]=d[2,1]=*"*lidocaine*".

$\lambda_{d[i,j-1]}$ és l'efecte del "*carry-over*" del tractament proporcionat al grup *i* en el període *j–1*; *d[1,1]*=efecte retardat del tractament *"cocaine"; d[2,1]*=efecte retardat del tractament *"lidocaine"*.

Les dades són les que apareixen seguidament.

Mechanism of Cocaine Induced Hyperthermia in Humans

PERIOD	GROUP	TREAT	PATNUM	TES	PREHS
1	1	2	4	38.360	36.88
1	1	2	6	38.060	36.48
1	1	2	7	37.580	36.89
1	2	1	1	37.530	36.70
1	2	1	2	37.980	37.09
1	2	1	3	37.410	36.62
1	2	1	5	38.300	36.48
2	1	1	4	38.510	36.86
2	1	1	6	37.800	36.66
2	1	1	7	37.920	37.10
2	2	2	1	38.163	36.69
2	2	2	2	38.470	37.22
2	2	2	3	37.710	36.71
2	2	2	5	38.540	36.45

```
proc glm;
 class GROUP PATNUM PERIOD TREAT;
 model TES=GROUP PATNUM (GROUP) PERIOD TREAT;
 random PATNUM (GROUP)/test;
 estimate 'cocaine-lidocaine' TREAT 1 -1;
run;
```

The GLM Procedure

Class Level Information

Class	Levels	Values
GROUP	2	1 2
PATNUM	7	1 2 3 4 5 6 7
PERIOD	2	1 2
TREAT	2	1 2

Number of observations 14

Dependent Variable: TES

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	8	1.74821665	0.21852708	7.67	0.0191
Error	5	0.14253171	0.02850634		
Corrected Total	13	1.89074836			

R-Square	Coeff Var	Root MSE	TES Mean
0.924616	0.444033	0.168838	38.02379

En primer lloc, notar que, al ser un disseny balancejat la descomposició de Sumes de Quadrats és única i, per tant, "Type III SS" no afecta coincidint amb "Type I SS":

Source	DF	Type I SS	Mean Square	F Value	Pr > F
GROUP	1	0.00222215	0.00222215	0.08	0.7913
PATNUM(GROUP)	5	1.39148171	0.27829634	9.76	0.0129
PERIOD	1	0.25596064	0.25596064	8.98	0.0302
TREAT	1	0.09855215	0.09855215	3.46	0.1221

Source	DF	Type III SS	Mean Square	F Value	Pr > F
GROUP	1	0.00222215	0.00222215	0.08	0.7913
PATNUM(GROUP)	5	1.39148171	0.27829634	9.76	0.0129
PERIOD	1	0.20783501	0.20783501	7.29	0.0428
TREAT	1	0.09855215	0.09855215	3.46	0.1221

Source	Type III Expected Mean Square
GROUP	Var(Error) + 2 Var(PATNUM(GROUP)) + Q(GROUP)
PATNUM(GROUP)	Var(Error) + 2 Var(PATNUM(GROUP))
PERIOD	Var(Error) + Q(PERIOD)
TREAT	Var(Error) + Q(TREAT)

Tests of Hypotheses for Mixed Model Analysis of Variance

Dependent Variable: TES

Source	DF	Type III SS	Mean Square	F Value	Pr > F
GROUP	1	0.002222	0.002222	0.01	0.9323
Error	5	1.391482	0.278296		
Error: MS(PATNUM(GROUP))					

Source	DF	Type III SS	Mean Square	F Value	Pr > F
PATNUM(GROUP)	5	1.391482	0.278296	9.76	0.0129
PERIOD	1	0.207835	0.207835	7.29	0.0428
TREAT	1	0.098552	0.098552	3.46	0.1221
Error: MS(Error)	5	0.142532	0.028506		

Dependent Variable: TES

Parameter	Estimate	Standard Error	t Value	Pr > |t|
cocaine-lidocaine	-0.16954167	0.09118306	-1.86	0.1221

El model ajustat es correspon al model complert (Model 4). A continuació, es pot veure que:

Source	Type III Expected Mean Square
GROUP	Var(Error) + 2 Var(PATNUM(GROUP)) + Q(GROUP)
PATNUM(GROUP)	Var(Error) + 2 Var(PATNUM(GROUP))
PERIOD	Var(Error) + Q(PERIOD)
TREAT	Var(Error) + Q(TREAT)

reprodueix les Esperances de les Sumes de Quadrats, equivalent GROUP a la font de variació "Carry-over", PATNUM(GROUP) a la "Residual Entre-Individus", Var(PATNUM(GROUP))=σ^2_τ, Var(Error)= σ^2_e i els termes Q a la variabilitat sistemàtica dels efectes a què fa referència.

Per a contrastar les hipòtesis d'interès:

Source	DF	Type III SS	Mean Square	F Value	Pr > F
GROUP	1	0.00222215	0.00222215	0.08	0.7913
PATNUM(GROUP)	5	1.39148171	0.27829634	9.76	0.0129
PERIOD	1	0.20783501	0.20783501	7.29	0.0428
TREAT	1	0.09855215	0.09855215	3.46	0.1221

S'observa que el valor FC=0,08 amb un p_valor de 0,7913 i, per tant, no hi ha evidència en les dades per rebutjar H_0:$\lambda_1=\lambda_2$. Assumint igualtat d'efectes "*carry-over*", s'està en condicions de contrastat la igualtat d'efectes directes dels tractaments i d'efectes període: FT=3,46 i el p_valor és 0,1221, pel que tampoc es pot rebutjar H_0:$\alpha_1=\alpha_2$; però sí hi ha efecte període al resultar FP=7,29 amb un p_valor de 0,0428<0,05.

2.6.3.- L'ÚS DE "BASELINES" I/O DE COVARIABLES

Una forma d'augmentar la potència del test per a la diferència d'efectes arrossegats del tractament (donat que usa la 'suma' d'on la variabilitat introduïda per l'individu no s'ha eliminat) és l'ús de les mesures basals preses a l'inici de l'estudi o durant el període "*wash-out*". Algunes vegades, per raons ètiques, per exemple, no es porta a terme el període "*wash-out*" i, en aquest cas, la mesura basal prèvia al tractament del primer període pot ser tractada com una covariable, en canvi, no es pot

dir el mateix de la mesura basal al segon període ja que estaria afectada pel primer tractament.

La importància de les covariables resideix en: en primer lloc, estudiar la interacció tractament-covariable. I, en segon lloc, estudiar l'efecte principal de la covariable en la resposta, si hi ha relació aleshores part de la variabilitat entre-individus s'explica amb els valors de la covariable. En aquest darrer cas, la variabilitat residual entre-individus es redueix.

En el disseny "*2x2 cross-over*", la segona utilitat de la covariable és rellevant només per contrastar l'efecte arrossegat o la interacció entre la tractament i període perquè la resta de contrastos es basen en les diferència intra-individus d'on la variabilitat inter-individu s'ha eliminat.

Si la covariable és categòrica s'introdueix com a factor (A) addicional. En presència d'interacció entre covariable i efecte directe del tractament, és necessari interpretar l'efecte del tractament per cada nivell del factor A. De manera similar s'examina la interacció entre la covariable i el període.

Pot passar que la interacció entre la covariable i l'efecte directe del tractament no resulti significativa, llavors l'avantatge de reduir la variabilitat entre-individus dependrà del seu ús com a factor bloc i la seva incorporació serà efectiva depenent de l'aportació dels quadrats mitjos de l'efecte principal de la covariable als quadrats mitjos residuals entre-individus.

Si és conegut el guany que pot representar la covariable abans de dissenyar l'experiment, els individus s'han d'aleatoritzar dins de cada nivell de la covariable ja que d'aquesta manera es garanteix un disseny balancejat i equilibrat. Per ser balancejat la descomposició de la suma de quadrats serà única i per ser equilibrat s'aconsegueix la màxima eficiència. Després, és essencial que sigui inclosa a l'anàlisi.

3.- RESULTATS

3.1.- SELECCIÓ DELS ARTICLES I DIAGRAMA DE FLUX

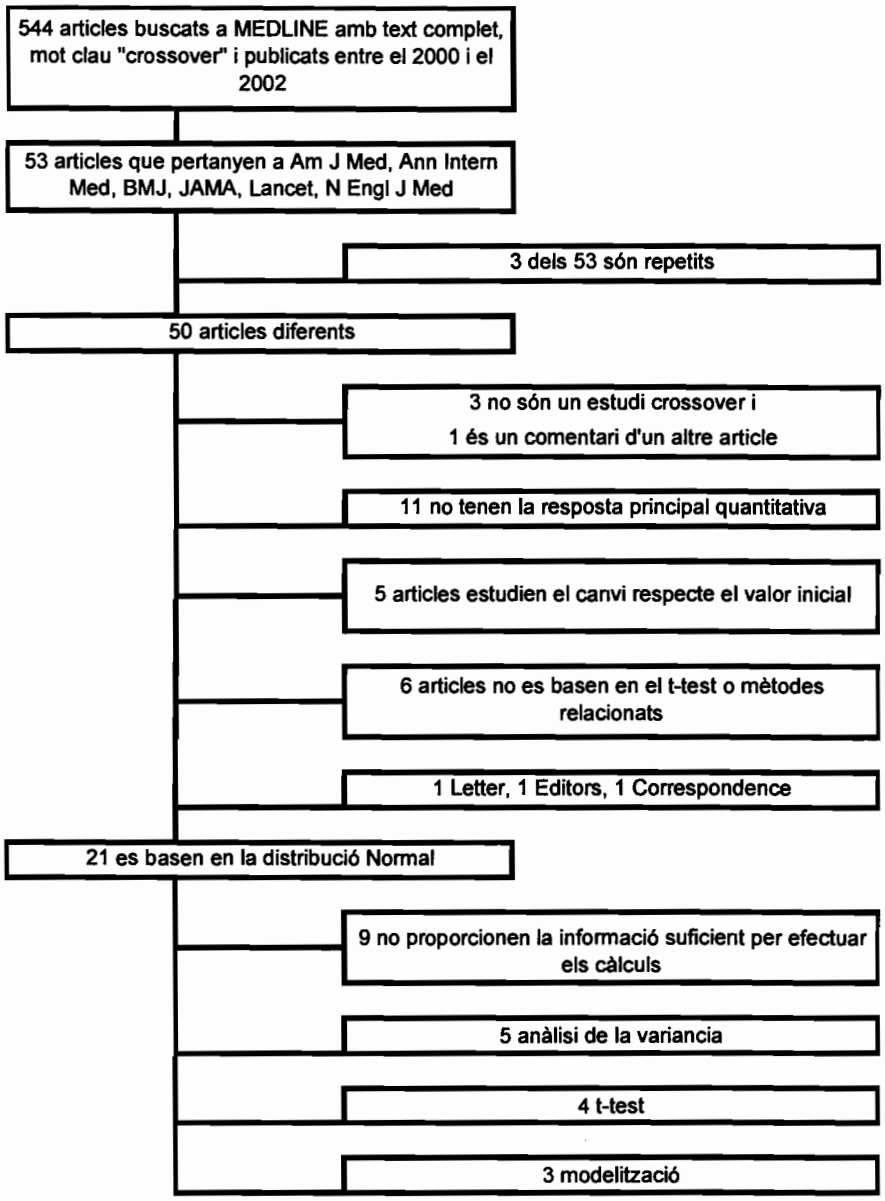


Figura 1: Diagrama de Flux

Com es pot veure a la figura 1, a partir de la cerca realitzada a MEDLINE descrita a l'apartat 2.1, s'obtenen un total de 544 articles de revistes mèdiques. D'aquests articles, es seleccionen els que pertanyen a les revistes següents [número d'articles obtinguts per cadascuna]:

- The American Journal of Medicine [4]
- Annals of Internal Medicine [9]
- The British Medical Journal [10]
- The Journal of the American Medical Association [8]
- The LANCET [7]
- The New England Journal of Medicine [12]

En total s'obtenen 50 articles dels quals s'inclouen els estudis amb disseny "*cross-over*", descartant els articles (numeració Annex 6.1):

- (212) de *Am J Med* perquè està planificat com a un disseny de grups paral·lels però aplica el "*cross-over*" al grup dels no-tractats (intervenció placebo) com a terapia de rescat.
- (119) de *Ann Intern Med* donat que és un "*Summary for patients*" i fa referència a l'article (120) de la mateixa revista.
- (196) de *BMJ* per tractar-se d'un disseny de grups paral·lels.
- (13) de *JAMA* perquè l'objectiu de la planificació del disseny "*cross-over*" no és el de reduir la variabilitat sinó el d'eliminar possibles biaixos.

Dels 46 articles restants s'escullen aquells amb resposta principal continua.

En la taula que figura a continuació es troba el codi Medline dels articles descartats per no verificar el criteri d'inclusió anterior, dels que estudien el canvi respecte dels valors inicials i de cartes de recerca:

		Resposta principal no contínua	Estudi del canvi respecte valors inicials	"*Research Letters*"
Revista	Am J Med		409	
	Ann Intern Med	120	319, 447	
	BMJ	39, 197, 282, 442		5, 408
	JAMA	16, 279	503	
	Lancet	341		
	N Engl J Med	189, 356, 528	442	6

A part, també s'ha buscat que l'estudi basi el mètode d'anàlisi de les dades en la teoria de la Normalitat: t-test, anàlisi de la variància, modelització o mètodes relacionats; excloent per aquest motiu els articles que apareixen amb numeració Medline a la taula següent:

		Mètode d'anàlisi		
	Revista	*"Wilcoxon signed-rank test"*	Estudi multinivell	*"Long-rank test"*
Revista	Am J Med			
	Ann Intern Med			
	BMJ	336		
	JAMA		335	
	Lancet	413		
	N Engl J Med	220, 309		20

Es pot trobar més informació sobre els articles rebutjats a l'Annex 6.2.

En total ens queden 21 articles dels quals:

- 9 no proporcionen la informació suficient per efectuar els càlculs
- 5 s'han realitzat els càlculs a partir d'informació sobre l'anàlisi de la variància
- 4 s'han calculat a partir de la prova t-test
- 3 han estat modelats per obtenir-ne els resultats.

3.2.- DADES OBTINGUDES

Dels 21 articles seleccionats per l'estudi, de bon principi només hi ha 11 que permeten fer els càlculs directament.

La resta no proporcionen informació suficient i per això es va decidir de posar-se en contacte amb els autors per demana'ls-hi la informació que mancava. D'aquests 10, s'escriu per mail tant a l'autor principal com als restants autors dels què es va poder aconseguir l'adreça de correu electrònic, en algun cas també es va telefonar.

Fins aquests moments només hem rebut resposta amb dades d'un autor (Craig Crandall). Dels restants, a l'Annex 6.6 es pot trobar la relació de mails enviats per solicitar la informació i les respostes, si n'hi ha.. Per tant, han quedat 9 articles als que no hem pogut realitzar els càlculs per falta d'informació.

3.2.1.- INFORMACIÓ DELS ARTICLES

En aquest apartat es proporciona la informació que s'ha extret de cada article per tal de realitzar els càlculs pertinents a l'estudi. Al final del títol de cada article, figura el codi Medline i la revista a la què s'ha publicat.

La presentació s'efectua en primer lloc, segons la metodologia emprada en l'anàlisi de les dades i, posteriorment, apareixen els articles dels quals no es disposa d'informació suficient, així:

ANÀLISI DE LA VARIÀNCIA

Mechanism of Cocaine hyperthemia in Humans [*30] (Ann Intern Med)	
N=7 healthy, cocaine naive volunteers	
Responses:	Internal temperature, skin temperature, cutaneous vascular conductance and sweat rate
Statistical analysis:	two-way repeated mesures analysis of variance
Main outcome:	Internal temperature

Comentari[*30]**:** L'anàlisi de la variància de les dades d'aquest estudi, sense ajustar pels valors basals, s'ha reproduit en l'apartat 2.6.2. on hi figuren les dades originals, proporcionades per l'autor Craig Crandall.

En l'apartat 3.3.2., el càlcul de la variància "within" es duu a terme a partir de la variància de la variable diferència d'efectes dels tractaments.

Thyroxine treatment in patients with symptoms of hypothyroidism but thyroid function tests within the reference range: randomised double blind placebo controlled crossover trial [*206] (BMJ)

Design: Randomised double blind placebo controlled cross-over trial

BASELINES: Table A on the BMJ website (Statistical methods: baseline characteristics were summarised by means (SD) or counts and percentages as appropriate. For each variable, participants were grouped into their sequence group (thyroxine-placebo or placebo-thyroxine) to create data summaries in relation to treatment

	Patients (n=25)	Controls (n=19)
Thyroid stimulating hormone (mU/l) at 0 mins	1.9 (1.1)	1.4 (0.9)

Participants: 22 patients with symptoms of hypothyroidism who had function tests within the reference range, and 19 controls (healthy volunteers)

Main outcome measures: thyroid function test (biochemical measures), measures of cognitive function and of psychological and physical wellbeing

Statistical analysis: In each set of participant, the effect of treatment on each variable was studied by using analysis of variance models in relation to patient, period, and treatment. The within patient treatment differences were adjusted by period and 95% confidence limits created.

Table 1.- Response to thyroxine of 22 patients with symptoms of hypothyroidism byt thyroid function test within the reference range. [Mean (SD)]

Outcome: Biochemical measures	Thyroxine	Placebo	Adjusted difference(95%CI)
Thyroid stimulating hormone	0,66 (0,77)	1,77 (1,21)	-1,17 (-1,76 to -0,59)

Table 2.- Response of 19 healthy participants to thyroxine. [Mean (SD)]

Outcome: Biochemical measures	Thyroxine	Placebo	Adjusted difference(95%CI)
Thyroid stimulating hormone	0,32 (0,38)	1,55 (1,54)	-1,17 (-1,80 to -0,53)

Comentari[*206]: l'efecte del tractament sobre la "*thyroid stimulating hormone*" es proporciona per separat pels pacients i pels controls, però no s'efectua cap comparació entre aquests dos grups (independents).

Effects of diet and simvastatin on serum lipids, insulin, and antioxidants in hypercholesterolemic men: a randomized controlled trial [*149] (JAMA)

N=60 previously untreated hypercholeterolemic men

n(simvastatin/placebo)=30 n(placebo/simvastatin)=30

Primary variables: Cholesterol and LDL-Cholesterol

Statistical analysis: Analysis of variance for repeated measures, with contrasts between baseline and simvastatin or placebo treatment periods, was used to test the significance of dietary changes within the dietary treatment and habitual diet groups. The analysis of variance model was fitted separately to the dietary treatment an habitual diet groups, where period and carryover effects were tested. Because no period or carryover effects were observed and baseline values affected the outcome, repeated analyses of covariance with baseline values as covariates, dietary treatment and habitual diet as intersubject factors, and placebo and simvastatin treatment as intersubject factors were included in the final models.

Table 2.- Dietary intake of cholesterol, fatty acids, fiber, vitamin E, ascorbic acid and beta-carotene (Mean(SD))

	Baselines	
Variable	Habitual diet	Dietary treatment
cholesterol, mg/dL	259 (24)	250 (21)

Table 3.- Effects of Dietary Treatment and Simvastatin (mean (SE) (95%CI))

Variable	Simvastatin-placebo (*)
cholesterol, mg/dL	-53 (2) (-58 to -48)

Comentari[*149]: El model ajustat finalment a les dades seria:

$$y_{ijlk}=\phi \cdot x_{ijlk}+\mu+\beta_l+\tau_{ik}+\alpha_{d[i,j]}+\varepsilon_{ijlk}$$

on

i=1≡"*simvastatin/placebo*" ,i=2≡" *placebo/simvastatin*"

x_{ijlk} representa la mesura basal del colesterol,

β_l és l'efecte de la covariable (factor) dieta amb l="*habitual diet*"; "*dietary treatment*".

$\alpha_{d[i,j]}$ és l'efecte directe del tractament; d[1,1]=d[2,2]="*simvastatin*", d[1,2]=d[2,1]="*placebo*".

La variància entre-individus, a més de la forma habitual en la majoria dels articles, s'ha calculat a partir de la comparació entre dietes, que es tracta d'una comparació de mitjanes de dos grups independents i que equival a l'anàlisi que es duria a terme si el disseny fos de grups paral·lels.

Entrainment of free-running circadian rhythms by melatonin in blind people [*388] (N Engl J Med)

N=7 subjects who had free-running circadian rhythms

Main outcome measure: timing of the increase in endogenous melatonin production (was determined as a marker of the circadian phase) and polysomnographic variables

Statistical analysis: two-sided t-test for repeated measures (differences in circadian period : TABLE 1)

Analysis of variance and post hoc two-sided t-test for repeated measures (differences in polysomnographic variables [total time asleep, sleep latency, sleep efficiency and time spent awake after the onset of sleep] according to treatment, stage of the trial and interaction : TABLE 2)

Table 1.- Characteristics of 7 blind subjects with free-running circadian rhythms at base line and during the administration of placebo or melatonin

	Circadian period		
Subject no.	Base-line	Placebo	Melatonin
1	24,2	24,2	24
2	24,3	24,3	24
3	24,4	24,3	24
4	24,5	-	24
5	24,4	24,3	24
6	24,6	24,5	24
7	24,9	24,8	24,3
Mean (SD)	24,5 (0,2)	24,4 (0,2)	24,0 (0,1)

Comentari [*388]: Com les dades per a cada individu estan disponibles i els resultats que es presenten fan referència a diferències entre-pacients (però no intra-), la variància "*between*" es calcula a partir de la variància de la resposta final, que s'obté ponderant les variàncies del "*circadian period*" per a les intervencions (placebo i "*melatonin*").

La variància "*within*" s'estima ponderant les variàncies de les tres diferències intra-individus ("*baseline-placebo*", "*baseline-melatonin*", "*melatonin-placebo*"), una altra

possibilitat seria estimar-la únicament de la diferència "*melatonin-placebo*" ja que no pressuposa res sobre el valor basal.

Beneficial effects of high dietary fiber intake in patients with type 2 diabetes mellitus [*481] (N Engl J Med)

N=13 patients with type 2 diabetes mellitus

n(high-fiber diet/ADA diet)=6 n(ADA diet/high-fiber diet)=7

Main outcome: plasma glucose and plasma lipid concentrations

Statistical analysis: repeated-measures analysis of variance

Table 4.- Fasting plasma lipid and lipoprotein concentrations during the last week of the study periods (days 38 through 42)

Variable	ADA diet	High-fiber diet	Difference between diets (95%CI)
plasma total cholesterol	210 (33)	196 (31)	-14 (-27 to -2)

Comentari [*481]: El càlcul de la variància "*between*" s'efectua a partir de la variància de la resposta final (No es disposa de valors sobre els valors basals).

t-TEST

Severe Hypertriglyceridemia with Insulin Resistence Is Associated with Systemic Inflamation: Reversal with Bezafibrate Therapy in a Randomized Controlled Trial [*131] (Am J Med)

N=18 patients with hypertriglyceridemia N=20 normolipidemic controls

Main outcome: Tumor necrosis factor (TNF)-alfa

Table 1.- Baseline Characteristics of Controls and Hypertriglyceridemic Patients

Variable	Controls (N=19)	Hypertriglyceridemic (N=13)	Mean Difference (95% CI)*
TNF-alfa production (pg/mL)	23200 (8700)	35000 (10000)	11700 (7800 to 15700)

Table 2.- Effects of Placebo and Bezafibrate on Serum Lipids, Insulin, Glucose, Insulin Resistance, and Inflammatory Parameters in 18 Hypertriglyceridemic Patients

Variable	Placebo	Bezafibrate	Mean Difference (95% CI)*
TNF-alfa production (pg/mL)	35000 (6500)	31400 (6500)	-3600 (-6500 to -600)

Comentari [*131]: Pel càlcul de la variància "*between*" s'usa la variància de la resposta final perquè per a la comparació dels efectes de placebo i "*bezafibrate*" els resultats que es proporcionen fan referència al grup "*hypertriglyceridemic*" únicament.

Substituting walnuts for monounsaturated fat improves the serum lipid profile of hypercholesterolemic men and women: a randomised crossover trial [*497] (Ann Intern Med)

Patients: adult men and women with polygenic hypercholesterolemia

n(walnut diet/control diet)=25 n(control diet/walnut diet)=24

Main outcome: Serum levels of total and LDL cholesterol

Table 3.- Serum lipid and lipoprotein levels, analytes related to low-density lipoprotein oxidation and body weight at the end of each diet period [Mean (SD)]

Variable	Baseline	control diet	walnut diet
Total cholesterol level	7,16 (0,85)	6,81 (0,79)	6,52 (0,90)
	Treatment effect (95%CI)		
Total cholesterol level	-0,28 (-0,43 to -0,12)		

Cardiovascular effects of sildenafil during exercise in men with known or probable coronary artery disease[*144] (JAMA)

N=105 men who had erectile dysfunction and known or highly suspected CAD (Coronary Artery Disease)

n(sildenafil/placebo)=53 n(placebo/sildenafil)=52

Main outcome measures: Hemodynamic effects of sildenafil during exercise (onset, extent, and severity of ischemia) assessed by exercise echocardiography

Statistical analysis: For continuous varibles, treatment effects were assessed by calculating the difference between first and second study data. These differences were then compared by using the 2-sample test wiht corresponding 95% CI. The matched-pairs paired t test was used to evaluate differences between premedication and postmedication data.

Results:

	BASELINES	
	Sildenafil	Placebo
Systolic blood pressure	135 (19)	135 (20)

Table 2.- Exercise test Hemodynamics (N=105) (Mean (SD))

Variable Systolic blood pressure	Sildenafil	Placebo	Mean difference (95%CI)
Rest	128 (17)	133 (19)	4,3 (0,9 to 7,7)
Exercise	174 (29)	176 (30)	2,4 (-2,5 to 7,4)

Rest was measured after medication receipt but before exercise testing

Comentari[*144]: Una de les premisses per a la comparació de tractaments és que aquests afecten a la mitjana de la resposta però no a la seva variància. En la taula 2, es disposa de la "*Systolic blood pressure*" a "*Rest*" i a "*Exercise*"; però només s'utilitza el moment "*Rest*" perquè la variabilitat de la resposta havent fet l'exercisi viola la premissa efectuada.

Sputum eosinophilia and short-term response to prednisolone in chronic obstructive pulmonary disease: a radomised controlled trial [*363] (Lancet)

Participants: 67 patients who had COPD (chronic obstructive pulmonary disease) treated with bronchodilators

n(placebo/prednisolone)=33 N(prednisolone/placebo)=26

Primary outcomes: Change in postbronchodilator FEV(1), total CRQ score, and shuttle walk distance after prednisolone compared with placebo

statistical analysis: Improvement in the primary and secondary outcome variables after prednisolone compared with placebo are reported as paired mean differences (95% CI); Mean paired difference between prednisolone and placebo (Mean (95%CI))

Results: Mean difference (95%CI) Postbronchodilator FEV(1) 0,07 (0,01 to 0,14)

Table 1.- Mean (SE) outcome measures before and after treatment

Variable		Before	After
postbronchodilator FEV(1)	prednisolone	1,10 (0,05)	1,15 (0,05)
	Placebo	1,15 (0,06)	1,14 (0,06)

MODELITZACIÓ

The effect of two different dosages of intravenous immuglobulin of the incidence of recurrent infections in patients with primary hypogammaglobulinemia [*240] (Ann Intern Med)

N(initial)=43 with primary hypogammaglobulinemia; 41 of whom completed the protocol

n(standard dose/high dose)=25 n(high dose/standard dose)=16

Primary outcomes: Ocurrence of infeccions

statistical analysis: Multivariate model in which the number of infections was the dependent variable and dosage, sequence of treatment, and patient were the independent variables. The total number of acute infections was compared by using the paired t-test

Table 1.- Infections in patients with primary hypogammaglobulinemia treated with 2 different dosages of intravenous immuglobulin (Mean +/- SD (95%CI),n)

Variable	Standard-dose therapy (n=41)	High-dose therapy (n=43)	Difference (95% CI)
Total immunodeficiency-related infections per patient	3,5 +/- 2,6 (2,7 to 4,3)	2,5 +/- 2,4 (1,8 to 3,2)	1,1 (0,4 to 1,8)

Comentari[*240]: La variància "*Between*" s'ha calculat a partir de la variància de la resposta final (No es disposa d'informació sobre la "*baseline*").

Effects of fexofenadine, diphenhydramine, and alcohol on driving performance [*511 (Ann Intern Med)]

Participants:	40 licensed drivers with seasonal allergic rhinitis
Primary end point:	coherence (correlation between the velocity of the participant's vehicle and the velocity of the head vehicle)
Statistical analysis:	Each treatment effect was estimated with equal precision in a model with treatment, period, and first-order carry-over effects...The statistical method we used (mixed general linear model) requires that the data be approximately normally distributed. Most of the outcomes we measured were significantly non-normally distributed, and it was necessary to re-express (transform) them to achieve normality. We use Box-Cox analysis to select an appropriate power transformation of each variable. Specifically, coherence (c) was re-expressed as 1-sqrt(1-c^2)^1,25)... To make statistic interpretable, we converted all statistical results-means, differences and CIs-back to the original, more interpretable measurement scales.

Table 1.- Primary end point : coherence [Participants: n = 40] Mean (SD) (95% CI)

	Mean Coherence value (SD) (95%CI)
alcohol	0,920 (0,014) (0,891 to 0,945)
diphenhydramine	0,877 (0,019) (0,837 to 0,911)
fexofenadine	0,915 (0,014) (0,884 to 0,940)
placebo	0,906 (0,015) (0,875 to 0,933)
alcohol vs diphenhydramine	0,043 (0,012) (0,021 to 0,068)
alcohol vs fexofenadine	0,005 (0,009) (-0,012 to 0,024)
alcohol vs placebo	0,014 (0,009) (-0,004 to 0,033)
Diphenhydramine vs fexofenadine	-0,038 (0,013) (-0,063 to -0,013)
diphenhydramine vs placebo	-0,029 (0,012) (-0,054 to -0,005)
fexofenadine vs placebo	0,009 (0,010) (-0,01 to 0,028)

Comentari[*511]: La resposta principal es transforma per tal de normalitzar-la i tant les mitjanes com les diferències i els intervals de confiança s'han transformat a l'escala original, però no les desviacions tipus (que estan en l'escala transformada). La variància de la resposta final (que és la suma de les variàncies intra- i entre-individus) es calcula ponderant les variàncies de la reposta pels tractaments per separat; mentre que la variància intra-individus és la meitat de la variància que s'obté ponderant les variàncies de les diferències d'efectes entre tractaments.

Effects on blood pressure of reducted dietary sodium and the dietary approaches to stop hypertension (DASH) diet [*334 (N Engl J Med)]

N=412 participants	
Primary end point:	systolic blood pressure
Study design:	participants were randomly assigned to follow one of the 2 diets (DASH diet/CONTROL diet) according to a parallel group design. They ate their assigned diet at each of the 3 sodium levels for 30 consecutive days in random order in a crossover study

Statistical analysis: A unified generalized-estimating-equation model with an exchangeable covariance matrix was used for primary analyses. Blood pressure was the outcome. The base-line blood pressure, the clinical center, and the cohort were represented in the model as fixed effects, whereas the intervention periods were included as random effects to allow for within-person correlation among blood-pressure measurements. The model included indicators of the cohort, the clinical center, and the carryover effect from the previous intervention. Results were similar with and without carryover in the model. Indicators for the subgroups specified in the study protocol (hypertensive status, race, and sex) and for the relevant interactions with the effects of the diet assignments and sodium levels were included in the subgroup analyses.

Table 1.- Base-line characteristics of the participants (Mean (SD))

Characteristic		
Blood pressure (mmHg)	Control diet (N=204)	DASH-diet (n=208)
systolic	135 (10)	134 (10)

Figure 1.- Mean changes in blood pressure for various sodium levels

Comentari [*334]: A part dels efectes fixos (valors basals, el centre i la cohort) que s'inclouen al model, el període en aquest cas és un efecte aleatori.

$$y_{ijlmk}=x_{ijk}+\mu+\beta_l+\delta_m+\tau_{ik}+A1_{d[i,j]k}+A2_{d[i,j]k}+A3_{d[i,j]k}+\pi_j+\alpha_{d[i,j]}+\lambda_{d[i,j-1]}+e_{ijk}+$$
$$+ \text{interaccions rellevants}[A(.)_{d[i,j]k}:\beta_l\ ;\ A(.)_{d[i,j]k}:\ \alpha_{d[i,j]}]$$

on

β_l és l'efecte fix del centre,

δ_m és l'efecte fix de la cohort

$A1_{d[i,j]k}$, $A2_{d[i,j]k}$ i $A3_{d[i,j]k}$ són els indicadors de l'estat hipertensiu, la raça i el sexe.

$\alpha_{d[i,j]}$ és l'efecte fix del tractament de sodi. Com no s'especifica el patró del quadrat llatí, per il·lustrar els valors de d[i,j] si suposem:

Seqüència	Període 1	Període 2	Període 3
1	High	Intermediate	Low
2	Intermediate	Low	High
3	Low	High	Intermediate

d[1,1]=d[2,3]=d[3,2]="*high*";

d[1,2]=d[2,1]=d[3,3]="*intermediate*";

d[1,3]=d[2,2]=d[3,1]="*low*".

π_j és l'efecte aleatori del període de tractament, amb mitjana zero i variància σ_π^2.

NO PROPORCIONEN INFORMACIÓ SUFICIENT PER EFECTUAR ELS CÀLCULS

Els efectes "absoluts" per separat no són identificables i no té cap valor proporcionar-los. El disseny serà eficient per:

a) Detectar l'efecte diferencial (prova d'hipòtesi)
b) Comunicar la seva magnitud informant del soroll aleatori (interval de confiança)

L'interval de confiança és més informatiu que la prova d'hipòtesi ja que quantifica la incertesa i, d'altra banda, és obligat reportar-lo (CONSORT) a diferència del p_valor.

Els articles que s'inclouen en aquest subapartat analitzen les dades adequadament d'acord amb el disseny aparellat del "*cross-over*"; però només presenten resultats pels efectes "absoluts" dels tractaments, pel que el benefici/eficiència del disseny amb intercanvi del tractament és nul en la presentació de resultats.

S'ha escrit a l'autor per completar la informació però fins al moment no s'ha rebut resposta, si no s'especifica el contrari.

The effects of Transdermal Estradiol on the response to mental stress in postmenopausal women: a randomized trial [*377] (Am J Med)

N=10 postmenopausal women

n(17-beta stradiol /placebo)=6 n(placebo/ 17-beta stradiol)=4

Main outcome measures: plasma catecholamine levels and the cardiovascular and metabolic responses to a 15-minute stress with mental arithmetic

Statistical analysis: repeated measures analysis of variance in which the effect of both time and treatment were evaluated. (interaction between order of treatment and effects on stress response were not statistically significant)

Results: Epinephrine levels (variable) (pmol/liter)

	Mean (SD)
Placebo (baseline)	273 (67)
Placebo treatment	431 (135)
Epinephrine (baseline)	279 (72)
Epinephrine treatment	357 (77)

Mediterranean and Low-Fat diets improve endothelial function in hypercholesterolemic men [*268] (Ann Intern Med)

N=22 hypercholesterolemic (volunteers) men

n(NCEP-1/mediterranean diet)=10 n(mediterranean diet/NCEP-1)=12

Measurements: endothelial function (flow-associated vasodilatation)

Statistical analysis: For each diet: analysis of variance for repeated measures to test for dietary effects on plasma lipid levels and edothelial function When significant effects were detected ($p<0,05$), the Bonferroni test was used for a post hoc comparison.

BASELINE: Saturated fat diet (initial 28 day period during which all patients consumed a diet high in saturated fat.)

Table 1.- Flow-associated vasodilation (endothelial-dependent), glyceryl trinitrate-induced vasodilatation (endothelial-independent) and plasma levels of soluble intracellular cell adhesion molecule-1, vascular cell adhesion molecule-1 and P-selectin at the end of each dietary period [mean(SD)]

Variable (main outcomes)	Saturated fat diet	Mediterranean diet	NCEP-1 diet
Flow-associated vasodil.	0,41 (0,05)	0,54 (0,04)	0,45 (0,03)

Effect of temazepan on ventilatory response at moderate altitude [*532] (BMJ)

N=7 healthy men aged 21 to 27

Measurements: arterial oxigen (Pao2) and carbon dioxide (Paco2) concentrations

Statistical analysis: Differences in blood gas concentrations Before and After temazepan or placebo at each altitude were analysed by paired t test

Table.- Pao2 and Paco2 concentrations (Kpa) of 7 men before and one hour after 10 mg temazepan at 171 and 3000m

The table shows the results of blood gas analysis before and after temazepan. At 171m blood gas concentrations did not change significantly after temazepan. At 3000m the arterial oxigen pressure decreased and carbon dioxide pressure increased significantly after temazepan. The mean decrease in arterial oxygen concentration between altitudes was 0,77 (95% CI -8,02 to -3,69) KPa (P<0,01) and the mean increase in arterial carbon dioxide concentration was 0,3 (0,46 to 4,11) KPa (P<0,05). Placebo did not affect blood gas concentrations at either altitude.

Comentari[*532]: Es tracta d'un "*Editor*" i només mostra les dades pel tractament actiu; sobre el tractament placebo comenta que no afecta a la resposta principal però no proporciona cap resultat estadístic. Les dades podrien ser analitzades com estudi del canvi aprofitant la informació Abans/Després.

Efficacy of 3 commonly used hearing aid circuits: a cross-over trial [*384 (JAMA)]

N=360 patients with bilateral sensorineural hearing loss

Trial design: 3-period 3-treatment (PC,CL and WDRC) cross-over design

Quality Rating Test: 52Q, 62Q, 74Q (Q = quiet)

52N, 62N, 74N (N= absolute S/B ratio of 10dB)

Main outcome measures: results of test of speech recognition; sound quality, and subjective hearing aid benefit, administered at baseline and after 3-month intervention with an without a hearing aid

Statistical analysis: mixed, repeated measures model was used to compare the 3 hearing aid circuits for the individual outcome variables. If the overall test was statistically significant, then pairwise comparisons were made between the groups usign the Bonferroni procedure to adjust the α level for multiple tests. No adjustment was made for multiple outcome.

Table 1.- Loudness ratings (10-point scale) obtained in the aided and unaided conditions

		AIDED	
Condition dB	Type of circuit	No of subjects	Mean (SD)
52Q	CL	335	4,14 (0,94)
	PC	333	4,15 (1,00)
	WDRC	330	4,43 (0,99)
52N	CL	335	3,92 (1,06)
	PC	333	3,81 (1,08)
	WDRC	328	4,21 (1,07)
62Q	CL	335	5,33 (0,90)
	PC	333	5,34 (0,90)
	WDRC	330	5,41 (1,00)
62N	CL	334	5,31 (1,19)
	PC	333	5,20 (1,09)
	WDRC	330	5,25 (1,31)
74Q	CL	335	7,96 (1,55)
	PC	333	8,31 (1,49)
	WDRC	330	7,73 (1,59)
74N	CL	335	7,57 (1,89)
	PC	333	8,11 (1,69)
	WDRC	330	7,26 (1,94)

Comentari[*384]: S'ajusta un model d'efectes mixtos, presentant els resultats que s'observen a la taula anterior pels diferents grups per separat, mentre que per a les comparacions entre grups només figura una indicació a peu de taula que indica si les comparacions dos a dos o entre els tres circuits són significatives. Com no dóna

informació de l'error estàndard ni de l'interval de confiança de les comparacions intra-individu (no es pot estimar σ_e^2).

Com la informació de la què es disposa per efectuar els càlculs és insuficient, s'ha escrit a l'autor (Vernon D. Larson) que ha remès la sol·licitud al bioestadístic de l'estudi, David Williams i, en aquests moments, s'està esperant resposta.

Topical butyrate for acute radiation proctitis: randomised, cross-over trial *370 (Lancet)

N=20 patients presenting with ARP (acute radiation proctitis) after completing a cycle of 35-52 Gy external-beam radiation therapy for pelvic malignant disease

n(butyrate/saline)=9 N(saline/butyrate)=9

Continuous outcome variables: bowel movements, night bowel movements

Statistical analysis: Student's test for paired data was used in comparison between groups for continuous outcome variables (number of bowel movements)

Table: Advantage of butyrate over placebo in the treatment of ARP

	Odds Ratio (95% CI)
Bowel movements	-0,95 (-0,83 to –0,16)
Night bowel movements	-0,35 (-0,57 to –0,50)

Results: All patients first treated with butyrate became symptom-free or improved greatly (clinical score from 8,2 (SE 1,6) to 1,5 (SE 0,7)) within the first 3 weeks of treatment. In the placebo group 3 patients had some improvement whereas three deteriorated slightly. Thus, the overall score was unchanged (7,9 (SE 1,8) vs 8,2 (SE 3,4). The number fo bowel movements decreased in the butyrate group from 3,7 (SE 0,6) to 1,7 (SE 0,2), but remained unchanged in the placebo group (from 2,8 (SE 0,3) to 2,6 (0,2))

Comentari*370: Hi ha un error en els "*Odds Rati*" calculats, tot i ser la resposta contínua, donat que són negatius i, en principi, estan transformats a l'escala original. El disseny està planificat com a "*cross-over*" i l'anàlisi efectuat és la comparació de dades aparellades via t-test, però no es presenten els resultats de forma adequada (s'haurien de proporcionar resultats per a la variable diferència), deixant de banda l'ambigüitat de l'explicació.

Comparison of pharmacokinetics and systemic effects of inhaled fluticasone propionate in patients with asthma and healthy volunteers: a randomised cross-over study *421 (Lancet)

Participants: 11 patients with asthma and 13 matched healthy controls

Main outcome variable: area under the curve for plasma fluticasone proportionate concentrations

Statistical analysis: Pharmacokinetic data were log-normalised for group comparisons and are presented as geometric means with 95%CI. The lung-function variables followed as normal distribution and ae presented as arithmetic mean (SD). We compared the results for asthma and cortisol groups by two-tailed Student's test. One-way and multiple ANOVA was applied to identify significant differences between the three different cortisol time points- screening, and samplind days 1 and 2- and the plasma cortisol profiles on the kinetic sampling days.
We used bivariate correlation with Pearson's correlation coefficient for the systemic availability and the area under the curve for fluticasone propionate after inhalation against FEV(1), carbon monoxide transfer, and body-mass index.

Results: The pharmacokinetics of fluticasone propionate differed significantly between the asthma and control groups, seen in plasma area-under-curve values for inhaled fluticasone propionate (1082 [850-1451] vs 2815 pg mL-1 h-1 [2262-3949], mean difference 0.38 [95% CI 0.28-0,55]; p<0.001 figure 2)...

The area-under-curve and systemic availability values were positively correlated with baseline transfer carbon monoxide (figure 3). Variability in systemic exposure within groups was explained partly by carbon monoxide transfer, but not the difference between groups.

Comentari[*421]: Només es proporciona la mitjana geomètrica, la desviació tipus i l'interval de confiança de la resposta principal sota el tractament actiu per a la diferència entre grups independents (asmàtics i controls) de manera que únicament és estimable la variància entre-individus.

Effect of consumption of red wine, spirits, and beer on serum homocysteine [*480 (Lancet)]

Measurement: Homocysteine concentration

Table.- Mean blood concentrations of homocysteine, folate, vitamins B6 and B12 after 3 weeks consumption of red wine, beer, spirits and water.

Comentari[*480]: Es tracta d'una "*Research Letter*" en la que només figuren les mitjanes de la "*homocysteine concentration*" pels diferents tractaments, pel què resulta impossible efectuar cap càlcul referent a variàncies sense disposar de les desviacions tipus ni intervals de confiança.

Transcranial magnetic sitmulation and auditory hallucinations in schizophrenia [*493 (Lancet)]

N= 12 right-handed patients with auditory hallucinations who met Diagnostic and Statistical Manual IV (DSM-IV) diagnostic criteria for schizophrenia

Main outcome measure: Auditory hallucinations were assessed with an individualised, composite scale. A score of ten corresponded to a narrative description of the patient's hallucinations at the time of study entry, with zero corresponding to no hallucination.

Statistical Analysis: Endpoint hallucination ratings were analysed by use of a repeated measure ANOVA with two additional factors; order of stimulation (active or shan first), and concomitant treatment with anticonvulsant drugs.

Results: Reductions in hallucination severity after active compared with sham stimulation were significant ($p<0.006$), as was the interaction between change in hallucination severity and anticonvulsant drugs ($p<0.02$) showing reduced treatment effects with these drugs. No effect of order of stimulation was seen.

Comentari[*493]: Es tracta d'una "*Research Letter*" on només figuren la mitjana i la desviació tipus pels valors basals; per la comparació entre tractaments només proporciona els p-valors.

Transdermal testosterone treatment in women with impaired sexual function after oophorectomy [*414 (N Engl J Med)]

N=75 women who had undergone bilateral salpingo-oophorectomy and hysterectomy before natural menopause

Primary end point: composite score on the Brief Index of Sexual Functioning for women, overall frequency index from the telephone-based diary

Statistical analysis: repeated-measures analysis of variance with terms for period, sequence, and carry over effects included in the model (for estimate the least-squares means corresponding to each treatment); t-test based on analysis of variance (for pairwise comparisons of values for each active dose with those for placebo (with baseline values substracted))

Table 3.- Mean (SD) scores on the Brief Index of sexual functioning for women, expressed as a percentages of the mean values in normal women* (N=65 women from intention-to-treat)

Dimension	Baseline	Placebo
Composite score	52 (27)	72 (37)
	150 (mu)g of testosterone per day	300 (mu)g of testosterone per day

Composite score	74 (37)	81 (37)

(*) Values are expressed as percentages of the mean values in normal women with partners; which were as follows: composite scores, 33.6

Comentari [*414]: Es disposa de resultats pels efectes dels tractaments, però no per a la diferència d'efectes; per tant, no seria eficient en quant a presentació de resulats.

3.3.- CÀLCULS

Tot seguit es mostra la fiabilitat (R) resultant per a cadascun dels articles que han permès el seu càlcul a partir dels procediments descrits als apartats 2.5 i 2.6.

3.3.1.- CÀLCULS DIRECTES

Els articles s'exposen en el mateix ordre que s'han mostrat en l'apartat d'informació, segons el mètode d'anàlisi.

ANOVA

Thyroxine treatment in patients with symptoms of hypothyroidism but thyroid function tests within the reference range: randomised double blind placebo controlled crossover trial [206] (BMJ)

Outcome measure: THYROID STIMULATING HORMONE (BIOCHEMICAL MEASURE)

Response of 22 patients with symptoms of hypothyroidism	baseline variance	1,21
	B-S variance from baseline	1,181399999
	final response variance	1,0285
	B-S variance from final resp	0,999899999
	Adjusted difference (95% CI):	-1,17 (-1,76 to –0,59)
	t(0.05,20):	2,085962478
	SD estimated from UCI:	0,237121073
	SD estimated from LCI:	0,241209368
	SD pooled:	0,239165221
	variance of Difference var.:	0,057200003
	W-S variance	0,028600001
	Reliability (R) from baseline	0,976363635
	Reliability (R) from final resp	0,972192512
Response of 19 healthy participants to thyroxine	baseline variance	0,81
	B-S variance from baseline	0,771752915
	final response variance	1,258
	B-S variance from final resp	1,219752915
	Adjusted difference (95% CI):	-1,17 (-1,80 to –0,53)
	t(0.05,17):	2,109818524
	SD estimated from UCI:	0,278753555
	SD estimated from LCI:	0,27439803
	SD pooled:	0,276575793
	variance of Difference var.:	0,076494169
	W-S variance	0,038247085
	Reliability (R) from baseline	0,952781377
	Reliability (R) from final resp	0,969596912

Effects of diet and simvastatin on serum lipids, insulin, and antioxidants in hypercholesterolemic men: a randomized controlled trial *149 (JAMA)

Measurement: CHOLESTEROL, mg/dL

baseline variance	508,5	Mean (SE) 95% CI; Dietary treatment-habitual diet	
B-S variance from baseline	321,1868883	-19 (3) [-26 to -12]	
Simvastatin-Placebo (n=60):	-53 (SE=2) (-58 to -48)		
t(0.05,59):	2,000997483	SE=SD pooled/sqrt(n)	3
SD estimated from UCI:	19,35526346	SD pooled (diets)	16,43167673
SD estimated from LCI:	19,35526346	Variance of Diff. (diets)	270
SD pooled:	19,35526346	**B-S variance from final resp**	82,68688829
Variance of Difference var.:	374,6262234	**Reliability (R) from final resp**	0,306247734
W-S variance	187,3131117		
Reliability (R) from baseline	0,631635965		

Entrainment of free-running circadian rhythms by melatonin in blind people *388 (N Engl J Med)

Measurement: timing of the increase in endogenous melatonin production (was determined as a marker of the circadian phase)

		from Final Response	from Baseline
	variance	0,023636364	0,04
	B-S variance	0,016724083	0,033087719

Baseline-Placebo	Baseline-Melatonin	Placebo-Melatonin
0	0,2	0,2
0	0,3	0,3
0,1	0,4	0,3
*	0,5	*
0,1	0,4	0,3
0,1	0,6	0,5
0,1	0,6	0,5
SD	0,051639778	0,149602648	0,122474487
Var	0,002666667	0,022380952	0,015
Pooled var			0,013824561
W-S variance	0,006912281		
Reliability (R)	0,707557355	0,827192982	

Beneficial effects of high dietary fiber intake in patients with type 2 diabetes mellitus *481 (N Engl J Med)

Measurement: PLASMA TOTAL CHOLESTEROL

final response variance	1025
B-S variance from final resp	816,588686
Difference between diets (95%CI):	-14 (-27 to -2)
t(0.05,12):	2,200986273
SD estimated from UCI:	19,59958504
SD estimated from LCI:	21,2328838
SD pooled:	20,41623442
variance of Difference var.:	416,822628
W-S variance	208,411314
Reliability (R) from final resp	0,796671889

t- TEST

Severe Hypertriglyceridemia with Insulin Resistence Is Associated with Systemic Inflamation: Reversal with Bezafibrate Therapy in a Randomized Controlled Trial *131 (Am J Med)

Measurement: TNF-alfa production (pg/mL)

baseline variance	42250000
B-S variance from baseline	30334355,01
Treatment effect (95% CI) (n=13):	-3600 (-6500 to -600)
t(0.05,12):	2,178812792
SD estimated from UCI:	4964,471415
SD estimated from LCI:	4798,989034
SD pooled:	4881,730225
Variance of Difference var.:	23831289,99
W-S variance	11915644,99
Reliability (R) from baseline	0,7179729

Substituting walnuts for monounsaturated fat improves the serum lipid profile of hypercholesterolemic men and women: a randomised crossover trial *497 (Ann Intern Med)

Measurement: CHANGES IN SERUM LEVELS OF TOTAL CHOLESTEROL

baseline variance	0,7225	final response variance	0,71705
B-S variance from baseline	0,612760716	**B-S var. from final resp**	0,607310716
Treatment effect (95% CI):	-0,28 (-0,43 to -0,12)	**Reliability (R) from final resp.**	0,846957278
t(0.05/2,47):	2,315491656		
SD estimated from UCI:	0,48359783		
SD estimated from LCI:	0,453372965		
SD pooled:	0,468485397		
variance of Difference var.:	0,219478568		
W-S variance	0,109739284		
Reliability (R) from base.	0,848111718		

Cardiovascular effects of sildenafil during exercise in men with known or probable coronary artery disease *144 (JAMA)

Measurement: SYSTOLIC BLOOD PRESSURE

baseline variance	380,25	**Rest**	
B-S variance from baseline	131,6471802	final response variance	325
Treatment effect (95% CI):	2,4 (-2,5 to 7,4)	**B-S var. from final resp**	76,39718022
t(0.05/2,103):	2,274637154	**Reliability (R) from final resp.**	0,566594146
SD estimated from UCI:	22,52334185		
SD estimated from LCI:	22,07287501		
SD pooled:	22,29810843		
variance of Difference var.:	497,2056396		
W-S variance	248,6028198		
Reliability (R) from baseline	0,346212177		

Sputum eosinophilia and short-term response to prednisolone in chronic obstructive pulmonary disease: a randomised controlled trial *363 (Lancet)

Measurement:	POSTBRONCHODILATOR FEV 1
basal/final response variance	0,178475
B-S variance	0,147369025
Treatment effect (95% CI) (n=59):	0,07 (0,01 to 0,14)
t(0.05,58):	2,001715984
SD estimated from UCI:	0,268609636
SD estimated from LCI:	0,230236831
SD pooled:	0,249423234
variance of Difference var.:	0,062211949
W-S variance	0,031105975
Reliability (R)	0,825712426

MODELITZACIÓ

The effect of two different dosages of intravenous immuglobulin of the incidence of recurrent infections in patients with primary hypogammaglobulinemia *240 (Ann Intern Med)

Measurement:	TOTAL IMMUNODEFICIENCY-RELATED INFECTIONS PER PATIENT
final response variance:	6,247804878
B-S variance from final resp	3,788653508
Difference (95% CI) (n=41):	1,1 (0,4 to 1,8)
t(0.05,40):	2,021074579
SD estimated from UCI:	2,217724676
SD estimated from LCI:	2,217724676
SD pooled:	2,217724676
variance of Difference var.:	4,91830274
W-S variance	2,45915137
Reliability (R) from final resp	0,60639754

Effects of fexofenadine, diphenhydramine, and alcohol on driving performance *511 (Ann Intern Med)

Measurement:	COHERENCE
final response variance	0,0002445
B-S variance from final resp	0,000184583
variance of Difference var.:	0,000119833
W-S variance	5,99167E-05
Reliability (R) from final resp	0,754942059

Effects on blood pressure of reducted dietary sodium and the dietary approaches to stop hypertension (DASH) diet *334 (N Engl J Med)

Measurement: SYSTOLIC BLOOD PRESSURE

baseline variance 100

	DASH-DIET (N=208)	Control diet (N=204)
B-S variance from baseline	54,77991657	53,93818243
B-S variance pooled from baseline	54,3590495	
Treat. effect (95% CI) HIGH-INTERM:	-1,3 (-2,6 to 0,0)	-2,1 (-3,4 to -0,8)
Treat. effect (95% CI) INTERM-LOW:	-1,7 (-3 to -0,4)	-4,6 (-5,9 to -3,2)
T(0.05,207/203):	1,971488928	1,971720849

	DASH-DIET (N=208)	Control diet (N=204)
SD estimated from UCI (h_i):	9,510003515	9,417009474
SD estimated from UCI (i_l):	9,510003515	10,14139482
SD estimated from LCI (h_i):	9,510003515	9,417009474
SD estimated from LCI (i_l):	9,510003515	9,417009474
SD pooled:	9,510003515	9,59810581
Variance of Difference var.:	90,44016686	92,12363515
W-S variance	45,22008343	46,06181757
W-S variance pooled	45,6409505	
Reliability (R) from baseline	0,543590495	

		DASH-Control diets	
t(0.05,410):	1,965768206	SD_UCI	SD_LCI
Mean (95% CI) DASH-Control (high):	-5,9 (-8 to -3,7)	11,35765095	10,84139409
Mean (95% CI) DASH-Control (inter):	-5 (-7,6 to -2,5)	12,90642153	13,42267839
Mean (95% CI) DASH-Control (low):	-2,2 (-4,4 to -0,1)	10,84139409	11,35765095
	SD pooled:	11,787865	
	(DASH-Control)_Variance	138,9537612	
	B-S variance	93,31281069	
	Reliability (R) from diff.	0,671538574	

3.3.2.- CÀLCULS A PARTIR D'INFORMACIÓ REBUDA

Mechanism of Cocaine hyperthemia in Humans [30] (Ann Intern Med)

Measurement:	ESOPHAGEAL INTERNAL TEMPERATURE
Before heat stress variance:	0,062613187
At the end of heat stress variance:	0,145442181
B-S variance from baseline	0,058456012
B-S variance from final resp	0,141285006
Standard Error from Difference var:	0,09118306
St. Dev of Difference var.:	0,241247701
variance of Difference var.:	0,058200453
W-S variance	0,029100227
Reliability (R) from baseline	0,535238055
Reliability (R) from final resp	0,799918935

3.4.- RESULTATS MÉS IMPORTANTS

Pels 21 articles amb resposta contínua i analitzats segons la teoria de la Normalitat, la matriu de variàncies-covariàncies a nivell teòric i pels dos tractaments és:

	Per1_B	Per1_F	Per2_B	Per2_F
Per1_B	$\sigma_\tau^2+\sigma_e^2$	σ_τ^2	σ_τ^2	σ_τ^2
Per1_F	σ_τ^2	$\sigma_\tau^2+\sigma_e^2$	σ_τ^2	σ_τ^2
Per2_B	σ_τ^2	σ_τ^2	$\sigma_\tau^2+\sigma_e^2$	σ_τ^2
Per2_F	σ_τ^2	σ_τ^2	σ_τ^2	$\sigma_\tau^2+\sigma_e^2$

Ja s'ha vist que la variància de $Y_{c\text{-}o}$ és $2\sigma_e^2$. Si el model complert per a la resposta observada, y_{ijk}, ajustant pel valor de la covariable, x_{ijk}, vindria modelat per:

$$y_{ijk}=\mu+\phi\cdot x_{ijk}+\tau_{ik}+\pi_j+\alpha_{d[i,j]}+\lambda_{d[i,j-1]}+\varepsilon_{ijk}$$

aleshores, la resposta analitzada ("*cross-over differences*", procedint d'igual forma amb les "*period differences*") via anàlisi de la covariància serà:

$$y_{\text{cross-over differences}}=\begin{cases} y_{AB1k}-y_{AB2k}=(\alpha_A-\alpha_B)+(\pi_1-\pi_2)+(\phi x_{AB1k}-\phi x_{AB2k})+(\varepsilon_{AB1k}-\varepsilon_{AB2k}) \\ y_{BA2k}-y_{BA1k}=(\alpha_A-\alpha_B)+(\pi_2-\pi_1)+(\phi x_{BA2k}-\phi x_{BA1k})+(\varepsilon_{BA2k}-\varepsilon_{BA1k}) \end{cases}$$

i, per tant, la variància d'aquestes variables serà:

$$\text{Var}(Y_{C-O})=2\sigma_e^2(1+\phi^2)$$

Notar que per ϕ=1, si la covariable és la "*baseline*" es correspon amb la variància de la resposta quan s'estudia el canvi respecte dels valors inicials, obtenint $4\sigma_e^2$.

D'aquests 21, un 42,86% presentarien eficiència "nul·la" en quant a presentació de resultats per no proporcionar la informació referent a la diferència d'efectes i, per tant, no beneficiar-se del disseny aparellat.

De la resta d'articles, els paràmetres que s'han calculat es separen segons la procedència dels càlculs de la variància "*between*": resposta final (F) i/o "*baseline*" (B); o a partir de la diferència de dos grups d'estudi independents: "Habitual Diet" i "Dietary Treatment" de l'article (149) de *JAMA* i "Control Diet" i "Dash diet" de l'article (334) de *N Engl J Med.*

A la taula següent es mostren els resultats:

Codi MEDLINE	Núm. Article	Revista	Mètode Càlcul	N	σ^2_τ	σ^2_e	R	'm'	Efficiency
206	1	BMJ	B	22	1,181	0,029	0,976	0,012	42,308
			F		1,000		0,972	0,014	35,962
			B	19	0,772	0,038	0,953	0,024	21,178
			F		1,220		0,970	0,015	32,891
30	2	Annals	B	7	0,058	0,029	0,668	0,166	3,009
			F		0,141		0,829	0,085	5,855
240	3	Annals	F	41	3,789	2,459	0,606	0,197	2,541
497	4	Annals	B	49	0,613	0,110	0,848	0,076	6,584
			F		0,607		0,847	0,077	6,534
511	5	Annals	F	40	0,00018	0,00006	0,755	0,123	4,081
363	6	LANCET	B	59	0,147	0,031	0,826	0,087	5,738
144	7	JAMA	B	105	131,647	248,603	0,346	0,327	1,530
			F		76,397		0,567	0,217	2,307
149	8	JAMA	B	60	321,187	187,313	0,632	0,184	2,715
			D		82,687		0,306	0,347	1,441
131	9	AMJMED	F	13	30334355,010	11915644,990	0,718	0,141	3,546
334	10	NEJM	B	412	54,359	45,641	0,544	0,228	2,191
			D		93,313		0,672	0,164	3,044
388	11	NEJM	B	7	0,033	0,007	0,827	0,086	5,787
			F		0,017		0,708	0,146	3,419
481	12	NEJM	F	13	816,589	208,411	0,797	0,102	4,918

Mètode de Càlcul:

B: variància entre-individus calculada a partir resposta basal

F: variància entre-individus calculada a partir resposta final

D: variància entre-individus calculada a partir diferència de grups independents (paral·lel)

Fiabilitat (R): proporció de variància deguda a la variabilitat entre-individus

Raó de Mostres (*m*): quociente entre la grandària mostral del disseny "*cross-over*" i la del disseny de grups paral·lels

Eficiència: quocient entre el cost del disseny paral·lel i del "*cross-over*", assumint unitari el cost del període

Els valors elevats de l'eficiència són indicadors de què el disseny amb intercanvi del tractament és molt adequat per al tipus de dades analitzades ja que el cost del disseny si hagués estat planificat com a paral·lel es dispararia per obtenir la mateixa precisió en les estimacions. Destacaríem, entre els articles estudiats, el (206) de *BMJ*, del qual cal destacar que les estimacions puntuals (i per això la divergencia) de la variància entre- segons el seu mètode de càlcul es poden considerar similars, no violant la premissa efectuada al model d'efectes fixos del tractament (el tractament afecta a la mitjana però no a la variància de la resposta).

D'altra banda, els valors propers a la unitat posen de manifest que els riscs del "*cross-over*" (pèrdua de pacients, "*carry*-over", etc.) no compensen el guany en quant a cost del "*cross-over*" respecte del disseny de grups paral·lels i, per tant, caldria sospesar

els avantatges i inconvenients d'ambdós dissenys. Aquest cas es pot aplicar als articles de la revista mèdica *JAMA*: (144), quan la variància entre-individus es calcula a partir de la variància de la resposta basal, i al (149), quan aquesta es calcula a partir de la variància de la diferència entre dietes.

La resta de valors de l'eficiència relativa intermitjos, no tan extrems, mostren la conveniència d'utilitzar el disseny amb intercanvi del tractament havent de duplicar, triplicar,... el cost en cas de fer servir el disseny de grups paral·lels.

Si ens fixem en les variàncies de la resposta final i basal per l'article (144) de *JAMA*, s'observa que la variància de la resposta basal és bastant superior a la variància de la resposta final i fa pensar en que potser caldria aplicar alguna transformació (logarítmica) a les dades que permeti modelar aquesta reducció de la variància amb un efecte proporcional.

D'altra banda, la discrepància entre les variàncies entre- observada en l'article (30) de l'*Ann Interm Med* segons s'estimin a partir de la variància de la resposta final o basal es pot atribuir a que la premissa efectuada sobre el tractament, que és un efecte fix no es compleix: per tant, existeix una interacció pacient-tractament que aconsellaria incloure l'efecte tractament com aleatori..

Pels articles (149) de *JAMA* i (334) de *N Engl J Med*, les estimacions de la variància entre-individus es calculen a partir del valor de la variància de la mesura basal i de la diferència entre dietes en ambdós casos:

- La nostra interpretació per l'article (149) és que la diferència entre dietes es realitza a nivell entre-individual (disseny paral·lel), però recordem que té en compte l'ajust dels valors basals com a covariable, resultant la variància intra- superior a la entre-.

  Per tant, en aquest cas, no hi ha benefici en la planificació del "*cross-over*" quan s'ajusta per la covariable, perquè augmenta la variància com s'ha vist a l'inici d'aquest apartat.

- Per a l'article (334) de *N Engl J Med*, en canvi, no és tanta la disparitat entre els valors obtinguts per a les variàncies calculades a partir de la variabilitat de la resposta basal i de la diferència entre dietes.

Ara, el resultat és pitjor per a la comparació mitjançant el modelatge estadístic que simplement el "*cross-over*" sense ajustar. L'explicació podria trobar-se en que s'ajusten les dades incloent-hi termes que podrien estar correlacionats (efectes principals i interaccions entre ells) i, això implica colinealitat que augmentaria la variància de les estimacions, amb intervals de confiança més amplis.

A continuació, es mostren les Figures 2, 3 referents a l'eficiència del "*cross-over*" respecte el paral·lel separades per revista i tipus de resposta a partir de la qual es calcula la variància "*between*", respectivament. La utilitat de la Figura 4 (que es mostra seguidament) és poder observar les discrepàncies o similituds entre valors de fiabilitat basats en l'estimació de la variància "*between*" a partir de la variància dels diferents tipus de resposta esmentats: Basal (B), Final (F) i Diferència entre grups independents(D).

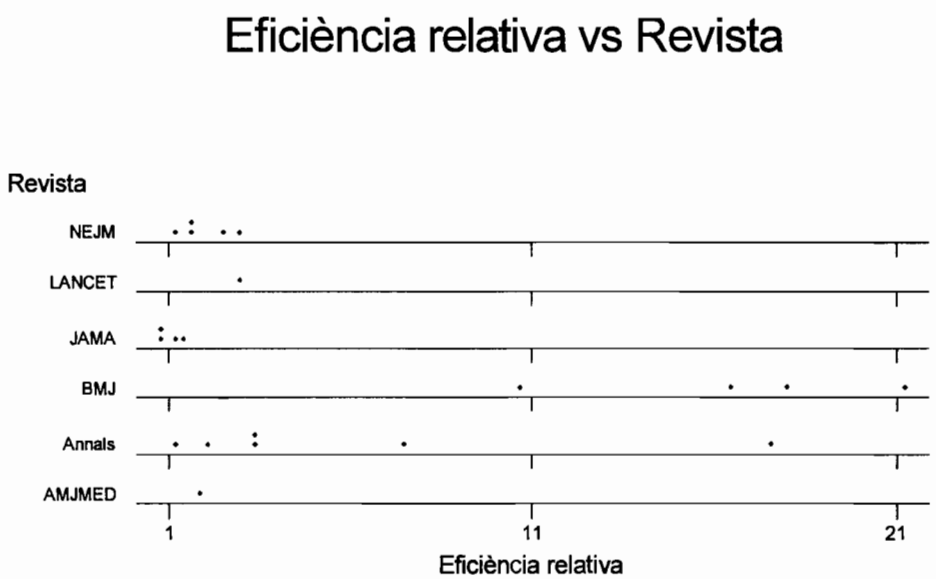


Figura 2

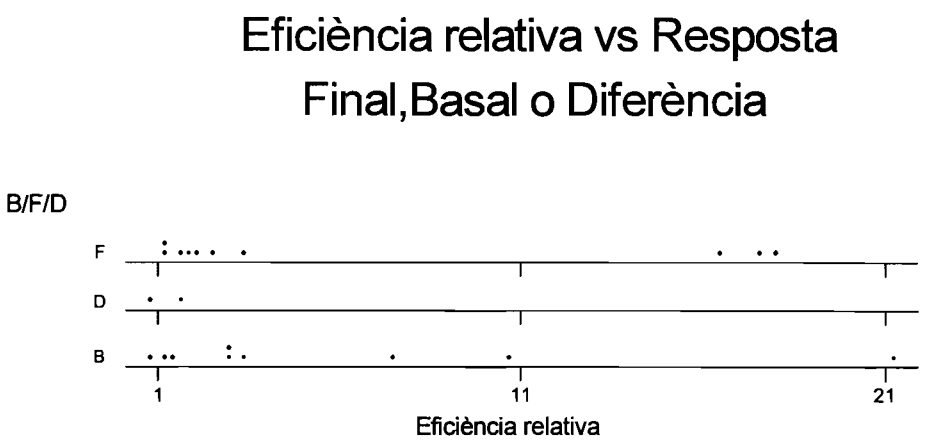


Figura 3

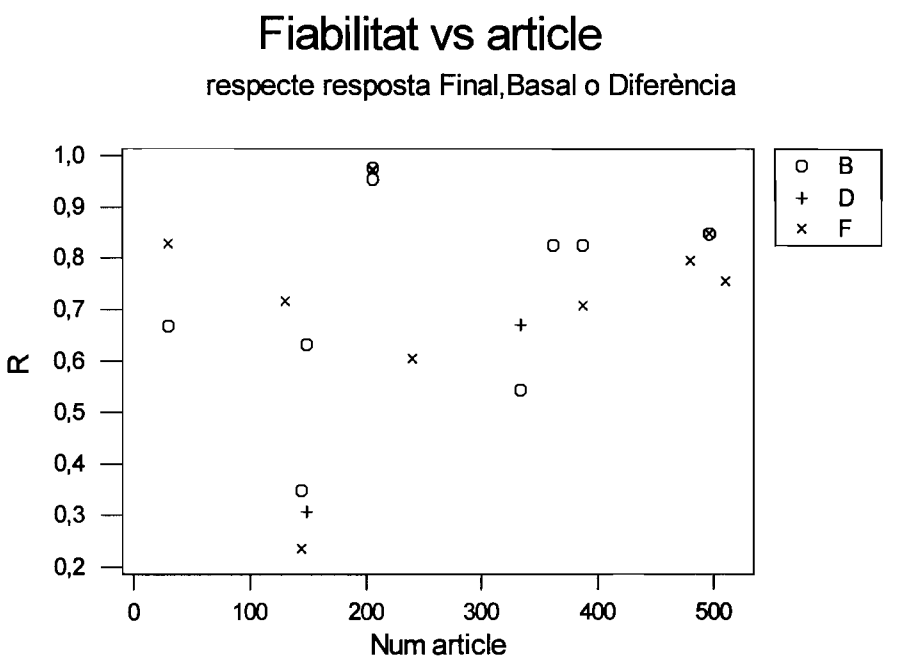


Figura 4

Finalment, es pot concloure que en els casos en què es compleixen les premisses de l'efecte fix del tractament sobre la variable de resposta, les estimacions de la variància entre-individus a partir de la variabilitat de la respota basal o final i, per tant, la resta de paràmetres basats en aquestes estimacions són estables. Però, en canvi, aquest resulat no es dóna si el model ajustat considera fix l'efecte directe del tractament quan és aleatori, inclou al model efectes correlacionats o la mesura basal com a covariable.

4.- CONCLUSIONS

L'eficiència relativa, del "*cross-over*" respecte del disseny de grups paral·lels obtinguda als 21 articles que compleixen el criteri d'inclusió, podríem separar-la en dos grups ben diferenciats:

- Un 42.86% (9 de 21) dels articles presenten eficiència nul·la pel que fa a la presentació dels resultats, per no proporcionar la informació sobre la diferència d'efectes dels tractaments bé via mitjana i desviació tipus o bé intervals de confiança a partir dels quals es deriva l'estimació de les variàncies entre i intra-individus (tot i que la CONSORT obliga a reportar els intervals de confiança).

- Del 57.14% (12 de 21) restant:

    - Un 8.3% (1 de 12) presenten valors de l'eficiència relativa compresos entre 21 i 43; que implica una reducció substancial del cost del disseny amb intercanvi del tractament respecte del disseny de grups paral·lels (aquesta relació d'equivalència és també aplicable a la variància dels estimadors).

    - D'un 58,3% (7 de 12), s'obtenen valors de l'eficiència relativa entre 3 i 7 que no implica un guany tan important com en el descrit anteriorment però suposaria augmentar d'entre tres i set vegades el cost del disseny paral·lel respecte del del "*cross-over*", perquè els estimadors fossin gairebé iguals en quant a precisió.

    - El 33.3% restant (4 de 12) presenten valors de l'eficiència més petits que 3 havent de duplicar o triplicar el cost del disseny "*cross-over*" per obtenir la mateixa eficiència en les estimacions si es planifica un disseny de grups paral·lels.

Si s'obtinguessin valors de l'eficiència propers a la unitat no hi hauria cap benefici addicional en planificar un disseny o l'altre, l'elecció del tipus de disseny podria dependre de factors com: l'econòmic (en contra del disseny de grups paral·lels degut a la major grandària de mostra requerida), la disponibilitat dels individus i el problema de les pèrdues (inconvenient més accentuat pel "*cross-over*"), la dificultat d'anàlisi del disseny, els efectes addicionals (període, efecte arrossegat,...) del disseny amb intercanvi de tractaments, el tipus de malaltia que s'estudia.

5.- BIBLIOGRAFIA

(1) Senn, Stephen (1993). A simple analysis ignoring the effect fo period, *Cross-over Trials in Clinical Research*, 3.2 33-37.

(2) Senn, Stephen (1993). Adjusting for a period effect: two-sample t approach, *Cross-over Trials in Clinical Research*, 3.5 41-42.

(3) Senn, Stephen (1993). Computer analysis, *Cross-over Trials in Clinical Research*, 3.16 60-68.

(4) Fleiss, J. L. (1986). Reliability of measurement, *The design and analysis of clinical experiments*, 1 1-32.

(5) Fleiss, J. L. (1986). The cross-over study, *The design and analysis of clinical experiments*, 10 263-90.

(6) Jones, Byron; Kenward, Michael G. (1989) The 2x2 cross-over trial with continuous data, *Design and analysis of cross-over trials (Monographs on statistics and applied probability)*, 2 16-88.

(7) Machin, David; Campbell, Michael; Fayers, Peter; Pinol, Alan. (second edition, 1987). *Sample size tables for clinical studies.*

(8) Prat Bartés, Albert; Tort-Martorell Llabrés, Xavier; Grima Cintas, Pere; Pozueta Fernández, Lourdes. (1997) Comparación de dos tratamientos, *Métodos estadísticos. Control y mejora de la calidad*, 5 85-104.

(9) Prat Bartés, Albert; Tort-Martorell Llabrés, Xavier; Grima Cintas, Pere; Pozueta Fernández, Lourdes. (1997) Comparación de más de dos tratamientos: análisis de la varianza, *Métodos estadísticos. Control y mejora de la calidad*, 6 105-126.

(10) Apunts "*Diseños EC con intercambio*" de Bioestadística, Erik Cobo, EIO, FME, UPC.

(11) Apunts "*Diseños EC específicos*" de Bioestadística, Erik Cobo, EIO, FME, UPC.

6.- ANNEX

6.1.- Llista d'articles de les revistes citades.

THE AMERICAN JOURNAL OF MEDICINE (Am J Med)

131. Jonkers IJ. Mohrschladt MF. Westendorp RG. van der Laarse A. Smelt AH. Severe hypertriglyceridemia with insulin resistance is associated with systemic inflammation: reversal with bezafibrate therapy in a randomized controlled trial. [Clinical Trial. Journal Article. Randomized Controlled Trial] American Journal of Medicine. 112(4):275-80, 2002 Mar.
UI: 11893366

• • •

212. Spertus JA. Jones PG. Coen M. Garg M. Bliven B. O'Keefe J. March RJ. Horvath K. Transmyocardial CO(2) laser revascularization improves symptoms, function, and quality of life: 12-month results from a randomized controlled trial.[comment]. [Clinical Trial. Journal Article. Multicenter Study. Randomized Controlled Trial] American Journal of Medicine. 111(5):341-8, 2001 Oct 1.
UI: 11583635

• • •

377. Ceresini G. Freddi M. Morganti S. Rebecchi I. Modena AB. Rinaldi M. Manca C. Amaducci A. Del Rio G. Valenti G. The effects of transdermal estradiol on the response to mental stress in postmenopausal women: a randomized trial. [Clinical Trial. Journal Article. Randomized Controlled Trial] American Journal of Medicine. 109(6):463-8, 2000 Oct 15.
UI: 11042235

• • •

409. Lipworth BJ. Dempsey OJ. Aziz I. Wilson AM. Effects of adding a leukotriene antagonist or a long-acting beta(2)-agonist in asthmatic patients with the glycine-16 beta(2)-adrenoceptor genotype. [Clinical Trial. Journal Article. Randomized Controlled Trial] American Journal of Medicine. 109(2):114-21, 2000 Aug 1.
UI: 10967152

• • •

ANNALS OF INTERNAL MEDICINE (Ann Intern Med)

30. Crandall CG. Vongpatanasin W. Victor RG. Mechanism of cocaine-induced hyperthermia in humans.[comment][summary for patients in Ann Intern Med. 2002 Jun 4;136(11):I20; PMID: 12044141]. [Clinical Trial. Journal Article. Randomized Controlled Trial] Annals of Internal Medicine. Online. 136(11):785-91, 2002 Jun 4.
UI: 12044126

• • •

93. Crandall CG. Vongpatanasin W. Victor RG. Mechanism of cocaine-induced hyperthermia in humans.[summary for patients in Ann Intern Med. 2002 Jun 4;136(11):I20; PMID: 12044141]. [Clinical Trial. Journal Article. Randomized Controlled Trial] Annals of Internal Medicine. 136(11):785-91, 2002 Jun 4.
UI: 12044126

• • •

119. Anonymous. Summary for patients. Splitting the evening insulin dose to avoid low blood sugar levels and to improve sugar control in patients with type 1 diabetes.[original report in Ann Intern Med. 2002 Apr 2;136(7):504-14; PMID: 11926785]. [Clinical Trial. Patient Education Handout. Randomized Controlled Trial] Annals of Internal Medicine. 136(7):I29, 2002 Apr 2.
UI: 11926807

• • •

120. Fanelli CG. Pampanelli S. Porcellati F. Rossetti P. Brunetti P. Bolli GB. Administration of neutral protamine Hagedorn insulin at bedtime versus with dinner in type 1 diabetes mellitus to avoid nocturnal hypoglycemia and improve control. A randomized, controlled trial.[comment][summary for patients in Ann Intern Med. 2002 Apr 2;136(7):I29; PMID: 11926807]. [Clinical Trial. Journal Article. Randomized Controlled Trial] Annals of Internal Medicine. 136(7):504-14, 2002 Apr 2.
UI: 11926785

• • •

240. Eijkhout HW. van Der Meer JW. Kallenberg CG. Weening RS. van Dissel JT. Sanders LA. Strengers PF. Nienhuis H. Schellekens PT. Inter-University Working Party for the Study of Immune Deficiencies. The effect of two different dosages of intravenous immunoglobulin on the incidence of recurrent infections in patients with primary hypogammaglobulinemia. A randomized, double-blind, multicenter crossover trial. [Clinical Trial. Journal Article. Randomized Controlled Trial] Annals of Internal Medicine. 135(3):165-74, 2001 Aug 7.
UI: 11487483

• • •

268. Fuentes F. Lopez-Miranda J. Sanchez E. Sanchez F. Paez J. Paz-Rojas E. Marin C. Gomez P. Jimenez-Pereperez J. Ordovas JM. Perez-Jimenez F. Mediterranean and low-fat diets improve endothelial function in hypercholesterolemic men. [Clinical Trial. Journal Article. Randomized Controlled Trial] Annals of Internal Medicine. 134(12):1115-9, 2001 Jun 19.
UI: 11412051

• • •

319. Cefalu WT. Skyler JS. Kourides IA. Landschulz WH. Balagtas CC. Cheng S. Gelfand RA. Inhaled Insulin Study Group. Inhaled human insulin treatment in patients with type 2 diabetes mellitus.[comment]. [Clinical Trial. Clinical Trial, Phase II. Journal Article. Randomized Controlled Trial] Annals of Internal Medicine. 134(3):203-7, 2001 Feb 6.
UI: 11177333

• • •

447. Swan SK. Rudy DW. Lasseter KC. Ryan CF. Buechel KL. Lambrecht LJ. Pinto MB. Dilzer SC. Obrda O. Sundblad KJ. Gumbs CP. Ebel DL. Quan H. Larson PJ. Schwartz JI. Musliner TA. Gertz BJ. Brater DC. Yao SL. Effect of cyclooxygenase-2

inhibition on renal function in elderly persons receiving a low-salt diet. A randomized, controlled trial.[comment]. [Clinical Trial. Journal Article. Randomized Controlled Trial] Annals of Internal Medicine. 133(1):1-9, 2000 Jul 4.
UI: 10877734

• • •

497. Zambon D. Sabate J. Munoz S. Campero B. Casals E. Merlos M. Laguna JC. Ros E. Substituting walnuts for monounsaturated fat improves the serum lipid profile of hypercholesterolemic men and women. A randomized crossover trial.[erratum appears in Ann Intern Med 2000 Oct 17;133(8):659]. [Clinical Trial. Journal Article. Randomized Controlled Trial] Annals of Internal Medicine. 132(7):538-46, 2000 Apr 4.
UI: 10744590

• • •

511. Weiler JM. Bloomfield JR. Woodworth GG. Grant AR. Layton TA. Brown TL. McKenzie DR. Baker TW. Watson GS. Effects of fexofenadine, diphenhydramine, and alcohol on driving performance. A randomized, placebo-controlled trial in the Iowa driving simulator.[comment]. [Clinical Trial. Journal Article. Randomized Controlled Trial] Annals of Internal Medicine. 132(5):354-63, 2000 Mar 7.
UI: 10691585

• • •

THE BRITISH MEDICAL JOURNAL (BMJ)

5. Niederhofer H. Staffen W. Mair A. Galantamine may be effective in treating autistic disorder.[comment]. [Clinical Trial. Comment. Letter. Randomized Controlled Trial] BMJ. 325(7377):1422, 2002 Dec 14.
UI: 12480867

• • •

39. Lenhardt R. Seybold T. Kimberger O. Stoiser B. Sessler DI. Local warming and insertion of peripheral venous cannulas: single blinded prospective randomised controlled trial and single blinded randomised crossover trial.[comment]. [Clinical Trial. Journal Article. Randomized Controlled Trial] BMJ. 325(7361):409, 2002 Aug 24.
UI: 12193353

• • •

63. Lenhardt R. Seybold T. Kimberger O. Stoiser B. Sessler DI. Local warming and insertion of peripheral venous cannulas: single blinded prospective randomised controlled trial and single blinded randomised crossover trial. [Clinical Trial. Journal Article. Randomized Controlled Trial] British Medical Journal. 325(7361):409, 2002 Aug 24.
UI: 12193353

• • •

196. Sahota P. Rudolf MC. Dixey R. Hill AJ. Barth JH. Cade J. Randomised controlled trial of primary school based intervention to reduce risk factors for obesity.[comment]. [Clinical Trial. Journal Article. Multicenter Study. Randomized Controlled Trial] BMJ. 323(7320):1029-32, 2001 Nov 3.
UI: 11691759

• • •

197. Sahota P. Rudolf MC. Dixey R. Hill AJ. Barth JH. Cade J. Evaluation of implementation and effect of primary school based intervention to reduce risk factors for obesity.[comment]. [Clinical Trial. Journal Article. Multicenter Study. Randomized Controlled Trial] BMJ. 323(7320):1027-9, 2001 Nov 3.
UI: 11691758

• • •

206. Pollock MA. Sturrock A. Marshall K. Davidson KM. Kelly CJ. McMahon AD. McLaren EH. Thyroxine treatment in patients with symptoms of hypothyroidism but thyroid function tests within the reference range: randomised double blind placebo controlled crossover trial. [Clinical Trial. Journal Article. Randomized Controlled Trial] BMJ. 323(7318):891-5, 2001 Oct 20.
UI: 11668132

• • •

282. Allan L. Hays H. Jensen NH. de Waroux BL. Bolt M. Donald R. Kalso E. Randomised crossover trial of transdermal fentanyl and sustained release oral morphine for treating chronic non-cancer pain.[comment]. [Clinical Trial. Journal Article. Multicenter Study. Randomized Controlled Trial] BMJ. 322(7295):1154-8, 2001 May 12.
UI: 11348910

• • •

336. Schrader H. Stovner LJ. Helde G. Sand T. Bovim G. Prophylactic treatment of migraine with angiotensin converting enzyme inhibitor (lisinopril): randomised, placebo controlled, crossover study. [Clinical Trial. Journal Article. Randomized Controlled Trial] BMJ. 322(7277):19-22, 2001 Jan 6.
UI: 11141144

• • •

408. Stone S. Kibbler C. How A. Balstrini A. Feedback is necessary in strategies to reduce hospital acquired infection.[comment]. [Comment. Letter] BMJ. 321(7256):302-3, 2000 Jul 29.
UI: 10979704

• • •

442. Emery J. Walton R. Murphy M. Austoker J. Yudkin P. Chapman C. Coulson A. Glasspool D. Fox J. Computer support for interpreting family histories of breast and ovarian cancer in primary care: comparative study with simulated cases.[comment]. [Journal Article] BMJ. 321(7252):28-32, 2000 Jul 1.
UI: 10875832

• • •

532. R]oggla G. Moser B. R]oggla M. Effect of temazepam on ventilatory response at moderate altitude. [Clinical Trial. Letter. Randomized Controlled Trial] BMJ. 320(7226):56, 2000 Jan 1.
UI: 10617541

• • •

THE JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION (JAMA)

13. Detmar SB. Muller MJ. Schornagel JH. Wever LD. Aaronson NK. Health-related quality-of-life assessments and patient-physician communication: a randomized controlled trial. [Clinical Trial. Journal Article. Randomized Controlled Trial] JAMA. 288(23):3027-34, 2002 Dec 18.
UI: 12479768

• • •

16. Parienti JJ. Thibon P. Heller R. Le Roux Y. von Theobald P. Bensadoun H. Bouvet A. Lemarchand F. Le Coutour X. Bensadoun H. Antisepsie Chirurgicale des mains Study Group. Hand-rubbing with an aqueous alcoholic solution vs traditional surgical hand-scrubbing and 30-day surgical site infection rates:
a randomized equivalence study.[comment][erratum appears in JAMA 2002 Dec 4;288(21):2689 Note: Bensadoun Herve [corrected to Bensadoun Henri]]. [Clinical Trial. Journal Article. Randomized Controlled Trial] JAMA. 288(6):722-7, 2002 Aug 14.
UI: 12169076

• • •

70. Parienti JJ. Thibon P. Heller R. Le Roux Y. von Theobald P. Bensadoun H. Bouvet A. Lemarchand F. Le Coutour X. Antisepsie Chirurgicale des mains Study Group. Hand-rubbing with an aqueous alcoholic solution vs traditional surgical hand-scrubbing and 30-day surgical site infection rates: a randomized equivalence study. [Clinical Trial. Journal Article. Randomized Controlled Trial] JAMA. 288(6):722-7, 2002 Aug 14.
UI: 12169076

• • •

144. Arruda-Olson AM. Mahoney DW. Nehra A. Leckel M. Pellikka PA. Cardiovascular effects of sildenafil during exercise in men with known or probable coronary artery disease: a randomized crossover trial.[comment]. [Clinical Trial. Journal Article. Randomized Controlled Trial] JAMA. 287(6):719-25, 2002 Feb 13.
UI: 11851538

• • •

149. Jula A. Marniemi J. Huupponen R. Virtanen A. Rastas M. Ronnemaa T. Effects of diet and simvastatin on serum lipids, insulin, and antioxidants in hypercholesterolemic men: a randomized controlled trial. [Clinical Trial. Journal Article. Randomized Controlled Trial] JAMA. 287(5):598-605, 2002 Feb 6.
UI: 11829698

• • •

279. O'Shea JC. Hafley GE. Greenberg S. Hasselblad V. Lorenz TJ. Kitt MM. Strony J. Tcheng JE. ESPRIT Investigators (Enhanced Suppression of the Platelet IIb/IIIa Receptor with Integrilin Therapy trial). Platelet glycoprotein IIb/IIIa integrin blockade with eptifibatide in coronary stent intervention: the ESPRIT trial: a randomized controlled trial. [Clinical Trial. Journal Article. Multicenter Study. Randomized Controlled Trial] JAMA. 285(19):2468-73, 2001 May 16.
UI: 11368699

• • •

355. Denke MA. Adams-Huet B. Nguyen AT. Individual cholesterol variation in response to a margarine- or butter-based diet: A study in families. [Journal Article] JAMA. 284(21):2740-7, 2000 Dec 6.
UI: 11105179

• • •

384. Larson VD. Williams DW. Henderson WG. Luethke LE. Beck LB. Noffsinger D. Wilson RH. Dobie RA. Haskell GB. Bratt GW. Shanks JE. Stelmachowicz P. Studebaker GA. Boysen AE. Donahue A. Canalis R. Fausti SA. Rappaport BZ. Efficacy of 3 commonly used hearing aid circuits: A crossover trial. NIDCD/VA Hearing Aid Clinical Trial Group. [Clinical Trial. Controlled Clinical Trial. Journal Article. Multicenter Study] JAMA. 284(14):1806-13, 2000 Oct 11.
UI: 11025833

• • •

503. Collacott EA. Zimmerman JT. White DW. Rindone JP. Bipolar permanent magnets for the treatment of chronic low back pain: a pilot study.[comment]. [Clinical Trial. Journal Article. Randomized Controlled Trial] JAMA. 283(10):1322-5, 2000 Mar 8.
UI: 10714732

• • •

THE LANCET (LANCET)

341. The ESPRIT Investigators. Enhanced Suppression of the Platelet IIb/IIIa Receptor with Integrilin Therapy. Novel dosing regimen of eptifibatide in planned coronary stent implantation (ESPRIT): a randomised, placebo-controlled trial.[comment][erratum appears in Lancet 2001 Apr 28;357(9265):1370]. [Clinical Trial. Journal Article. Multicenter Study. Randomized Controlled Trial] Lancet. 356(9247):2037-44, 2000 Dec 16.
UI: 11145489

• • •

363. Brightling CE. Monteiro W. Ward R. Parker D. Morgan MD. Wardlaw AJ. Pavord ID. Sputum eosinophilia and short-term response to prednisolone in chronic obstructive pulmonary disease: a randomised controlled trial.[comment]. [Clinical Trial. Journal Article. Randomized Controlled Trial] Lancet. 356(9240):1480-5, 2000 Oct 28.
UI: 11081531

• • •

370. Vernia P. Fracasso PL. Casale V. Villotti G. Marcheggiano A. Stigliano V. Pinnaro P. Bagnardi V. Caprilli R. Topical butyrate for acute radiation proctitis: randomised, crossover trial.[comment]. [Clinical Trial. Journal Article. Randomized Controlled Trial] Lancet. 356(9237):1232-5, 2000 Oct 7.
UI: 11072942

• • •

413. Cromheecke ME. Levi M. Colly LP. de Mol BJ. Prins MH. Hutten BA. Mak R. Keyzers KC. Buller HR. Oral anticoagulation self-management and management by a specialist anticoagulation clinic: a randomised cross-over comparison.[comment]. [Clinical Trial. Journal Article. Randomized Controlled Trial] Lancet. 356(9224):97-102, 2000 Jul 8.
UI: 10963245

• • •

421. Brutsche MH. Brutsche IC. Munawar M. Langley SJ. Masterson CM. Daley-Yates PT. Brown R. Custovic A. Woodcock A. Comparison of pharmacokinetics and systemic effects of inhaled fluticasone propionate in patients with asthma and healthy volunteers: a randomised crossover study.[comment]. [Clinical Trial. Journal Article. Randomized Controlled Trial] Lancet. 356(9229):556-61, 2000 Aug 12.
UI: 10950233

• • •

421. Brutsche MH. Brutsche IC. Munawar M. Langley SJ. Masterson CM. Daley-Yates PT. Brown R. Custovic A. Woodcock A. Comparison of pharmacokinetics and systemic effects of inhaled fluticasone propionate in patients with asthma and healthy volunteers: a randomised crossover study.[comment]. [Clinical Trial. Journal Article. Randomized Controlled Trial] Lancet. 356(9229):556-61, 2000 Aug 12.
UI: 10950233

• • •

480. van der Gaag MS. Ubbink JB. Sillanaukee P. Nikkari S. Hendriks HF. Effect of consumption of red wine, spirits, and beer on serum homocysteine.[comment]. [Clinical Trial. Letter. Randomized Controlled Trial] Lancet. 355(9214):1522, 2000 Apr 29.
UI: 10801179

• • •

493. Hoffman RE. Boutros NN. Hu S. Berman RM. Krystal JH. Charney DS. Transcranial magnetic stimulation and auditory hallucinations in schizophrenia. [Clinical Trial. Letter. Randomized Controlled Trial] Lancet. 355(9209):1073-5, 2000 Mar 25.
UI: 10744097

• • •

THE NEW ENGLAND JOURNAL OF MEDICINE (N Engl J Med)

6. Gottlieb SS. Fisher ML. Cardiac resynchronization therapy for heart failure.[comment]. [Comment. Letter] New England Journal of Medicine. Online. 347(22):1803-4; author reply 1803-4, 2002 Nov 28.
UI: 12462232

• • •

20. Wyse DG. Waldo AL. DiMarco JP. Domanski MJ. Rosenberg Y. Schron EB. Kellen JC. Greene HL. Mickel MC. Dalquist JE. Corley SD. The Atrial Fibrillation Follow-up Investigation of Rhythm Management (AFFIRM) Investigators. A comparison of rate control and rhythm control in patients with atrial fibrillation.[comment]. [Clinical Trial. Journal Article. Multicenter Study. Randomized Controlled Trial] New England Journal of Medicine. Online. 347(23):1825-33, 2002 Dec 5.
UI: 12466506

• • •

189. Anonymous. The safety of inactivated influenza vaccine in adults and children with asthma.[comment]. [Clinical Trial. Journal Article. Multicenter Study. Randomized Controlled Trial] New England Journal of Medicine. 345(21):1529-36, 2001 Nov 22.
UI: 11794219

• • •

220. The Deep-Brain Stimulation for Parkinson's Disease Study Group. Deep-brain stimulation of the subthalamic nucleus or the pars interna of the globus pallidus in Parkinson's disease.[comment]. [Clinical Trial. Journal Article. Multicenter Study. Randomized Controlled Trial] New England Journal of Medicine. 345(13):956-63, 2001 Sep 27.
UI: 11575287

• • •

309. Cazeau S. Leclercq C. Lavergne T. Walker S. Varma C. Linde C. Garrigue S. Kappenberger L. Haywood GA. Santini M. Bailleul C. Daubert JC. Multisite Stimulation in Cardiomyopathies (MUSTIC) Study Investigators. Effects of multisite biventricular pacing in patients with heart failure and intraventricular conduction delay.[comment]. [Clinical Trial. Journal Article. Multicenter Study. Randomized Controlled Trial] New England Journal of Medicine. 344(12):873-80, 2001 Mar 22.
UI: 11259720

• • •

334. Sacks FM. Svetkey LP. Vollmer WM. Appel LJ. Bray GA. Harsha D. Obarzanek E. Conlin PR. Miller ER 3rd. Simons-Morton DG. Karanja N. Lin PH. DASH-Sodium Collaborative Research Group. Effects on blood pressure of reduced dietary sodium and the Dietary Approaches to Stop Hypertension (DASH) diet. DASH-Sodium Collaborative Research Group.[comment]. [Clinical Trial. Journal Article. Multicenter Study. Randomized Controlled Trial] New England Journal of Medicine. 344(1):3-10, 2001 Jan 4.
UI: 11136953

• • •

356. Rai KR. Peterson BL. Appelbaum FR. Kolitz J. Elias L. Shepherd L. Hines J. Threatte GA. Larson RA. Cheson BD. Schiffer CA. Fludarabine compared with chlorambucil as primary therapy for chronic lymphocytic leukemia.[comment]. [Clinical Trial. Journal Article. Multicenter Study. Randomized Controlled Trial] New England Journal of Medicine. 343(24):1750-7, 2000 Dec 14.
UI: 11114313

• • •

388. Sack RL. Brandes RW. Kendall AR. Lewy AJ. Entrainment of free-running circadian rhythms by melatonin in blind people.[comment]. [Clinical Trial. Controlled Clinical Trial. Journal Article] New England Journal of Medicine. 343(15):1070-7, 2000 Oct 12.
UI: 11027741

• • •

414. Shifren JL. Braunstein GD. Simon JA. Casson PR. Buster JE. Redmond GP. Burki RE. Ginsburg ES. Rosen RC. Leiblum SR. Caramelli KE. Mazer NA. Transdermal testosterone treatment in women with impaired sexual function after oophorectomy.[comment]. [Clinical Trial. Journal Article. Multicenter Study. Randomized Controlled Trial] New England Journal of Medicine. 343(10):682-8, 2000 Sep 7.
UI: 10974131

• • •

422. van Hilten BJ. van de Beek WJ. Hoff JI. Voormolen JH. Delhaas EM. Intrathecal baclofen for the treatment of dystonia in patients with reflex sympathetic dystrophy.[comment]. [Clinical Trial. Journal Article. Randomized Controlled Trial] New England Journal of Medicine. 343(9):625-30, 2000 Aug 31.
UI: 10965009

• • •

481. Chandalia M. Garg A. Lutjohann D. von Bergmann K. Grundy SM. Brinkley LJ. Beneficial effects of high dietary fiber intake in patients with type 2 diabetes mellitus.[comment]. [Clinical Trial. Journal Article. Randomized Controlled Trial] New England Journal of Medicine. 342(19):1392-8, 2000 May 11.
UI: 10805824

• • •

528. Schein OD. Katz J. Bass EB. Tielsch JM. Lubomski LH. Feldman MA. Petty BG. Steinberg EP. The value of routine preoperative medical testing before cataract surgery. Study of Medical Testing for Cataract Surgery.[comment]. [Clinical Trial. Journal Article. Multicenter Study. Randomized Controlled Trial] New England Journal of Medicine. 342(3):168-75, 2000 Jan 20.
UI: 10639542

6.2.- Articles descartats.

NO ESTUDIS AMB INTERCANVI DEL TRACTAMENT

Transmyocardial CO(2) laser revascularization improves symptoms, function, and quality of life: 12-month results from a randomized controlled trial *212 (Am J Med)

Participants: Eligible patients had medically refractory symptoms, defined as Canadian Cardiovascular Society classifications III and IV, such as reversible ischemia on the left ventricular free wall, a left ventricular ejection fraction greater than 20%, and a coronary anatomy that was not amenable to traditional revascularization, as assessed by their attending therapy

n(transmyocardial revascularization)=93 n(medical management)=99

Methods: Patients were randomized to either transmyocardial revascularization or continued medical therapy. Those randomized to continued medications were eligible for crossover to transmyocardial revascularization after 6 months, or earlier if they developed unstable angina requiring 48 hours or more of intravenous anti-anginal therapy.

Main outcome: health status assessed with the Seattle Angina Questionnaire and the Short Form-36

Statistical analysis: (Primary analysis:) Student t test were used to compare serial scores and the area under the curve for each of the health status.

(Secondary analysis: patients assigned to medical theraphy crossed-over to the transmyocardial revascularization arm) Efficacy or on-treatment analysis to estimate the relative improvement associated with transmyocardial revascularizatin vs medial therapy alone. In this analysis, scores for medically treated patients were included only up to the time of cross-over. Changes in scores were calculated using the last precrossover value, and both the area under their health status curves and their benefit statistics were truncated at the time they crossed over.

NOT CROSS-OVER (RESCUE THERAPY)

Table 2.- Intention-to-treat and efficacy analyses comparing health status scores (range: 0 to 100) over time between Transmyocardial Revascularization and Medical Management (mean(SD))

Seattle Angina Questionnaire	Transmyocardial revascularization	Medical Management	
		Intention-to-treat	Efficacy analysis
Month 0 (BASELINES)	27 (23)	28 (23)	28 (23)
3	60 (30)	36 (27)	28 (26)
6	59 (31)	41 (31)	26 (26)
12	53 (32)	48 (31)	27 (27)

Splitting the evening insulin dose to avoid low blood sugar levels and to improve sugar control in patients with type 1 diabetes *119 (Ann Intern Med)

Service provided by "Annals" to help patients better understand the complicated and often mystifying language of modern medicine

FULL REPORT: "Administration of Neutral Protamine Hagedorn Insulin at Bedtime versus with Dinner in Type 1 Diabetes Mellitus to Avoid Nocturnal Hypoglycemia and Improve control: A randomized controlled trial" (120) 2 April 2002 (Volume 136, pages 504-514)

SUMMARY FOR PATIENTS

Randomised controlled trial of primary school based intervention to reduce risk factors for obesity *196 (BMJ)

N=636 children

Main outcome measures: Body mass index, diet, physical activity and psychological state

Design: Group-randomised controlled trial; 2 groups: Intervention schools (5) and comparison schools (5)

PARAL·LEL GROUP TRIAL

Table 1.- Characteristics at baseline of children in intervention and control schools. [mean (SD)]

	Intervention (n=314)	Comparison (n=322)
Body mass index SD score	0,12 (1,01)	0,04 (1,17)

Table 3.- Weighted mean difference in body mass index standard deviation score and...

School	Weighted mean difference (95%CI)	% weight of shool
1	0 (-0,2 to 0,1)	25,8
2	0,1 (0 to 0,2)	18
3	0,1 (-0,1 to 0,2)	22,5
4	-0,1 (-0,3 to 0)	19,8
5	-0,2 (-0,3 to 0)	13,9
Overall	0 (-0,1 to 0,1)	

Health-Related Quality-of-Life Assessments and Patient-Physician Communication: A Randomized Controlled Trial *13 (JAMA)

N = 10 physicians and 214 patients undergoing palliative chemotherapy

Main outcome measure: Audiotapes of the consultations were content analyzed to evaluate patient-physician communication. Physician's awareness of their patients' health problems was assessed by comparing physicians' and patients' ratings on the Dartmouth Primary Care Cooperative Information Functional Health Assessment (COOP) and the World Organisation Project of National Collegues and Academics (WONCA) charts

COMMENT: Although the use of a crossover in the study design enabled us to largely neutralize any effect that might be attributed to physicians' background characteristics, it also carried with it the risk of a carryover or contamination effect. In fact, some evidence was found suggesting that the physicians who began in the experimental condition and subsequently crossed over to the control condition may have been sensitized to HRQL issues and may have changed their behaviur during the period in which they were no longer exposed explicitly to the intervention.

CROSS-OVER TO CONTROL CONTAMINA-TION EFFECTS

RESPOSTA PRINCIPAL NO QUANTITATIVA

Administration of Neutral protamine hagedorn insulin at bedtime versus with dinner in type 1 diabetes mellitus to avoid nocturnal hypoglycemia and improve control *120 (Ann Intern Med)

N=22 C peptide-negative persons with type 1 diabetes mellitus

Primary end point: frequency of nocturnal hypoglycemia

Statistical analysis: McNemar test

Local warming and insertion of peripheral venous cannulas: single blinded prospective randomised controlled trial and single blinded randomised crossover trial *39 (BMJ)

N (inicial) = 42 patients with leukaemia who required chemotherapy [2 failed (both had initially been assigned to passive insulation)]

n(active/passive insulation)=21 n(passive/active insulation)=19

Primary outcome measure: Succes rate for insertion of 18 gauge cannula into vein on back of hand

Statistical analysis: the results in the warmed and unwarmed patients with leukaemia were compared with McNemar tests

Evaluation of implementation and effect of primary school based intervention to reduce risk factor for obesity *197 (BMJ)

Main outcome measures: Response RATE to questionaires, teacher's evaluation of training and input, succes of school action plans, content of school meals and children's knowledge of healthy living and self reported behaviour

Statistical analysis: Tables show Numbers and/or percentages

Randomized crossover trial of transdermal fentanyl and sustained release oral morphine for treating chronic non-cancer pain *282 (BMJ)

Main outcome measures: PREFERENCE for transdermal fentanyl or sustained release oral morphine, pain control, quality of life and safety assessment

Statistical analysis: binomial test ("preferred"/"not preferred")

Computer support for interpreting family histories of breast and ovarian cancer in primary care: comparative study with simulated cases *442 (BMJ)	
Main outcome measures:	number of appropiate management decisions made (maximum of 6), mean time taken to reach a decision, number of pedigrees accurately drawn (maximum of 6)
Statistical analysis:	Friedman's two way analysis of variance to compare effects overall for each outcome (median-range)
Table 1.- Median (range) outcome measures for 36 doctors managing family histories using three different methods of support	
Table 2.- Values are numbers (percentages)	

Hand-rubbing with an aqueous alcoholic solution us traditional surgical hand-scrubbing and 30-day surgical site infection rates: a randomized equivalence study *16 (JAMA)	
Primary end point:	Nosomical SSI rates (Surgical Site Infection)
Statistical analysis:	The 2 protocols were considered equivalent is the 95% CI of the SSI rate difference (calculated according to the Wallenstein method) was within the limits of -2% to 2% and contained the bound zero in both analyses, the as-treated as well as in the conservative intent-to-treat. The level of significance for equivalence was given by the highest P value related to the lowest Chi-square value of the continuity-corrected one-sided test (described by Dunnet and Gent).

Platelet glycoprotein IIb/IIIa integrin blockade with eptifibatide in coronary stent intervention *279 (JAMA)	
Primary end point:	rate of death, MI, urgent target vessel revascularization or bailout Gp IIb/IIIa inhibitor use at 48 h
Statistical analysis:	survival analysis methods were used for the 6-months analyses; pairwise comparison between 2 treatment groups were made using the long-rank test with event rates calculated by the kaplan-Meier method

Novel dosing regimen of eptifibatide in planned coronary stent implantation (ESPRIT): a randomised, placebo-controlled trial *341 (Lancet)	
Primary end point:	composite of death, myocardial infarction, urgent target vessel revascularisation and thrombotic bailout glycoprotein iib/IIIa therapy within 48h after randomisation
Statistical analysis:	pairwise comparisons by means of conventional Chi-square analyses. For secondary analyses, Chi-square for categorical variables and Wilcoxon rank-sum test for continuous variables.

The safety of inactivated influenza vaccine in adults and children with asthma *169 (N Engl J Med)
N=2032 patients with asthma (1952 receive both injections and completed both 14-day post-injection diaries)
Primary end point: exacerbation of asthma in the 2 weeks after the injections
Statistical analysis: 2-group test of the equivalence of binomial proportions

Fludarabine compared with chlorambucil as primary therapy for chronic lymphocytic leukemia *356 (N Engl J Med)	
Clinical responses:	complete remission, partial remission, complete or partial remission, and stable or progressive disease
Statistical analysis:	Chisquare test to compare the response rates. All time-to-event distribution were calculated by the Kaplan-Meier method and compared with the use of the log-rank test

The value of routine preoperative medical testing before cataract surgery *528 (N Engl J Med)	
2 types of crossover:	complete and partial
Principal outcomes:	cumulative rate of medical events; rate of medical events on the day of surgery; rate of events during post operative period
Statistical analysis:	stratified analysis

Effects of adding a Leukotriene antagonist or a long-acting beta(2)-agonist in asthmatic patients with the glycine-16 beta(2)-adrenoceptor genotype [409 (Am J Med)]

N=24 patients with moderate asthma

Primary end point: methacholine provocative dose transformed to a logscale to normalize the distribution (**geometric means**)

BASELINE: Metacholine provocative dose [Mean (SD)]62 (13)

Statistical analysis: The geometric mean metacholine provocative dose was determined for each treatment, 95% CI for treatment effects were also calculated. The geometric mean-fold protection ratio was calculated for each active treatment as the difference between the effects after 1 week of treatment vs baseline, in comparison with placebo. A multifactorial analysis of variance was performed using subject, treatment, sequence, and visit as factors.

Results: For both formoterol and zafirlukast, there were significant ($p<0,05$) improvements in the geometric mean metacholine provocative doses for the difference after 1 week versus baseline, as compared with placebo. This amounted to a geometric mean 1.9-fold difference (95%CI: 1,2- to 2,9-fold) for formoterol and 1,5-fold difference (95%CI: 1,1- to 2,2-fold) for zafirlukast. There was no significant difference in the geometricmean methacholine provocative doses between the formoterol and zafirlukast groups after 1 week fo treatment, either as absolute values or as changes from baseline

Inhaled human insulin treatment in patients with type 2 diabetes mellitus [*319 (Ann Intern Med)]

N=26 participants

Measurements: Glycemic control (hemoglobin A1c level) obtained at baseline and monthly for 3 months. Pulmonary function tests were done at baseline and at the end of the study.

Design: Randomised, open-label, 3-month study consisting of a screening visit, a 4-week baseline lead-in phase, and a 12-week treatment phase.

Statistical analysis: Efficacy was assessed by the 12-week change in hemoglobin A(1C) level from baseline. The 95% CI for the mean of the changes was calculated on the basis of the SE and the sampling t-distribution. Data are presented as the mean +/- SD

Effect of cyclooxigenase-2 inhibition on renal function in elderly persons receiving a low-salt diet [*447 (Ann Intern Med)]

N = 15 patients in good general health who were receiving a sodium-restricted diet

Measurements: glomerular filtration rate

Design: Randomized three-period balanced single-dose cross-over study and a randomised, parallel-group, multiple-dose study.

Statistical analysis: analysis of covariance model. The final model included terms for patient, period, treatment, sequence, and predose value as covariates. The normality and homogeneity of variance assumptions of the analysis of covariance were tested.

Objective: Determine effect of rofecoxib (treatment)

GLOMERULAR FILTRATION RATE was calculated measuring inulin clearance before and after treatment

PEAK REDUCTION response in inulin clearance from baseline: maximal difference between the minimum glomerular filtration rate after drug administration and the mean glomerular filtration rate measured during urinary stabilization (baseline)

Bipolar permanent magnets for the treatment of chronic low back pain: a pilot study [*503 (JAMA)]

N=20 with stable low back pain

Primary outcome: Visual Analog Scale (VAS) (used to quantify pain intensity)

Statistical analysis: The difference between baseline (preapplication) and posttreatment VAS scores was analysed using repeated measures analysis of variance.

BASELINE: VAS (mean (SD)) 4,8 (2,2)

Table 2.- Baseline and **Change in VAS Scores after treatment**, by day (mean(SD))

Day	Baseline (pretreatment) Sham	Magnet
Monday	5,2 (2,3)	4,7 (2,8)
Wednesday	5,2 (2,5)	4,8 (2,8)
Friday	4,5 (2,5)	4,6 (3,1)
overall	5,0 (2,4)	4,7 (2,9)
Day	**Change after treatment Sham**	**Magnet**
Monday	-0,58 (2,4)	-0,41 (1,3)
Wednesday	-0,75 (1,1)	-0,80 (1,1)
Friday	0,36 (1,5)	-0,31 (1,4)
overall	-0,44 (1,4)	-0,49 (0,96)

Intrathecal baclofen for the treatment of dystonia in patients with reflex sympathetic dystrophy *422 (N Engl J Med)

N=7 women who had had reflex sympathetic dystrophy

Main outcome measure:

Main outcome measure: Before and 4 and 8 hours after each injection, each woman assessed the overall severity of her dystonia on a 100-mm VAS on which the severity of symptoms was rated from 0 to 100. Changes in dystonia were calculated as the percentage change form the score at baseline to the lower score at 4 and 8 hours

Statistical analysis: To evaluate the differences in outcome among the 3 injections of baclofen and the 2 injections of placebo: repeated-measures analysis of variance

Table 2.- Percent change from base line in scores of the VAS after single intrathecal injections of placebo or baclofen in **the** 7 women with reflex sympathetic distrophy (Mean (SD))

Patient	Placebo 1rst injection	2nd injection	
1	-5	3	
2	-2	1	
3	-38	-19	
4	-9	-3	
5	-4	3	
6	9	4	
7	21	-47	
All patients	-4 (18)	-8 (19)	
Patient	**Baclofen** 25 (mu)g	50 (mu)g	75 (mu)g
1	4	-31	-51
2	-33	-67	-87
3	-51	-89	-88
4	1	-85	-3
5	-24	-36	-64
6	-6	-17	-42
7	-17	-	-42
All patients	-18 (20)	-54 (30)	-54 (30)

"RESEARCH LETTERS"

Galantamine may be effective in treating autistic disorder [Letters] *5 (BMJ)

Feedback is necessary in strategies to reduce hospital acquired infection (Editor) *408 (BMJ)

Outcome: Incidence of 'Clostridium difficile' associated diarrhoea

Cardiac Resynchronization Therapy for Heart Failure *6 (N Engl J Med) — NOT CROSS-OVER (RESCUE THERAPY)

Correspondence: Ten patients crossed over to pacing; in seven of these patients, the reason was an exacerbation of chronic heart failure.

Original article: **Cardiac Resynchronization in Chronic Heart Failure**

Study design: After this initial evaluation, patients underwent implantation of a cardiac-resynchronization device (InSync model 8040, Medtronic) along with three pacing leads: a standard right atrial lead, a standard right ventricular lead, and a specialized

leads: a standard right atrial lead, a standard right ventricular lead, and a specialized left ventricular lead, [25] which was placed into a distal cardiac vein by way of the coronary sinus through a guiding catheter. Patients who had undergone successful implantation were randomly assigned to atrial-synchronized biventricular pacing (the resynchronization group) or to a control group (no pacing) for six months, during which time medications for heart failure were to be kept constant. Randomization occurred in permuted blocks to ensure a balance between groups within centers. Base-line variables were reevaluated one, three, and six months after randomization.

Crossover from the control mode to the cardiac-resynchronization mode before the six-month assessment was prohibited, except for patients in whom a bradyarrhythmia that required cardiac pacing developed. Neither the patients nor the physicians treating them for heart failure and performing the study evaluations were aware of the treatment assignment

"WILCOXON SIGNED-RANK TEST"

Prophylactic treatment of migraine with angiotensin converting enzyme inhibitor (lisipronil):randomized, placebo controlled, crossover study *336 (BMJ)

Primary end points:	number of hours with headache (not normally distributed); number of days with headache and number of days with migraine
Statistical analysis:	Wilcoxon signed rank test to compare end point variables. For comparison of adverse events and acceptability, McNemar's matched pairs test

Oral anticoagulation self-management and management by a specialist anticoagulation clinic: a randomised cross-over comparison *413 (Lancet)

Primary end point:	number of measurements that were within the OS INR units from the therapeutic target value INR during each study period
Statistical analysis:	For comparison of groups the Wilcoxon singned-rank test and 2-tailed Fisher's exact test

Deep-brain stimulation of the subthalamic nucleus or the pars interna of the globus pallidus in parkinson's disease *220 (N Engl J Med)

N=143 patients with advanced parkinson's disease; 134 received bilateral implants in the subthalamic nucleus or the pars interna of the globus pallidus and 9 did not receive bilateral implants

Primary outcome:	difference between scores on the motor subscale of the UPDRS
Statistical analysis:	Wilcoxon rank-sum test was used to assess treatment, period, and carryover effects.

Effects fo multisite biventricular pacing in patients with heart failure and intraventricular conduction delay 309 (N Engl J Med)

N=67 patients with severe heart failure; 48 completed both study periods

Primary end point:	distance walked in 6 minutes
Statistical analysis:	The responses obtained for all criteria assessing clinical efficacy were compared with the use of the Wilcoxon test.

ESTUDI MULTINIVELL

Individual cholesterol variation in response to a margarine- or butter-based diet: a study in families *335 (JAMA)

Main outcome measures: mean LDL-C levels during the last 2 weeks of each dietary period

Statistical analysis:	generalized estimating equations (GEEs) to compare the 2 diets and construct Cis to adjust for the lack of independence within families

N=46 families; n(butter-margarine)= 23 [104 persons] and n(margarine-butter)=23 [102 persons]

"LONG RANK TEST"

A Comparison of Rate Control and Rhythm Control in Patients with Atrial Fibrilation [20] (N Engl J Med)	
Primary end point:	overall mortality
Statistical Analysis:	The base-line characteristics of patients were compared with chi-square tests and t-tests. The primary analysis was an intention-to-treat comparison of the time to death from any cause, adjusted to 10 interim analyses. For all time-to-event analyses, rates were estimated by the method of Kaplan and Meier and were compared by the long-rank test. Secondary analyses were conducted to evaluate results within prespecified subgroups and to adjust the primary end point for base-line characteristics. The prespecified covariates... were used to construct a multivariate Cox proportional-hazards survival model.

6.3.- Exemple Senn (ref. secció 2.1)

Senn and Auclair 1990 (p. 1290). There are measurements of peak expiratory flow (PEF), a measure of lung function (National Asthma Education Program, 1991, pp. 6-9), made on 13 children aged 7 to 14 with moderate or severe asthma in a two-treatment two-period cross-over comparing the effect of a single inhaled dose of 200 μg salbutamol, a well-established bronchodilator, and 12 μg formoterol, a more recently developed bronchodilator (Faulds et al. 1991). The children were randomized to one of two sequence groups. In one group they were given a dose of formoterol in the morning and observed for 8h in the clinic. They then traveled home where they or their parents took further measurements 10, 11 and 12h after treatment. On a subsequent occasion after a wash-out of at least one day they presented at the clinic again and were given a single dose of salbutamol. Measurements in the clinic followed as before and were again succeeded by measurements at home. For the second sequence group, the procedure was as for the first except that they received salbutamol on the first visit to the clinic and formoterol on the second visit to the clinic. [TREAT=1 → salbutamol ; TREAT=2 → formoterol]

EXAMPLE CROSSOVER: SENN,S: page 32

GROUP	PERIOD	PATNUM	TREAT	PEF	BASE
1	1	1	1	310	290
1	1	4	1	310	300
1	1	6	1	370	250
1	1	7	1	410	390
1	1	10	1	250	250
1	1	11	1	380	365
1	1	14	1	330	190
2	1	2	2	370	350
2	1	3	2	310	350
2	1	5	2	380	350
2	1	9	2	290	280
2	1	12	2	260	270
2	1	13	2	90	220
1	2	1	2	270	270
1	2	4	2	260	270
1	2	6	2	300	210
1	2	7	2	390	390
1	2	10	2	210	240
1	2	11	2	350	380
1	2	14	2	365	260
2	2	2	1	385	345
2	2	3	1	400	370
2	2	5	1	410	360
2	2	9	1	320	290
2	2	12	1	340	310
2	2	13	1	220	220

Grup	Pacient	Període 1		Període 2	
		for	sal	for	Sal
1: for/sal	1	310			270
	4	310			260
	6	370			300
	7	410			390
	10	250			210
	11	380			350
	14	330			365
2: sal/for	2		370	385	
	3		310	400	
	5		380	410	
	9		290	320	
	12		260	340	
	13		90	220	

A SIMPLE ANALYSIS IGNORING THE EFFECT OF PERIOD

A simple way of analysing data arranged in such a way is to perform a *matched pairs t-test*, sometimes aslo known as a *correlated t test*. (The first designation reflects the fact that the 26 PEF values may be matched in 13 pairs corresponding to the 13 patients. The second reflects the correlation that may be expected between values obtained for a given patient under one treatment and those obtained under another).

```
proc glm;
  class PATNUM TREAT;
  model PEF=PATNUM TREAT;
  estimate 'for-sal' TREAT 1 -1;
run;
```

The GLM Procedure

Class Level Information

Class	Levels	Values
PATNUM	13	1 2 3 4 5 6 7 9 10 11 12 13 14
TREAT	2	1 2

Number of observations 26

Dependent Variable: PEF

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	13	128601.9231	9892.4556	12.01	<.0001
Error	12	9886.5385	823.8782		
Corrected Total	25	138488.4615			

R-Square	Coeff Var	Root MSE	PEF Mean
0.928611	9.013107	28.70328	318.4615

Source	DF	Type I SS	Mean Square	F Value	Pr > F
PATNUM	12	115213.4615	9601.1218	11.65	<.0001
TREAT	1	13388.4615	13388.4615	16.25	0.0017

Source	DF	Type III SS	Mean Square	F Value	Pr > F
PATNUM	12	115213.4615	9601.1218	11.65	<.0001
TREAT	1	13388.4615	13388.4615	16.25	0.0017

Parameter	Estimate	Standard Error	t Value	Pr > |t|
for-sal	45.3846154	11.2583521	4.03	0.0017

ADJUSTING FOR A PERIOD EFFECT: TWO-SAMPLE t APPROACH

A very simple procedure allows us to adjust for period effects. Consider the period differences. If a contrast trend is present this must affect each of the period differences identically. Thus any differences between them cannot be due to the period effect. Differences between period differences in the same sequence group can be regarded as being random. On the other hand differences between any two period differences in different sequences would also reflect treatment differences. (They would also reflect carry-over if it were present, but for the purpose of the current discussion this will be assumed not to be a

problem.) Thus by comparing the means of the period differences for the two sequences we may examine the treatment effect. This may be done by using a two sample *t* test for the period differences.

```
proc glm;
 class PATNUM TREAT PERIOD;
 model PEF=PATNUM TREAT PERIOD;
 estimate 'for-sal' TREAT 1 -1;
run;
```

```
                        The GLM Procedure

                  Class Level Information
        Class         Levels    Values
        PATNUM            13    1 2 3 4 5 6 7 9 10 11 12 13 14
        TREAT              2    1 2
        PERIOD             2    1 2
                  Number of observations    26

Dependent Variable: PEF
                                    Sum of
Source                      DF       Squares    Mean Square   F Value   Pr > F
Model                       14    130233.9973     9302.4284     12.40   <.0001
Error                       11      8254.4643      750.4058
Corrected Total             25    138488.4615

            R-Square     Coeff Var      Root MSE      PEF Mean
            0.940396      8.601835      27.39354      318.4615

Source                      DF      Type I SS    Mean Square   F Value   Pr > F
PATNUM                      12    115213.4615     9601.1218     12.79   <.0001
TREAT                        1     13388.4615    13388.4615     17.84   0.0014
PERIOD                       1      1632.0742     1632.0742      2.17   0.1683

Source                      DF    Type III SS    Mean Square   F Value   Pr > F
PATNUM                      12    115213.4615     9601.1218     12.79   <.0001
TREAT                        1     14035.9203    14035.9203     18.70   0.0012
PERIOD                       1      1632.0742     1632.0742      2.17   0.1683

                                    Standard
Parameter                Estimate       Error    t Value   Pr > |t|
for-sal               46.6071429  10.7765596       4.32     0.0012
```

ANALYSIS FITTING PERIOD EFFECTS IN ADDITION TO TREATMENT:

The reader who is interested in testing for carry-over will require a different representation of the patient effect; he will need to split this into two sources: *GROUPS* and *PATIENTS* within *GROUP*. (It will be necessary to include a variable which identifies for each of the *2n* observations to which sequence *GROUP* the particular patient on whom that observation was made belongs).

```
proc glm;
 class GROUP PATNUM PERIOD TREAT;
 model PEF=GROUP PATNUM (GROUP) PERIOD TREAT;
 random PATNUM (GROUP)/test;
 estimate 'for-sal' TREAT 1 -1;
run;
```

The *model* statement has now been amended to split the variation between patients into two sources: that between groups, *GROUP*, and the variation between patients within groups, *PATIENT(GROUP)*. This split is made necessary because we wish to use the difference between sequence groups to say something about carry-over. In addition the *random* statement has been used to identify *PATIENTs* as random effect.

The GLM Procedure

Class Level Information

Class	Levels	Values
GROUP	2	1 2
PATNUM	13	1 2 3 4 5 6 7 9 10 11 12 13 14
PERIOD	2	1 2
TREAT	2	1 2

Number of observations 26

Dependent Variable: PEF

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	14	130233.9973	9302.4284	12.40	<.0001
Error	11	8254.4643	750.4058		
Corrected Total	25	138488.4615			

R-Square	Coeff Var	Root MSE	PEF Mean
0.940396	8.601835	27.39354	318.4615

Source	DF	Type I SS	Mean Square	F Value	Pr > F
GROUP	1	335.1877	335.1877	0.45	0.5177
PATNUM(GROUP)	11	114878.2738	10443.4794	13.92	<.0001
PERIOD	1	984.6154	984.6154	1.31	0.2763
TREAT	1	14035.9203	14035.9203	18.70	0.0012

Source	DF	Type III SS	Mean Square	F Value	Pr > F
GROUP	1	335.1877	335.1877	0.45	0.5177
PATNUM(GROUP)	11	114878.2738	10443.4794	13.92	<.0001
PERIOD	1	1632.0742	1632.0742	2.17	0.1683
TREAT	1	14035.9203	14035.9203	18.70	0.0012

Source	Type III Expected Mean Square
GROUP	Var(Error) + 2 Var(PATNUM(GROUP)) + Q(GROUP)
PATNUM(GROUP)	Var(Error) + 2 Var(PATNUM(GROUP))
PERIOD	Var(Error) + Q(PERIOD)
TREAT	Var(Error) + Q(TREAT)

Tests of Hypotheses for Mixed Model Analysis of Variance

Dependent Variable: PEF

Source	DF	Type III SS	Mean Square	F Value	Pr > F
GROUP	1	335.187729	335.187729	0.03	0.8611
Error	11	114878	10443		
Error: MS(PATNUM(GROUP))					

Source	DF	Type III SS	Mean Square	F Value	Pr > F
PATNUM(GROUP)	11	114878	10443	13.92	<.0001
PERIOD	1	1632.074176	1632.074176	2.17	0.1683
TREAT	1	14036	14036	18.70	0.0012
Error: MS(Error)	11	8254.464286	750.405844		

Parameter	Estimate	Standard Error	t Value	Pr > |t|
for-sal	46.6071429	10.7765596	4.32	0.0012

L'estimació de la variància within que s'obté és: σ_W^2=754,87

ANALYSIS USING A SINGLE BASELINE:

We use a single baseline before treatment and no longer regard this as a concomitant value associated with the first outcome measures but as an extra 'outcome' in its own right. Thus where we have 13 patients we now have 3*13=39 values of the variable *OUTCOME* as opposed to the 26 we had previously. Corresponding to each outcome value we record the *PATIENT* it was measured for and the *PERIOD* in which it was measured as before, only now *PERIOD* is a variable with 3 values 0 (for the baseline period) and 1 and 2 (for the treatment periods). We now have as our *TREAT*ment variable, -1 (for salbutamol) and 0 (for baseline). The effect of this is that the baseline itself contributes nothing to the treatment estimate.

```
proc glm;
  class PATIENT PERIOD;
  model OUTCOME = PATIENT PERIOD TREAT;
  estimate 'treatment effect' TREAT 2;
run;
```

Note the change in the form of the *estimate* statement. Because *TREAT* is a numerical variable *estimate* provides a slope. To move from –1 for salbutamol to +1 for formoterol is a change of 2 units. This is why the '2' is included in the *estimate* statement. This produces an estimate of the treatment effect, 46.61, which is identical to that obtained by our calculations. The same value would result using the 26 genuine *OUTCOMEs* only and fitting *PATIENT PERIOD* and *TREAT* as class variables. What changes is the standard error This is now 14.59, based on 23 degrees of freedom, as opposed to 10.78, based on 11 degrees of freedom. The extra 12 degrees of freedom have arisen because we have used 13 further *OUTCOMEs* as the cost of only one further parameter being fitted for *PERIOD*.

It is interesting to trace the origin of the increase in the standard error. The SAS® output for these two approaches shows that the estimated error variance for 11 degrees of freedom is 750.41, that for 23 degrees of freedom is 1375.25. The ratio of the square of the standard errors to the error variance will be found to be the same in each case: $10.78^2/750.41=14.59^2/1375.25=0.155$. This is as it should be. Both methods will always produce identical estimates, hence the variance of the two methods must be the same. This variance is, in fact, $\sigma_W^2(1/6+1/7)/2=0.155\sigma_W^2$ where σ_W^2 is the within-patient variance and is half the variance of a basic estimator (cross-over differences). What has changed from one case to the other is not the estimate itself, nor the variance of the estimate, but the estimate of the variance of the estimate. We have used the baseline values in addition for this purpose. Their effect has been to increase our estimate of σ_W^2 considerably. In fact if we isolate these 12 degrees of freedom we find that the estimated error variance associated with them alone is 1948.02 compared to the 750.41 the other 11.

There are many possible explanations for this difference, not least among them chance. One possible explanation is that patients respond to different degrees to beta-agonists as a class of drug. Because each patient acts as his own control, in comparing two beta-agonists this source of patient heterogeneity is eliminate. Comparing to baseline it is not eliminated.

Of the two approaches, I prefer the one which does not use the baseline and estimates the error variance using 11 degrees of freedom rather than 23. This is not because it leads to a lower estimated standard error. I should still prefer it if the reverse were the case. I prefer it because it is based on the direct comparison of the two treatments. Further discussion of this general issue is given in Chapter 5.

6.4.- Confusió d'efectes en el disseny cross-over AB/BA

Els factors del disseny AB/BA són els següents:

- Tractament: α_i ; $_{i=A,B}$
- Període o temps: π_j ; $_{j=1,2}$
- Ordre o seqüència: O_k ; $_{k=1(AB),\,2(BA)}$

Per tant, obtindriem 8 possibles combinacions però només 4 d'elles observables (certes combinacions són incompatibles, es corresponen amb les cel·les pintades de la taula de la pàgina 7) i, conseqüentment, el model amb totes les interaccions ($\alpha\pi$, αO, πO, $\alpha\pi O$) estaria sobredimensionat. L'algorisme dels signes d'un disseny factorial 2x2x2 que defineix els contrastos per a cada efecte, essent m_{ijk} la mitjana en el tractament i, període j i seqüència k seria:

	μ	α	π	O	$\alpha\pi$	αO	πO	$\alpha\pi O$
m_{A11}	+	+	+	+	+	+	+	+
m_{B11}	+	−	+	+	−	−	+	−
m_{A21}	+	+	−	+	−	+	−	−
m_{B21}	+	−	−	+	+	−	−	+
m_{A12}	+	+	+	−	+	−	−	−
m_{B12}	+	−	+	−	−	+	−	+
m_{A22}	+	+	−	−	−	−	+	+
m_{B22}	+	−	−	−	+	+	+	−

Però com les mitjanes m_{B11}, m_{A21}, m_{A12}, m_{B22} no s'observen:

	μ	α	π	O	$\alpha\pi$	αO	πO	$\alpha\pi O$
m_{A11}	+	+	+	+	+	+	+	+
m_{B21}	+	−	−	+	+	−	−	+
m_{B12}	+	−	+	−	−	+	−	+
m_{A22}	+	+	−	−	−	−	+	+

el resultat és que es troben confosos els efectes: α amb πO, π amb αO, O amb $\alpha\pi$ i $\alpha\pi O$ amb μ; i, per tant, per a poder fer inferència sobre els efectes d'interès serà necessari assumir que les interaccions $\alpha\pi$, αO, πO i $\alpha\pi O$ són zero.

6.5.- "*Independent t-test based on matched-pairs differences*"

Recordem que la variable "*cross-over differences*" es defineix com:

$$c_{1k} = y_{11k} - y_{12k} \qquad \text{si } k \in AB\,(1)$$
$$c_{2k} = y_{22k} - y_{21k} \qquad \text{si } k \in BA\,(2)$$

D'acord amb la premissa formulada sobre la distribució d'aquesta variable, el model que ajusta les dades és $Y_{ijk}=\mu+\tau_{ik}+\alpha_{d[i,j]}+\varepsilon_{ijk}$ i, per tant, les esperances $E[c_{1k}]=E[c_{2k}]=\alpha_1-\alpha_2$. Aleshores, per contrastar H_0: $\alpha_1=\alpha_2$, s'aplicarà el conegut t-test per a dades aparellades a la variable "*cross-over differences*".

Si definim $\alpha_c=\alpha_1-\alpha_2$ i $\hat{\alpha}_c=\bar{c}_{..}$ i notem per $n=n_1+n_2$:

$$E[\hat{\alpha}_c]=\alpha_c$$
$$V[\hat{\alpha}_c]=\frac{2\sigma_e^2}{n}=\frac{\sigma_c^2}{n}$$

Per estimar σ_c^2 s'usa:

$$S_c^2=\frac{\sum_{i=1}^{2}\sum_{k=1}^{n_i}(c_{ik}-\bar{c}_{..})^2}{n-1}$$

amb n−1 graus de llibertat. Sota H_0, l'estadístic per a contrastar l'igualtat d'efectes:

$$t_c=\sqrt{n}\cdot\bar{c}_{..}/S_c$$

es distribueix segons una *t-Student* amb n-1 graus de llibertat. L'interval de confiança al 95% per al veritable efecte del tractament, α_d, serà:

$$\alpha_c \in \bar{c}_{..} \pm t_{n-1}^{\alpha/2}\cdot S_c/\sqrt{n}$$

A partir de l'interval de confiança, es tracta d'aillar *S* i d'obtenir una sola estimació promitjant els dos valors obtinguts a partir dels límits superior (Upper Confidence Interval, UCI) i l'inferior (Lower Confidence Interval, LCI) de l'interval de confiança:

$$S_{UCI}=\frac{(UCI-\bar{c}_{..})\sqrt{n}}{t_{n-1}^{0.05/2}} \qquad S_{LCI}=\frac{(\bar{c}_{..}-LCI)\sqrt{n}}{t_{n-1}^{0.05/2}}$$

6.6.- Llistat mails dels articles on falta informació

Codi Medline	Revista	Títol
377	AMJMED	**The effects of Transdermal Estradiol on the response to mental stress in postmenopausal women: a randomized trial**
		petició / **adreça** 26/02/03 geriat@ipruniv.cce.unipr.it 18/02/03 gdelrio@regione.emilia-romagna.it 12/02/03 graziano.ceresini@unipr.it
		respostes
268	ANNALS	**Mediterranean and Low-Fat diets improve endothelial function in hypercholesterolemic men**
		petició / **adreça** 11/03/03 fjfuentes@hrs.sas.junta-andalucia.es 20/02/03 fjfuentes@hrs.sas.junta-andalucia.es 18/02/03 renee.g.tagert@uth.tmc.edu 12/02/03 fperez@sofia.hrs.sas.cica.es
		respostes 12/03/03
		En los próximos días recibirán los datos sin problemas y perdonen por la tardanza, un saludo
532	BMJ	**Effect of temazepan on ventilatory response at moderate altitude**
		petició / **adreça** 18/02/03 interne@khneunkirchen.at
		respostes
384	JAMA	**Efficacy of 3 commonly used hearing aid circuits: a cross-over trial**
		petició / **adreça** 01/04/03 vlarson@howardleight.com
		respostes 01/04/03
		Thank you for your inquiry. I'm forwarding your request to David Williams, the biostatistician on the study. His e-mail is: williams@research.hines.med.va.gov
370	LANCET	**Topical butyrate for acute radiation proctitis: randomised, cross-over trial**
		petició / **adreça** 12/02/03 disciclin@uniroma1.it
		Mail al departament de ciències clíniques, per demanar informació sobre el Dr. Piero Vernia
		respostes
413	LANCET	**Comparison of pharmacokinetics and systemic effects of inhaled fluticasono propionate in patients with asthma and healthy volunteers: a randomised crossover study**
		petició / **adreça** 20/04/03 ashley@nwlung.u-net.com
		Mail al departament de ciències clíniques, per demanar informació sobre el Dr. Piero Vernia
		respostes

480	LANCET	**Effect of consumption of red wine, spirits, and beer on serum homocysteine**	
		petició	**Adreça**
		01/04/03	blseni@uta.fi
		18/02/03	Hendriks@voeding.tno.nl
		respostes	
493	LANCET	**Transcranial magnetic sitmulation and auditory hallucinations in schizophrenia**	
		petició	**Adreça**
		23/04/03	Ralph.Hoffman@yale.edu
		11/04/03	ralph.hoffman@yale.edu
		respostes	
		11/04/03 Is this for rTMS? Effect sizes are going to vary from study to study. I don't really understand what you are trying to show. Can you give me more details? RH 23/04/03 THis is going to take a while to track down since the study was finished quite a long time ago	
414	NEJM	**Transdermal testosterone treatment in women with impaired sexual function after oophorectomy**	
		petició	**adreça**
		01/04/03	jshifren@partners.org
		18/02/03	peter.casson@vtmednet.org
		13/02/03	janshifren@hotmail.com
		respostes	
		01/04/03 I will be out of the office until Monday, March 31, 2003.	jshifren@partners.org

6.7.- Article originals que compleixen criteris de selecció

(54) Mechanism of Cocaine Induced Hyperthermia in Humans

[30]

Craig G. Crandall, PhD; Wanpen Vongpatanasin, MD; and Ronald G. Victor, MD

Background: The lethal effects of cocaine are unique among those of other illicit drugs because cocaine has the propensity to cause hyperthermia. The traditional view is that cocaine causes a hypermetabolic state with increased heat production. However, because cocaine induced hyperthermia occurs primarily in hot weather, it is hypothesized that cocaine also impairs thermoregulatory adjustments that mediate heat dissipation.

Objective: To test the effects of cocaine on body temperature regulation in humans.

Design: Randomized, double blind, placebo controlled crossover trial.

Setting: A cardiovascular physiology laboratory in Dallas, Texas.

Participants: 7 healthy, cocaine naive volunteers.

Intervention: Progressive passive heat stress, during which each participant received intranasal cocaine (2 mg/kg of body weight) or placebo (lidocaine, 2 mg/kg).

Measurements: Esophageal temperature, skin blood flow, sweat rate, and perceived thermal sensation.

Results: Three major new findings were noted. First, cocaine substantially augmented the progressive increase in esophageal temperature during heat stress ($P < 0.001$). Second, this augmentation was explained by a rightward shift in the esophageal temperature threshold for the onset of both cutaneous vasodilation (37.37 ± 0.09 °C for cocaine vs. 37.06 ± 0.07 °C for lidocaine; $P = 0.01$) and sweating (37.38 ± 0.09 °C for cocaine vs. 37.07 ± 0.06 °C for lidocaine; $P = 0.002$). Third, cocaine paradoxically impaired the perception of heating by attenuating the progressive increase in thermal discomfort associated with heat stress.

Conclusions: In humans, impaired heat dissipation is a major mechanism by which cocaine elevates body temperature. When healthy, cocaine naive persons are subjected to passive heating, pretreatment with even a small dose of intranasal cocaine impairs sweating and cutaneous vasodilation (the major autonomic adjustments to thermal stress) and heat perception (the key trigger for behavioral adjustments).

Ann Intern Med. 2002;136:785-791. www.annals.org
For author affiliations, see end of text.

Cocaine abuse is a major cause of life-threatening cardiovascular emergencies, including hypertensive crisis, acute myocardial infarction, and ventricular arrhythmias. The lethal effect of cocaine is unique among those of other illicit drugs because it is related not only to dose but also to cocaine's propensity to cause hyperthermia. Although fatal cocaine overdose typically is associated with high blood cocaine levels (3 to 6 mg/L) (1), cocaine-related deaths can also occur when hyperthermia is present at blood levels 10 to 20 times lower (2). The intrinsic thermogenic property of cocaine underlies recent epidemiologic data indicating that mortality rates for cocaine overdose increase substantially in hot weather (3). This temperature-dependent increase in mortality rates is specific to cocaine and is not seen with opiates or other illicit drugs. These clinical observations are bolstered by experiments in dogs demonstrating that cocaine-induced death can be eliminated by multiple strategies that prevent hyperthermia (4). In humans, however, the underlying mechanisms mediating cocaine-induced hyperthermia are poorly understood.

The hyperthermic properties of cocaine have been attributed largely to a hypermetabolic state (agitation with increased locomotor activity) that increases heat production (5). Indeed, cocaine-induced hyperthermia has been likened to two other syndromes, neuroleptic malignant syndrome and malignant hyperthermia, that are characterized by hyperpyrexia, delirium, tachycardia, and elevated blood pressure (6, 7). In both of those syndromes, excessive heat production due to involuntary muscle contraction is sufficient to raise body temperature even in a cool environment (for example, an operating room). In contrast, cocaine-induced hyperthermia is seen primarily in the setting of high ambient temperatures. We therefore hypothesized that, in addition to increased heat production, impaired heat dissipation is another important mechanism contributing to the hyperthermic properties of cocaine.

When body temperature increases above thermoregulatory set points, heat dissipation depends on both autonomic and behavioral adjustments. The major autonomic adjustments include activation of sympathetic cholinergic fibers that leads to sweating and cutaneous vasodilation (8). The behavioral adjustments are

Context

The cause of hyperthermia in cocaine abuse is not well understood. Both excess heat generation and defective heat dissipation are potential causes of fatal hyperthermia.

Contribution

Nasal administration of cocaine causes greater increase in core body temperature, decrease in heat perception, and greater impairment of sweating and skin blood flow compared with nasal lidocaine administration.

Implications

Excessive heat production, impaired heat dissipation, and alteration of behavioral responses to increased body temperature may lead to fatal hyperthermia in cocaine abusers.

–The Editors

prompted by thermal discomfort and include heat avoidance and external cooling. In humans, these simple behavioral responses can be more important than autonomic adjustments in maintaining normal body temperature during heat stress (9).

We sought to test the effect of cocaine on thermoregulatory adjustments in humans. Because cocaine has a profound influence on both cardiovascular and cognitive function, we conducted a randomized, placebo controlled study to assess the effects of a low, nonеuphoric dose of intranasal cocaine on autonomic and behavioral thermoregulation in healthy cocaine naive volunteers.

Methods

Participants

We studied 7 healthy male and female volunteers who ranged in age from 23 to 37 years. The Institutional Review Board of the University of Texas Southwestern Medical Center at Dallas, Texas, approved the protocol, and all volunteers provided written informed consent. All participants were normotensive and had no history of cardiovascular disease, cocaine abuse, or other recreational drug use. No participant was taking prescription or nonprescription drugs with cardiovascular or autonomic effects. Participants refrained from smoking cigarettes or drinking alcohol or caffeine containing beverages for at least 12 hours before the experiment.

Measurements

Internal temperature was measured by using a thermistor placed in the esophagus at a depth equal to 25% of each participant's standing height, and skin temperature was measured from the weighted electrical average of six thermocouples attached to the skin. Each participant wore a tube lined suit that controlled skin temperature by changing the temperature of the water perfusing the suit. The suit covered the entire body surface with the exception of the head, arms, and feet. We assessed skin blood flow from each participant's forearm during heat stress. The water perfused suit did not cover the forearm where skin blood flow measurements were obtained; thus, any change in skin blood flow during heat stress was not directly related to mechanisms associated with local heating. Skin blood flow was measured by using multifiber laser Doppler flowmeter (Perimed, North Royalton, Ohio). Heart rate was obtained from the electrocardiogram by using a cardiotachometer (CWE, Inc., Ardmore, Pennsylvania), and blood pressure was obtained by using the oscillometric technique (Welch Allyn, Beaverton, Oregon). Mean arterial blood pressure was calculated as one third of the pulse pressure plus diastolic blood pressure. Cutaneous vascular conductance, which is the reciprocal of vascular resistance, was calculated from the ratio of the laser Doppler signal to mean arterial pressure.

Forearm sweat rate was measured continuously with the ventilated capsule method, in which a capsule with a window of 2.8 cm^2 is attached to the surface of the skin. Nitrogen gas is perfused through the capsule at a fixed rate, and as the person begins to sweat, the water on the skin evaporates through the window into the nitrogen gas. The humidity of the effluent nitrogen gas is quantified through a humidity detector "downstream" from the capsule. The sweat rate is then calculated from the humidity, temperature, and flow of the nitrogen gas. Throughout the heat stress, participants rated thermal sensation by using a standardized 9 point thermal comfort scale ranging from 4.0 (neutral) to 8.0 (unbearably hot) at increments of 0.5 (10).

Experimental Protocol

All experiments were conducted with the participants in a supine position. In a randomized, double blind, crossover trial, each participant was exposed to

whole body heating after receiving intranasal cocaine hydrochloride (2 mg/kg of body weight, 10% solution) or lidocaine hydrochloride (2 mg/kg, 10% solution). Lidocaine was used as an internal control for the local anesthetic property of cocaine (11). The experiments were conducted on 2 separate days, at the same time of day, 3 to 6 days apart. The dose of cocaine administered is half the standard dose administered for rhinolaryngological procedures (12). The heat stress test was performed by perfusing the tube lined suit with water at 46 °C. To assess maximal vasodilaror responses, a heater surrounding the laser Doppler flow probe was used on completion of the heat stress to increase local skin temperature to 42 °C for 30 minutes. Heating the skin at this temperature and duration elicits maximal cutaneous vasodilarion (13). Skin blood flow during the final minutes of this 30 minute procedure was averaged, and maximal cutaneous vascular conductance was then calculated from the ratio of that value and mean arterial blood pressure. Values for cutaneous vascular conductance were then normalized to percentages of maximum for that site (13).

Data Collection

Temperatures, skin blood flow, and sweat rate data were sampled at 50 Hz (Biopac, Santa Barbara, California). The electrocardiogram was sampled by a cardiotachometer at 1000 Hz, using a multichannel digital data recorder (CWE, Inc.). These values were reduced to 20 second averages before statistical analysis. When cutaneous vascular conductance and sweat rate are plotted relative to esophageal temperature, two important variables can be identified to assess thermoregulatory function (14): the esophageal temperature threshold at which cutaneous vasodilation and sweating begin, and the slope of the line relating progressive elevations in cutaneous vascular conductance and sweating to increasing esophageal temperature. Esophageal temperature thresholds at which cutaneous vascular conductance and sweat rate began to increase were identified from individual plots of esophageal temperature versus cutaneous vascular conductance and esophageal temperature versus sweat rate. These points were selected as the internal temperatures at which pronounced and sustained increases in cutaneous vascular conductance and sweating were evident during the heating procedure. The investigator identifying these thresholds was blinded to the participants and to their experimental conditions (lidocaine or cocaine).

Statistical Analysis

Statistical comparisons of temperature thresholds and slopes during cocaine and lidocaine administration were performed by using the paired *t* test. Differences in responses (internal temperature, skin temperature, cutaneous vascular conductance, and sweat rate) during cocaine and lidocaine administration throughout the heat stress were statistically analyzed by using two way repeated measures analysis of variance (SigmaStat 2.0, SPSS Science, Chicago, Illinois). Treatment order was also assessed in the models, and no effect of treatment order on any outcome variables was found. The effects of cocaine on thermal sensation were statistically analyzed by comparing the internal temperature at each perceived heating score between 5.0 and 7.5 using paired *t* tests. All values are reported as the mean (±SE), and the α level for all statistical analyses was set at 0.05.

Role of the Funding Sources

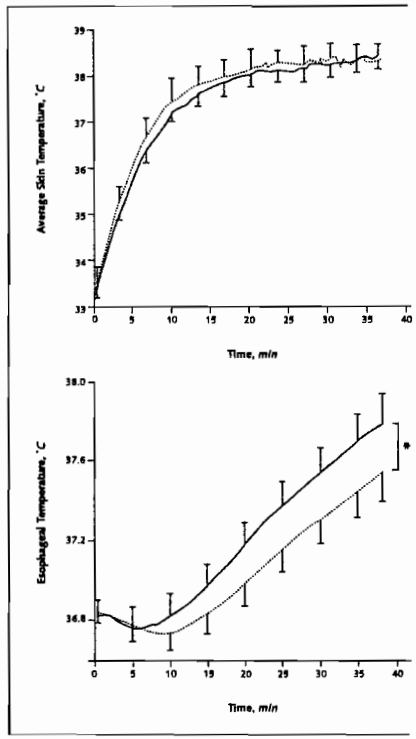
The funding sources had no role in the design, analysis, or interpretation of the data or in the decision to submit the manuscript for publication.

Results

Before heat stress, cocaine had no effect on esophageal temperature compared with lidocaine (36.79 ± 0.01 °C vs. 36.75 ± 0.08 °C). However, cocaine significantly increased mean arterial pressure (93 ± 3 mm Hg vs. 85 ± 3 mm Hg; $P = 0.001$) and heart rate (74 ± 5 beats/min vs. 67 ± 4 beats/min; $P = 0.01$). Cocaine had no significant effect on cutaneous vascular conductance before heat stress (9 ± 2% of maximum vs. 7 ± 2% of maximum; $P = 0.09$).

Heat stress had no effect on mean arterial pressure during cocaine or lidocaine administration (93 ± 3 to 92 ± 2 mm Hg and 85 ± 3 to 87 ± 2 mm Hg, respectively) but significantly increased heart rate for both conditions (74 ± 5 to 107 ± 5 beats/min and 67 ± 4 to 105 ± 5 beats/min, respectively). However, mean arterial pressure at the end of heat stress was still significantly higher during cocaine administration than during lidocaine administration (92 ± 2 mm Hg vs. 87 ± 2 mm Hg, respectively; $P < 0.01$), whereas heart rate did not differ significantly between agents (107 ± 5 beats/min vs. 105 ± 5 beats/min; $P > 0.2$). Passive heating yielded comparable increases in skin temperature during both cocaine and lidocaine administration, confirming that the external heat stress applied was identical on both days (Figure 1, *top*).

Figure 1. **Responses of average skin temperature and esophageal temperature to cocaine (*solid lines*) and lidocaine (*dotted lines*) during whole body heating.**



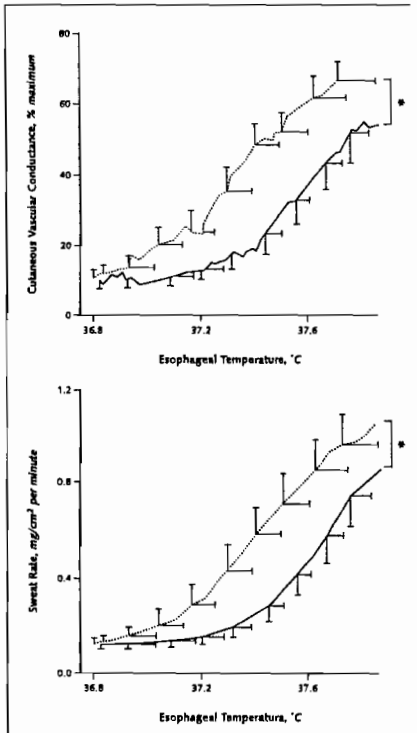
Data are the mean (±SE). *$P < 0.001$ for cocaine vs. lidocaine.

These results represent three major new findings. First, compared with lidocaine, cocaine significantly augmented the progressive increase in esophageal temperature during heat stress ($P < 0.001$) (Figure 1, *bottom*). Second, cocaine also significantly attenuated the progressive increase in cutaneous vascular conductance and sweating during heat stress (Figure 2). At the highest common esophageal temperature of 37.86 °C, cutaneous vascular conductance (49 ± 9% of maximum vs. 70 ± 5% of maximum; $P = 0.03$) and sweat rate (0.8 ± 0.2 mg/cm^2 per minute vs. 1.0 ± 0.1 mg/cm^2 per minute; $P = 0.04$) were significantly lower during cocaine administration than during lidocaine administration (Figure 2). The cocaine induced decrease in cutaneous vasodilation and sweating during heat stress was related to a delay in the onset of these thermoregulatory responses. This is evident by significant increases in the esophageal temperature thresholds for cutaneous vasodilation (37.37 ± 0.09 °C for cocaine vs. 37.06 ± 0.07 °C for lidocaine; $P = 0.01$) and sweating (37.38 ± 0.09 °C for cocaine vs. 37.07 ± 0.06 °C for lidocaine; $P = 0.002$), with no change in the slopes of these relationships. Third, although cocaine augmented the progressive increase in esophageal temperature during heat stress, it paradoxically attenuated the increase in thermal discomfort in the same participants (Figure 3). The difference in heat perception between cocaine and lidocaine administration first became statistically significant when esophageal temperature increased above 37.0 °C and became larger as esophageal temperature progressively increased ($P < 0.05$) (Figure 3).

Discussion

Although hyperthermia increases the risk for death after cocaine administration, the underlying mechanisms mediating cocaine-induced hyperthermia are poorly understood. The thermogenic properties of cocaine have been attributed largely to a hypermetabolic state that increases heat production. In contrast to this traditional view, our study shows that impaired heat dissipation is another major mechanism by which cocaine elevates body temperature in humans. When healthy cocaine-naive persons were subjected to passive heating, pretreatment with even a small dose of intranasal cocaine impaired sweating and cutaneous vasodilation (the major autonomic adjustments to thermal stress) as well as heat perception (the key trigger for behavioral adjustments).

Figure 2. **Effects of cocaine on autonomic adjustments to heat stress.**

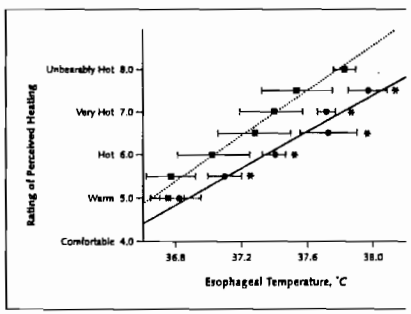


Solid lines indicate cocaine; dotted lines indicate lidocaine. Data are the mean (±SE). $^*P < 0.01$ for cocaine vs. lidocaine.

When a person experiences repeated bouts of heat stress, the threshold temperature for the onset of sweating is remarkably constant, with an average test–retest variability of 0.09 °C (15). During the passive heating used in our experiments, even a small dose of intranasal cocaine elevated the threshold for both sweating and cutaneous vasodilation by more than three times this value, thereby markedly augmenting the increase in core temperature. That skin temperature increased similarly during both cocaine and lidocaine administration indicates that the thermal stress was equivalent under both conditions. In the absence of external heat stress, cocaine did not affect core temperature, suggesting that impaired heat dissipation (rather than augmented heat production) was the dominant mechanism underlying cocaine's hyperthermic properties. This pattern of cocaine-induced hyperthermia after exposure to heat stress resembles hyperthermia caused by impaired sweat gland function in patients with hereditary ectodermal dysplasia (16, 17), poisoning by muscarinic receptor antagonist (18), or growth hormone deficiency (19). However, the slope of the line relating increased sweating to increased core temperature was markedly reduced in these conditions (16, 19) but was unaffected by cocaine in our study, suggesting differences in pathogenetic mechanisms.

The precise mechanism by which cocaine impairs cutaneous vasodilation and sweating is still unknown. Both peripheral and central neural mechanisms could be involved. For example, cocaine is thought to inhibit the peripheral norepinephrine transporter, thereby increasing the norepinephrine concentration in the synaptic cleft (20). During thermal stress, augmented α adrenergic vasoconstrictor tone in the cutaneous bed could contribute to the delayed onset of vasodilation. Inhibition of

Figure 3. **Effects of cocaine on thermal perception.**



Solid lines indicate cocaine; dotted lines indicate lidocaine. Data are the mean (±SE). The ratings of perceived heating were modified from reference 10. $^*P < 0.05$ for cocaine vs. lidocaine.

norepinephrine reuptake, however, cannot explain the delayed onset of sweating during cocaine; sweating is regulated not by noradrenergic fibers but rather by sympathetic cholinergic fibers (9, 21).

The most plausible explanation for this mechanism is that cocaine-induced shifts in the threshold for cutaneous vasodilation and sweating are mediated centrally. This notion is consistent with recent findings indicating that cocaine acts centrally to stimulate sympathetic vasoconstrictor drive (11). We previously showed that under normothermic conditions, the dose of intranasal cocaine used in the current study increases sympathetic vasoconstrictor drive targeted to both the skeletal muscle and cutaneous beds (11). Our current data suggest that cocaine produces a distinctive pattern of altered central sympathetic outflow: augmentation of vasoconstrictor pathways accompanied by attenuation of active vasodilator and sudomotor pathways. This unique property of cocaine is not shared by other sympathomimetic drugs, such as methamphetamine and ephedrine, which cause hyperthermia mainly by increasing skeletal muscle metabolism and thermogenesis (22, 23).

In our current study, cocaine impaired the perception of heat stress; this is the most clear-cut evidence that cocaine acts centrally to alter thermoregulatory responses. This impairment was dramatic: Participants experienced less thermal discomfort with cocaine even though core temperature was higher than with lidocaine. We should emphasize, however, that this effect was observed with a very low dose of cocaine that did not produce euphoria (11). Larger doses of cocaine that produce intoxication, agitation, and increased locomotor activity would be expected to increase heat production, thereby compounding the hyperthermic effects of impaired heat dissipation. Furthermore, impaired heat perception would impair behavioral responses to hyperthermia, such as seeking a cooler environment or adjusting the thermostat on the air conditioner. In humans, these behavioral responses have the most powerful thermoregulatory effects (9).

We speculate that when recreational doses of cocaine are taken in a warm environment, such as in hot weather, in crowded nightclubs, or at "rave parties," the hyperthermic effects of the drug will be greatly amplified (6, 24, 25). Because our study was performed only in cocaine-naive, healthy persons, the results may not be applicable to persons with long-term cocaine use. However, long-term exposure to cocaine does not produce tolerance to cocaine-induced hyperthermia in rodents and can even produce "reverse tolerance," an augmented response (26–28). Although our findings provide a new explanation for cocaine-induced hyperthermia at the organ systems level, the cellular and molecular mechanisms are unknown. Elucidating the precise molecular mechanisms of cocaine-induced hyperthermia could lead to identification of new drug targets that may reduce cocaine-related deaths.

From University of Texas Southwestern Medical Center and Institute for Exercise and Environmental Medicine, Presbyterian Hospital of Dallas, Dallas, Texas.

Grant Support: By the National Institutes of Health (HL 61388) (Dr. Crandall); the American Heart Association, Texas Affiliate (0060010Y) (Dr. Vongpatanasin); and the National Institute on Drug Abuse (RO 1 DA10064) (Dr. Victor).

Requests for Single Reprints: Craig G. Crandall, PhD, Institute for Exercise and Environmental Medicine, Presbyterian Hospital of Dallas, 7232 Greenville Avenue, Dallas, TX 75231; e-mail, CraigCrandall@texashealth.org.

Potential Financial Conflicts of Interest: None disclosed.

Current author addresses and author contributions are available at www.annals.org.

References

1. Escobedo LG, Ruttenber AJ, Agnew MM, Anda RF, Wetli CV. Emerging patterns of cocaine use and the epidemic of cocaine overdose deaths in Dade County, Florida. Arch Pathol Lab Med. 1991;115:900 5. [PMID: 1929786]

2. Konen TR, Kleber HD. Rapid death during cocaine abuse: a variant of the neuroleptic malignant syndrome? Am J Drug Alcohol Abuse. 1988;14:335 46. [PMID: 3055940]

3. Marzuk PM, Tardiff K, Leon AC, Hirsch CS, Portera L, Iqbal MI, et al. Ambient temperature and mortality from unintentional cocaine overdose. JAMA. 1998;279:1795 800. [PMID: 9628710]

4. Catravas JD, Waters IW. Acute cocaine intoxication in the conscious dog: studies on the mechanism of lethality. J Pharmacol Exp Ther. 1981;217:350 6. [PMID: 7229977]

5. Lenhardt R, Kurz A, Sessler DI. Thermoregulation and hyperthermia. Acta Anaesthesiol Scand. 1996;109(Suppl):34 8.

6. Daras M, Kaldoudras L, Tuchman AJ, Koppel BS. Rhabdomyolysis and hyperthermia after cocaine abuse: a variant of the neuroleptic malignant syndrome? Acta Neurol Scand. 1995;92:161 5. [PMID: 7484066]

7. Sessler DI. Malignant hyperthermia. Acta Anaesthesiol Scand. 1996; 109(Suppl):25 30.

8. Kellogg DL Jr, Pérgola PE, Piest KL, Kosiba WA, Crandall CG, Grossman M, et al. Cutaneous active vasodilation in humans is mediated by cholinergic nerve cotransmission. Circ Res. 1995;77:1222 8. [PMID: 7586235]

9. Sessler DI. Mild perioperative hypothermia. N Engl J Med. 1997;336:1730 7. [PMID: 9180091]

10. Toner MM, Drolet LL, Pandolf KB. Perceptual and physiological responses during exercise in cool and cold water. Perceptual and Motor Skills. 1986;62: 211 20.

11. Vongpatanasin W, Mansour Y, Chavoshan B, Arbique D, Victor RG. Cocaine stimulates the human cardiovascular system via a central mechanism of action. Circulation. 1999;100:497 502. [PMID: 10430763]

12. Johns ME, Henderson RL. Cocaine use by the otolaryngologist: a survey. Transactions of the American Academy of Ophthalmology and Otolaryngology. 1977;84:969 73.

13. Taylor WF, Johnson JM, O'Leary D, Park MK. Effect of high local temperature on reflex cutaneous vasodilation. J Appl Physiol. 1984;57:191 6. [PMID: 6469780]

14. Gisolfi CV, Wenger CB. Temperature regulation during exercise: old concepts, new ideas. Exerc Sport Sci Rev. 1984;12:339 72. [PMID: 6376137]

15. Brengelmann GL, Savage MV, Avery DH. Reproducibility of core temperature threshold for sweating onset in humans. J Appl Physiol. 1994;77:1671 7. [PMID: 7836185]

16. Brengelmann GL, Freund PR, Rowell LB, Olerud JE, Kraning KK. Absence of active cutaneous vasodilation associated with congenital absence of sweat glands in humans. Am J Physiol. 1981;240:H571 5. [PMID: 7223909]

17. Totel GL. Physiological responses to heat of resting man with impaired sweating capacity. J Appl Physiol. 1974;37:346 52. [PMID: 4415071]

18. Brown JH, Taylor P. Muscarinic receptor agonists and antagonists. In: Hardman JG, Limbird LL, Gilman AG, eds. Goodman and Gilman's The Pharmacological Basis of Therapeutics. 10th ed. New York: McGraw Hill; 2001: 162 8.

19. Juul A, Hjortskov N, Jepsen LT, Nielsen B, Hallkjaer Kristensen J, Vahl N, et al. Growth hormone deficiency and hyperthermia during exercise: a controlled study of sixteen GH deficient patients. J Clin Endocrinol Metab. 1995;80:3335 40. [PMID: 7593447]

20. Lange RA, Hillis LD. Cardiovascular complications of cocaine use. N Engl J Med. 2001;345:351 8. [PMID: 11484693]

21. Rowell LB. Control of regional blood flow during dynamic exercise. In: Rowell LB, ed. Human Cardiovascular Control. New York: Oxford Univ Pr; 1993:220 8.

22. Makisumi T, Yoshida K, Watanabe T, Tan N, Murakami N, Morimoto A. Sympatho adrenal involvement in methamphetamine induced hyperthermia through skeletal muscle hypermetabolism. Eur J Pharmacol. 1998;363:107 12. [PMID: 9881575]

23. Astrup A, Bulow J, Madsen J, Christensen NJ. Contribution of BAT and skeletal muscle to thermogenesis induced by ephedrine in man. Am J Physiol. 1985;248(5 Pt 1):E507 15. [PMID: 3922230]

24. Bauwens JE, Boggs JM, Hartwell PS. Fatal hyperthermia associated with cocaine use. West J Med. 1989;150:210 2. [PMID: 2728444]

25. Callaway CW, Clark RF. Hyperthermia in psychostimulant overdose. Ann Emerg Med. 1994;24:68 76. [PMID: 8010552]

26. Arush TA, Wade LH, Shockley DC. Changes in locomotor activity, core temperature, and heart rate in response to repeated cocaine administration. Physiol Behav. 1996;60:1261 7. [PMID: 8916180]

27. Lomax P, Daniel KA. Cocaine and body temperature in the rat: effects of ambient temperature. Pharmacology. 1990;40:103 9. [PMID: 2345771]

28. Chung LL, Pechnick RN, Bova MP, George R. Acute and chronic administration of cocaine produces hyperthermia in the rat. Proc West Pharmacol Soc. 1991;34:27 8. [PMID: 1788296]

(144) [206]

Thyroxine treatment in patients with symptoms of hypothyroidism but thyroid function tests within the reference range: randomised double blind placebo controlled crossover trial

M Anne Pollock, Alison Sturrock, Karen Marshall, Kate M Davidson, Christopher J G Kelly, Alex D McMahon, E Hamish McLaren

Abstract

Objectives To determine whether thyroxine treatment is effective in patients with symptoms of hypothyroidism but with thyroid function tests within the reference range, and to investigate the effect of thyroxine treatment on psychological and physical wellbeing in healthy participants.

Design Randomised double blind placebo controlled crossover trial.

Setting Outpatient clinic in a general hospital.

Participants 25 patients with symptoms of hypothyroidism who had thyroid function tests within the reference range, and 19 controls.

Methods Participants were given thyroxine 100 µg or placebo to take once a day for 12 weeks. Washout period was six weeks. They were then given the other to take once a day for 12 weeks. All participants were assessed physiologically and psychologically at baseline and on completion of each phase.

Main outcome measures Thyroid function tests, measures of cognitive function and of psychological and physical wellbeing.

Results 22 patients and 19 healthy controls completed the study. At baseline, patients' scores on 9 out of 15 psychological measures were impaired when compared with controls. Patients showed a significantly greater response to placebo than controls in 3 out of 15 psychological measures. Healthy participants had significantly lower scores for vitality when taking thyroxine compared to placebo (mean (SD) 60 (17) *v* 73 (16), $P < 0.01$). However, patients' scores from psychological tests when taking thyroxine were no different from those when taking placebo except for a poorer performance on one visual reproduction test when taking thyroxine. Serum concentrations of free thyroxine increased and those of thyroid stimulating hormone decreased in patients and controls while they were taking thyroxine, confirming compliance with treatment. Although serum concentrations of free triiodothyronine increased in patients and controls taking thyroxine, the difference between the response to placebo and to thyroxine was significant only in the controls.

Conclusions Thyroxine was no more effective than placebo in improving cognitive function and psychological wellbeing in patients with symptoms of hypothyroidism but thyroid function tests within the reference range. Thyroxine did not improve cognitive function and psychological wellbeing in healthy participants.

Department of Clinical Biochemistry, Stobhill Hospital, Glasgow G21 3UW
M Anne Pollock *principal biochemist*
Department of Medicine, Stobhill Hospital
Alison Sturrock *senior house officer*
Christopher J G Kelly *specialist registrar*
E Hamish McLaren *consultant physician*
Department of Psychological Medicine, Gartnavel Royal Hospital, Glasgow G12 0XH
Karen Marshall *trainee clinical psychologist*
Kate M Davidson *research nurse*
Robertson Centre for Biostatistics, University of Glasgow, Glasgow G12 8QQ
Alex D McMahon *consultant statistician*
Correspondence to: M A Pollock
anne.pollock@northglasgowacs.nhs.uk

BMJ 2001;323:891–5

Details of baseline measurements and scores are available on the BMJ's website

Introduction

The classic symptoms of hypothyroidism are wide ranging and non-specific, therefore biochemical testing has become the cornerstone of diagnosis in patients for whom there is a clinical suspicion of thyroid dysfunction. However, recent anecdotal evidence has suggested there may be some clinical benefit in giving thyroxine to patients with symptoms of hypothyroidism who have thyroid function tests within the reference range.[1–3] After a series of reports in our local newspaper suggesting that such patients benefited from thyroxine therapy we treated two patients empirically with thyroxine, and they both reported symptomatic relief.[1]

To investigate this further, we conducted a double blind placebo controlled crossover trial of thyroxine in patients who had symptoms of hypothyroidism but whose thyroid function tests were within the reference range. A group of controls, who were similar in age and sex to the patient group, took part in a parallel trial. The same protocol was used for controls and patients to test the clinical belief that thyroxine treatment would have an effect on wellbeing even in participants without symptoms of hypothyroidism. We assessed response to thyroxine by using a battery of biochemical, physical, and psychological tests.

Methods

Participants

Patients were required to have had at least three of the following symptoms for six months: tiredness, lethargy, weight gain or inability to lose weight, intolerance to cold, hair loss, or dry skin or hair. We recruited patients either by referral from their general practitioner or hospital clinician, or through an article, published in a local newspaper, which described the trial and asked for volunteers. Controls were healthy volunteers recruited by personal contact with the investigators. All participants were required to have no current medical disorder, no history of thyroid disease, and recent thyroid function tests within the reference range.

Because of limited resources, the number of patients was restricted to the first 25 who met the criteria. Three patients withdrew at an early stage: one was anxious about the thyroxine dosage, one was ill, and the third failed to attend for unknown reasons. We did not enter one interviewed patient into the study because she was unwilling to have a thyrotrophin releasing hormone test. Unfortunately, we were unable to recruit sufficient controls to match the patients strictly for age and sex. However, the control group was similar to the patient group (see table A on the *BMJ*'s website). No controls were ineligible and all completed the study. The local ethical committee approved the study and we obtained informed consent from all participants.

Study protocol

The study was a randomised double blind placebo controlled crossover trial. The two treatment periods of twelve weeks each were separated by a washout period of six weeks. Half of the participants were given thyroxine for the first treatment period and placebo for the second treatment period; half were given placebo first and thyroxine second. The first 20 participants received the treatment in 100 µg capsules; thereafter 100 µg tablets were supplied. In each case a visually identical placebo was used. A 14 week supply of tablets, to be taken once a day, was provided for each phase and participants were asked to bring the remaining tablets to their assessment to assess compliance.

Randomisation was by toss of a coin in batches of four. Controls and patients attended their baseline assessment simultaneously, which led to unequal sequence groups. Two of the patients who withdrew had been assigned to the thyroxine-placebo group. The code was broken to two investigators (MAP and KMD) after each participant had completed the trial.

Evaluation

Serum thyroid stimulating hormone, free thyroxine, free triiodothyronine, cholesterol, and prolactin were measured at each visit. At baseline, serum ferritin and antithyroid peroxidase antibodies were measured and a thyrotrophin releasing hormone test was performed. Physiological and psychological assessments were performed at baseline and on completion of each phase.

At the end of the trial participants were asked to identify which treatment they thought they had received in each phase.

Biochemical measurements

Serum thyroid stimulating hormone, free thyroxine, cholesterol, and prolactin were analysed at the time the blood was collected. All blood samples were stored at −80°C and free triiodothyronine, ferritin, and antithyroid peroxidase antibodies were analysed in single batches to minimise interassay variation. Serum thyroid stimulating hormone, free thyroxine, free triiodothyronine, prolactin, and ferritin were measured by fluorescent microparticle enhanced immunoassay (Abbott Laboratories Ltd, Maidenhead, UK). An increment in thyroid stimulating hormone of >25 mU/l constituted an abnormal result in the thyrotrophin releasing hormone test. Antithyroid peroxidase antibodies were measured by a solid phase chemiluminescent enzyme immunoassay (DPC, Llanberis, UK). Cholesterol was measured on a multichannel discrete analyser (Olympus Diagnostica, Hamburg, Germany) using a cholesterol oxidase method (Randox Laboratories Ltd, Co. Antrim, UK). Interassay coefficients of variation were <15% for thyroid stimulating hormone, <9% for free thyroxine, <8.5% for free triiodothyronine, <6% for prolactin, and <2% for cholesterol.

Physical and psychological evaluation

We recorded supine blood pressure and pulse of participants using a Criticare (CE1023, Waukesha, WI, USA) non-invasive blood pressure monitor after they had rested for five minutes. We measured their weight on SECA scales (Hamburg, Germany).

We measured cognitive functioning with logical memory, verbal paired associates, visual reproduction, and digit span tests from the revised Weschler memory scale.[5] These tests assess attention and concentration, visual memory, and verbal memory. Psychomotor speed, attention, and sequencing were assessed by the trail making test.[6]

Psychological and physical wellbeing were measured by using two questionnaires. The hospital anxiety and depression scale assessed emotional disorder[7] and the SF-36 health survey measured five health concepts.[8] Two patients failed to complete one page of

Table 1 Response to thyroxine of 22 patients with symptoms of hypothyroidism but thyroid function tests within the reference range. Values are mean (SD) unless otherwise stated

Outcome	Thyroxine	Placebo	Adjusted difference† (95% CI)	P value
Biochemical measures				
Thyroid stimulating hormone (mU/l)	0.86 (0.77)	1.77 (1.21)	−1.17 (−1.76 to −0.59)	<0.001***
Free thyroxine (pmol/l)	17.95 (3.03)	13.88 (3.37)	4.75 (2.67 to 6.83)	<0.001***
Free triiodothyronine (pmol/l)	3.72 (0.68)	3.50 (0.54)	0.23 (−0.11 to 0.56)	0.177
Cholesterol (mmol/l)	8.33 (1.17)	8.27 (1.25)	0.05 (−0.27 to 0.37)	0.739
Prolactin (mU/l)	250 (158)	307 (331)	−37 (−189 to 116)	0.622
Clinical measures				
Pulse (beats/min)	86 (14)	85 (16)	−1 (−7 to 6)	0.775
Blood pressure (mm Hg)	83 (12)	80 (12)	1 (−5 to 8)	0.657
Weight (kg)	84 (19)	83 (19)	1 (−1 to 2)	0.389
Cognitive functioning scores				
Logical memory I	25 (7)	25 (7)	−1 (−3 to 1)	0.231
Logical memory II	22 (8)	21 (7)	0 (−2 to 3)	0.965
Verbal paired associates I	20 (3)	19 (4)	0 (−1 to 2)	0.599
Verbal paired associates II	7 (1)	7 (1)	0 (−1 to 0)	0.671
Visual reproduction I	30 (7)	32 (6)	−1.9 (−3.3 to −0.4)	0.016*
Visual reproduction II	27 (8)	27 (6)	1 (−2 to 3)	0.594
Digits forward	10 (1)	9 (2)	1 (0 to 1)	0.133
Digits backward	8 (3)	8 (2)	0 (−1 to 1)	0.885
Trail making test	81 (31)	80 (31)	1 (−9 to 12)	0.779
Psychological functioning scores				
Hospital anxiety and depression scale	15 (10)	15 (8)	0 (−4 to 5)	0.874
SF-36 health survey:				
Role emotional	48 (47)	64 (44)	−18 (−48 to 12)	0.220
Physical functioning	46 (32)	46 (34)	0 (−14 to 14)	0.979
Role physical	35 (41)	43 (40)	−8 (−33 to 17)	0.515
General health	42 (24)	48 (24)	−6 (−17 to 4)	0.228
Vitality	36 (27)	42 (28)	−5 (−24 to 14)	0.571

*P<0.05, ***P<0.001.
†Adjusted by subject and period effects.

the SF-36 health survey. We calculated the missing values by taking the means of their scores on their remaining two visits.

Statistical methods

Baseline characteristics were summarised by means (SD) or counts and percentages as appropriate. For each variable, participants were grouped into their sequence group (thyroxine-placebo or placebo-thyroxine) to create data summaries in relation to treatment.

In each set of participants, the effect of treatment on each variable was studied by using analysis of variance models in relation to patient, period, and treatment. The within patient treatment differences were adjusted by period and 95% confidence limits created. Some of the results were not normally distributed; therefore we repeated all analyses using non-parametric tests and rank data.

The placebo effect was calculated by using participants who received placebo as their first treatment. We calculated changes from baseline by subtracting the baseline value from placebo response.

Results

Baseline measures

All participants had thyroid function results within the reference range, with the exception of one patient who had a concentration of thyroid stimulating hormone of 5.8 mU/l (see table A on *BMJ*'s website). Concentrations of the hormone in this patient, however, normalised to 4.5 mU/l when taking placebo. Other biochemical tests for the same patient were within the reference range. Three participants—the patient described above, one control, and one patient who failed to complete the study—had exaggerated responses to thyrotrophin releasing hormone and raised concentrations of antithyroid peroxidase antibodies (range 262-1656 U/ml). One patient and one control had raised concentrations of antithyroid peroxidase antibodies (211 U/ml and 76 U/ml, respectively) but normal thyrotrophin releasing hormone responses.

Patients' scores on psychological testing were significantly impaired for logical memory I and II, mail making test, hospital anxiety and depression scale, and all components of the SF-36 health survey in comparison to healthy participants (see table B on *BMJ*'s website).

Response to thyroxine or placebo

Biochemical measures

In both groups the serum concentrations of thyroid stimulating hormone decreased and free thyroxine increased in response to thyroxine, confirming compliance with treatment. Although free triiodothyronine concentration increased in both groups when participants were taking thyroxine, the difference between the response to placebo and the response to thyroxine was only significant in the healthy group (tables 1 and 2). This finding was replicated by non-parametric analysis of the data. The response to thyroxine in patients with positive autoantibodies or abnormal thyrotrophin releasing hormone responses did not differ from that in other participants; numbers were too small for detailed analysis. No other significant biochemical changes were observed.

Table 2 Response of 19 healthy participants to thyroxine. Values are mean (SD) unless otherwise stated

Outcome	Thyroxine	Placebo	Adjusted difference† (95% CI)	P value
Biochemical measures				
Thyroid stimulating hormone (mU/l)	0.32 (0.38)	1.55 (1.54)	−1.17 (−1.80 to −0.53)	0.001**
Free thyroxine (pmol/l)	20.21 (4.33)	14.29 (5.10)	5.83 (2.02 to 9.63)	0.005**
Free triiodothyronine (pmol/l)	4.39 (1.03)	3.62 (0.52)	0.82 (0.34 to 1.31)	0.002**
Cholesterol (mmol/l)	5.03 (0.84)	5.22 (0.78)	−0.21 (−0.46 to 0.04)	0.095
Prolactin (mU/l)	195 (71)	211 (118)	−14 (−73 to 46)	0.636
Clinical measures				
Pulse (beats/min)	77 (12)	75 (12)	3 (−2 to 8)	0.222
Blood pressure (mm Hg)	74 (9)	72 (11)	3 (−3 to 9)	0.358
Weight (kg)	66 (12)	67 (12)	−1 (−1 to 0)	0.240
Cognitive functioning scores				
Logical memory I	34 (5)	33 (6)	1 (−1 to 4)	0.380
Logical memory II	31 (7)	31 (6)	1 (−2 to 3)	0.435
Verbal paired associates I	20 (3)	21 (3)	−1 (−1 to 0)	0.128
Verbal paired associates II	7 (1)	7 (1)	0 (−1 to 0)	0.413
Visual reproduction I	35 (4)	36 (2)	−0.8 (−1.8 to 0.3)	0.148
Visual reproduction II	34 (4)	34 (4)	0 (−1 to 2)	0.723
Digits forward	10 (2)	10 (2)	0 (−1 to 1)	0.605
Digits backward	8 (3)	9 (2)	0 (−1 to 0)	0.338
Trail making test	57 (19)	54 (12)	3 (−3 to 9)	0.368
Psychological functioning scores				
Hospital anxiety and depression scale	9 (7)	8 (5)	2 (0 to 5)	0.095
SF-36 health survey:				
Role emotional	82 (28)	92 (19)	−7 (−21 to 7)	0.318
Physical functioning	91 (14)	93 (10)	−1 (−8 to 3)	0.515
Role physical	79 (27)	89 (19)	−10 (−26 to 6)	0.216
General health	80 (12)	84 (10)	−5 (−11 to 2)	0.135
Vitality	60 (17)	73 (16)	−13 (−22 to −4)	0.007**

**P<0.01.
†Adjusted by subject and period effects.

Physical health and psychological measures

We compared the differences in the scores at baseline and after placebo in participants taking placebo first. The patients showed a significant symptomatic improvement in the general health, role physical, and hospital anxiety and depression scale scores after placebo, compared with healthy participants (mean (95% confidence interval) 8 (2 to 15) *v* −8 (−14 to −3), 25 (4 to 46) *v* −9 (−32 to 13), and −7 (−11 to −3) *v* −1 (−2 to 0), respectively; table 3). We observed no changes in measures of cognitive function. There was no placebo effect with regard to psychological or cognitive function scores for the controls (table 3).

In the comparison of the response to thyroxine or placebo, the only difference in cognitive function tests was in the patient group, which showed a significant improvement in visual reproduction I score with placebo (32 (6) *v* 30 (7), P = 0.016; table 1). The difference was less significant on non-parametric analysis (P = 0.035) and was not replicated in the delayed recall test, which suggests that the finding may be spurious. No other tests of cognitive or psychological function showed differences (table 1). The vitality score for the healthy control group was significantly better with placebo compared with thyroxine (73 (16) *v* 60 (17), P = 0.007; table 2), but otherwise no differences were observed.

Clinical measures

No significant changes occurred in patients with respect to blood pressure, pulse rate, or weight during the study.

Table 3 Effect of placebo when given first for patients with symptoms of hypothyroidism but thyroid function tests within the reference range and controls. Values are mean or mean difference (95% confidence intervals)

Outcome	Patients (n=14)			Controls (n=9)			Patients *v* controls (P value)
	Baseline	Placebo	Placebo-baseline	Baseline	Placebo	Placebo-baseline	
Biochemical measures							
Thyroid stimulating hormone (mU/l)	1.64	1.78	0.08 (−0.47 to 0.35)	1.25	1.08	−0.17 (−0.42 to 0.08)	0.685
Free thyroxine (pmol/l)	13.93	14.50	0.50 (−1.75 to 2.75)	14.83	15.25	0.83 (−0.83 to 1.88)	0.932
Free triiodothyronine (pmol/l)	3.20	3.56	0.51 (0.17 to 0.84)	3.55	3.86	0.29 (−0.23 to 0.80)	0.107
Cholesterol (mmol/l)	6.19	6.20	−0.13 (−0.40 to 0.14)	4.99	5.14	0.15 (−0.28 to 0.58)	0.212
Prolactin (mU/l)	286	339	56 (−142 to 254)	240	193	−46 (−94 to 2)	0.410
Clinical measures							
Pulse (beats/min)	79	82	5 (−2 to 12)	68	72	5 (−4 to 13)	0.918
Blood pressure (mm Hg)	81	79	−2 (−6 to 1)	77	69	−8 (−18 to 3)	0.175
Weight (kg)	85	85	−1 (−2 to 1)	67	67	0 (−1 to 0)	0.627
Cognitive functioning							
Logical memory I	19	28	6 (3 to 9)	25	34	5 (−1 to 12)	0.801
Logical memory II	15	20	5 (2 to 8)	28	31	5 (−2 to 11)	0.777
Verbal paired associates I	16	19	3 (2 to 4)	19	21	2 (0 to 4)	0.211
Verbal paired associates II	7	7	0 (0 to 1)	8	8	0 (−1 to 1)	0.526
Visual reproduction I	32	34	2 (0 to 4)	35	35	1 (−2 to 4)	0.555
Visual reproduction II	28	29	1 (−3 to 5)	33	35	2 (−2 to 7)	0.658
Digits forward	9	9	1 (−1 to 2)	10	10	1 (0 to 1)	0.862
Digits backward	7	8	1 (0 to 2)	10	10	0 (−2 to 1)	0.124
Trail making test	97	82	−16 (−29 to −2)	65	53	−11 (−22 to −1)	0.655
Psychological functioning							
Hospital anxiety and depression scale	20	13	−7 (−11 to −3)	8	7	−1 (−2 to 0)	0.034*
SF-36 health survey:							
Role emotional	53	71	21 (−4 to 47)	79	80	1 (−17 to 19)	0.240
Physical functioning	42	50	8 (−7 to 23)	98	97	−1 (−2 to 1)	0.354
Role physical	18	45	25 (4 to 46)	97	88	−9 (−32 to 13)	0.031*
General health	40	50	8 (2 to 15)	94	86	−8 (−14 to −3)	0.001**
Vitality	29	47	16 (−4 to 35)	71	71	0 (−7 to 7)	0.214

*P<0.05. **P<0.01.

At the end of the study neither group was able to identify accurately which treatment period was thyroxine or placebo (table 4).

Discussion

This is the first randomised double blind placebo controlled trial of thyroxine treatment in patients who have symptoms of hypothyroidism but are biochemically euthyroid.

Biochemical results

Compliance was confirmed in both groups by the rise in free thyroxine and fall in thyroid stimulating hormone while participants were taking thyroxine. The lack of significant increase in free triiodothyronine in patients taking thyroxine might reflect impairment of the peripheral conversion of thyroxine to triiodothyronine. Although this finding requires further investigation, anecdotal evidence suggests that patients could benefit from thyroxine treatment alone.[9]

Comparison of thyroxine and placebo treatments

Psychological testing showed that patients differed from the controls at baseline. Cognitively, they scored worse on immediate and delayed verbal recall and had slower motor movements. They also perceived themselves to have poorer general health, more fatigue, increased problems with routine tasks and activities related to work, and higher levels of anxiety and depression. These findings may be consistent with a depressive illness, although no formal assessment was performed.

Controls showed no significant changes in psychological measurements after treatment with either thyroxine or placebo. This suggests that, contrary to widespread belief, thyroxine does not have a non-specific effect on wellbeing. In the participants who received placebo first, patients showed a small but significant improvement in general health, physical wellbeing, and anxiety and depression after placebo when compared with baseline. Thyroxine treatment, however, had no greater effect than placebo in this group of patients. This contrasts with previous studies in biochemically hypothyroid patients, where thyroxine treatment was associated with psychological improvement.[10–12]

Table 4 Participants' ability to distinguish between thyroxine and placebo at end of trial

	Patient group	Control group
Correct	9	6
Wrong	10	8
"Don't know"	3	5

Numbers in the study

The small number of participants in this preliminary study means that, although there was no significant difference between placebo and thyroxine in 13 of the 14 well validated psychological tests, the power of the study may not have been sufficient to eliminate definitively a possible biological effect of thyroxine. If this was the case we would have expected to see a trend in favour of thyroxine over placebo in the test results, especially as a recent open intervention study of thyroxine (mean dose 125 µg daily) reported self

Recent anecdotal accounts suggest that patients with symptoms of hypothyroidism but who are biochemically euthyroid may benefit from thyroxine treatment

No controlled trials in this area have been reported

This study suggests that thyroxine is no more effective than placebo in improving psychological and physical wellbeing in patients who show symptoms of being clinically hypothyroid but whose thyroid function tests are within the reference range

Thyroxine replacement did not improve psychological and physical wellbeing in healthy participants

assessed improvements in energy and poor memory in 80% of 139 participants.[4] Our study showed no discernible trend (table 1).

Conclusion

We can find no support for the hypothesis that people with symptoms of hypothyroidism but thyroid function tests within the reference range benefit from treatment with 100 μg thyroxine daily. However, our results require confirmation in a larger study. The improvement noted anecdotally and in open studies may be due to the placebo effect shown in our study.

We are grateful to the staff of the clinical biochemistry department at Stobhill Hospital for performing the biochemical analyses and to the staff of the pharmacy department for performing the randomisation and preparing the capsules. We thank Goldshield Pharmaceuticals for supplying the placebo tablets.

Contributors: MAP conceived the study and coordinated the laboratory component. MAP, EHMcL, CJGK and KMD designed the study. KMD and KM chose the cognitive function tests and questionnaires, and KM tested all participants. AS coordinated all contact with the study participants and undertook the clinical assessment of the participants at each visit. ADMcM performed the statistical analysis and provided further statistical advice. All authors were involved in writing the paper, with MAP providing coordination. EHMcL is guarantor for the study.

Funding: MAP received a scientific development scholarship from the Association of Clinical Biochemists.

Competing interests: None declared.

1 Skinner GRB, Thomas R, Taylor M, Sellarajah M, Boh S, Kretz S, et al. Thyroxine should be tried in clinically hypothyroid, but biochemically euthyroid patients. *BMJ* 1997;314:1764.
2 Williams G. Distinguishing hypothyroid symptoms from common non-specific complaints is difficult. *BMJ* 1997;315:814.
3 Holmes, Diana. *Tears behind closed doors*. London: Avon, 1998.
4 Mclaren EH, Kelly CJG, Pollock MA. Trial of thyroxine treatment for biochemically euthyroid patients has been approved. *BMJ* 1997;315:1465.
5 Wechsler D. *Wechsler memory scale–revised manual*. San Antonio: Psychological Corporation, 1987.
6 Rehan RM. A research programme on the psychological effects of brain lesions in human beings. In: Ellis NR, ed. *International review of research in mental retardation*. New York: Elsevier, 1966.
7 Zigmund AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983;67:361-70.
8 Ware JE. *SF36 health survey: manual and interpretation guide*. Health Institute: New England Medical Centre, 1997.
9 Skinner GRB, Holmes D, Ahmad A, Davies JA, Benfield J. Clinical response to thyroxine sodium in clinically hypothyroid but biochemically euthyroid patients. *J Nutr Environ Med* 2000;10:115-24.
10 Beckwith BC, Tucker DM. Thyroid disorders. In: Turner RE, Van Thiel DH, Edwards KL, eds. *Medical neuropsychology: the impact of disease on brain/mind*. New York: Plenum Press, 1998.
11 Denicoff KD, Joffe RT, Lakshmanan MC, Robbins J, Rubinow DR. Neuropsychiatric manifestations of altered thyroid state. *Am J Psychiatr* 1990;147:94-9.
12 Osterweil D, Syndulko K, Cohen SN. Cognitive function in non-demented older adults with hypothyroidism. *J Am Geriatr Soc* 1992;40:325-35.

(Accepted 8 May 2001)

A patient who changed my practice

"That's very nice, but will it get me pregnant?"

I have to confess, I didn't understand the question at the time. A year later, I am starting to. I was an eager new general practitioner and member of a research team investigating how the internet might influence patients' use of primary care, and we had just begun offering free internet access at our deprived inner city practice. I had filled the waiting room with posters advertising the service, and a member of staff was on hand to guide newcomers through the web. With our new PC and a cup of coffee, we would help patients empower themselves with the information they'd been waiting for.

The patient's recent blood tests for infertility had indicated that she had polycystic ovary syndrome, and she had come for her follow up appointment. I explained the likely diagnosis, suggested she might use our internet service to find more detailed information, and offered to demonstrate what was available. Going straight to a website that I knew had some excellent patient information on the syndrome, I briefly talked her through it and gave her a printed copy to take away. She and I had always had a good rapport, I believed. Now she looked disdainfully, first at me and then at the sheets of paper I offered. "That's very nice," she said, "but will it get me pregnant?"

In my enthusiasm, I had completely misunderstood what the patient wanted from the consultation. She wanted to conceive and wasn't in the least interested in explanations of why she had failed to thus far—she had come to get a prescription, not information. However helpful our internet service was intended to be, she didn't want that kind of help. Neither, it seems, did many of our 15 000 patients, only nine of whom used the internet service in the three months before we closed it. (All but one of the nine were well educated, had good jobs, and had already used the internet—not representative of residents of inner city Manchester.)

Since this rather awkward encounter, I have noticed that many of my assumptions about what my patients want are equally mistaken. Very few want information from me. Many don't even want a prescription: they want money, or a job, or an escape from their often appalling lives. I think the internet is a valuable source of patient information, but I've been reminded of a maxim drummed into me by the first consultant who taught me clinical medicine: "Assume nothing about your patients—check everything, every time."

Our research group is conducting a questionnaire and interview study to investigate the influence of deprivation on attitudes to health information and the internet. Meanwhile, I persist in trying to persuade my patients to expect a little more from me than hissing letters and prescriptions. But I try to assume less before opening my mouth or offering information.

Robert Varnam *general practitioner, Robert Darbishire Practice, Rusholme Health Centre, Manchester (robert.varnam@man.ac.uk)*

(103) [1497]

Effects of Diet and Simvastatin on Serum Lipids, Insulin, and Antioxidants in Hypercholesterolemic Men

A Randomized Controlled Trial

Antti Jula, MD, PhD
Jukka Marniemi, PhD
Risto Huupponen, MD, PhD
Arja Virtanen, MSc
Merja Rastas, MSc
Tapani Rönnemaa, MD, PhD

Context Limited information exists on the interaction between diet and 3-hydroxy-3-methylglutaryl coenzyme A reductase inhibitors (statins) and the interaction's effect on serum lipid and lipoprotein levels, insulin sensitivity, and circulating antioxidant vitamin and provitamin levels.

Objective To evaluate the separate and combined effects of diet and simvastatin therapy on serum levels of lipids, lipoproteins, antioxidants, and insulin.

Design, Setting, and Participants Randomized, controlled crossover trial conducted from August 1997 to June 1998 in 120 previously untreated hypercholesterolemic men aged 35 to 64 years who were recruited from the community in Turku, southwestern Finland.

Interventions After a 4- to 6-week placebo run-in period, participants were randomly allocated to a habitual diet (n=60) or dietary treatment group (n=60), and each of these groups was further randomized in a double-blind crossover fashion to receive simvastatin (20 mg/d) or placebo, each for 12 weeks (n=30 in each group). The main goals of the dietary treatment were to reduce energy intake from saturated plus *trans*-unsaturated fats to no more than 10% by replacing them partly with monounsaturated and polyunsaturated fats rich in omega-3 fatty acids and to increase intake of fruits, vegetables, and dietary fiber.

Main Outcome Measures Changes in levels of total, low-density lipoprotein (LDL), and high-density lipoprotein (HDL) cholesterol; triglycerides; apolipoprotein B; insulin; glucose; and antioxidants at week 12 of each treatment period, compared among the 4 groups.

Results Dietary treatment decreased levels of total cholesterol by 7.6% ($P<.001$), LDL cholesterol by 10.8% ($P<.001$), HDL cholesterol by 4.9% ($P=.01$), apolipoprotein B by 5.7% ($P=.003$), serum insulin by 14.0% ($P=.02$), and α-tocopherol by 3.5% ($P=.04$). Simvastatin decreased levels of total cholesterol by 20.8%, LDL cholesterol by 29.7%, triglycerides by 13.6%, apolipoprotein B by 22.4%, α-tocopherol by 16.2%, β-carotene by 19.5%, and ubiquinol-10 by 22.0% ($P<.001$ for all) and increased levels of HDL cholesterol by 7.0% ($P<.001$) and serum insulin by 13.2% ($P=.005$). Glucose levels remained unchanged in all groups. The effects of dietary treatment and simvastatin were independent and additive.

Conclusions A modified Mediterranean-type diet rich in omega-3 fatty acids efficiently potentiated the cholesterol-lowering effect of simvastatin, counteracted the fasting insulin–elevating effect of simvastatin, and, unlike simvastatin, did not decrease serum levels of β-carotene and ubiquinol-10.

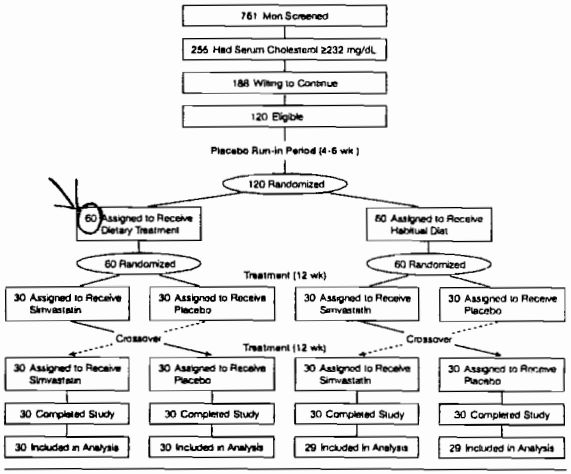
JAMA. 2002;287:598-605 www.jama.com

Author Affiliations: Research and Development Centre of the Social Insurance Institution, Turku, Finland (Drs Jula and Marniemi and Mss Virtanen and Rastas); Department of Pharmacology and Clinical Pharmacology (Dr Huupponen), and Department of Medicine (Dr Rönnemaa), University of Turku, Finland. **Corresponding Author and Reprints:** Antti Jula, MD, PhD, Research and Development Centre of the Social Insurance Institution, Peltolantie 3, FIN-20720 Turku, Finland (e-mail: antti.jula@kela.memonet.fi).

erides and an increase in serum high-density lipoprotein cholesterol (HDL-C) concentration. Effects of a modified Mediterranean-type diet low in saturated fatty acid intake and high in omega-3 fatty acid intake may be mediated mainly by factors other than those associated with serum lipid levels.[4]

Only limited information exists on the effects of HMG-CoA reductase inhibitors and dietary therapy on insulin sensitivity, circulating antioxidant vitamin and provitamin levels, LDL oxidation, and their interactions on circulating lipids and lipoproteins. Apparently, no data exist on the interaction between diet and drug treatment and insulin sensitivity and antioxidants. The aim of this study was to characterize the effects of an HMG-CoA reductase inhibitor (simvastatin) and a diet low in saturated fatty acids and enriched in monounsaturated and polyunsaturated fatty acids (especially α-linolenic acid), cereals, fruits, berries, and vegetables on serum lipids, glucose, insulin, and antioxidants.

Figure 1. Study Design



METHODS

Subjects and Study Design

Previously untreated hypercholesterolemic men 35 to 64 years of age were screened from the clients of the occupational health service of 5 industrial plants and government offices in Turku in southwestern Finland. Subjects with a fasting serum cholesterol concentration of at least 232 mg/dL (≥6.0 mmol/L) at screening were invited for briefing about the study. After the subjects had given their informed consent, their fasting serum cholesterol, triglyceride, and glucose concentrations were measured and routine biochemical tests were performed. An electrocardiogram was taken, and blood pressure, weight, and height were measured. An internist performed a physical examination and checked questionnaires for medical history and cardiovascular symptoms. If fasting serum cholesterol concentration was between 232 and 309 mg/dL (6.0 and 8.0 mmol/L) and fasting serum triglyceride concentration was no higher than 266 mg/dL (3.0 mmol/L), the subject could be included in the study. Subjects with a body mass index higher than 32 kg/m^2, coronary artery disease, cerebrovascular disease, claudication and pharmacologically treated hypertension, hyperlipidemia, or diabetes were excluded from the study.

Subjects included in the study entered first a 4- to 6-week open placebo run-in period, at the end of which they were randomly allocated to a habitual diet or a dietary treatment group (FIGURE 1). In both groups, a second randomization was performed, and the subjects received simvastatin (20 mg/d) or a matching placebo for 12 weeks in a double-blind, crossover fashion. A washout period was not included, since no period or carryover effects were seen in a preceding pilot study of 20 men. The sample size was calculated with the assumption that a difference of 15 mg/dL (0.4 mmol/L) in primary outcome variables (cholesterol and LDL-C) can be detected with 80% power and 5% type I error (n=88). To ensure a sufficient sample size, a total of 120 subjects were included in the study. All subjects completed the study.

The study was conducted according to the latest revision of the Declaration of Helsinki and was approved by the Ethical Committee of the Social Insurance Institution of Finland.

Measurements and Analyses

Blood pressure and weight were measured, diet was recorded, physical exercise frequency and intensity were determined, and 12-hour fasting blood samples were taken before randomization at the end of the placebo run-in period (baseline) and at the end of both 12-week drug-treatment periods. Two blood samples were taken 1 week apart at the end of each period. All measurements and analyses were done blinded to the treatment allocation of the subject. The serum samples were frozen and stored at −70°C until assayed. The baseline and follow-up samples were analyzed always in 1 analytical run. Subjects' body weight was measured while they wore light clothing and no shoes, with an ac-

curacy of 0.1 kg; height measurements had an accuracy of 1 cm. Seated blood pressure was measured by a trained nurse with a mercury sphygmomanometer and averaged across 2 readings. Diet was monitored through 7-day food records by using household measures. The records were analyzed by means of the Nutrica (Social Insurance Institution, Turku, Finland) food and nutrient calculation software and the databases on the nutrient composition of Finnish food.[6] Leisure-time physical activity was assessed by a questionnaire with questions about average frequency (5-point scale: 0, >0 but <1, 1, 2, ≥3 times per week) and intensity (4-point scale: no physical exercise, 0; exercise does not cause sweating or labored breathing, 1; exercise causes sweating and some degree of labored breathing, 2; exercise causes strong sweating and labored breathing, 3) during the preceding 12 weeks.

Serum cholesterol, HDL-C, and triglyceride concentrations were averaged across 2 fasting blood samples taken at a 1-week interval at the end of each period. Concentrations were determined by enzymatic methods (Merck Diagnostica, Darmstadt, Germany). High-density lipoprotein cholesterol was analyzed after the magnesium-phosphotungstate precipitation of very low-density lipoprotein and LDL. Except for diene conjugation and total peroxyl radical trapping antioxidant potential analyses, LDL-C content was estimated by the Friedewald formula.[7] Apolipoprotein A1 and B concentrations were determined immunoturbidimetrically (Orion Diagnostica, Espoo, Finland). Serum glucose was analyzed enzymatically (Merck Diagnostica). Serum insulin was determined by microparticle enzyme immunoassay (Abbott Laboratories, Dainabot, Tokyo, Japan). The homeostasis model assessment formula was used to assess insulin resistance as follows: insulin resistance = (fasting insulin × fasting glucose level)/22.5.[8] Serum α-tocopherol (the most active form of vitamin E) and β-carotene were analyzed simultaneously by high-performance liquid chromatography.[9] The concentration of serum ascorbic acid (vitamin C) was determined spectrophotometrically.[10] Erythrocyte folate levels were assayed by radioimmunoassay (ICN Pharmaceuticals, Orangeburg, NY). Concentrations of serum ubiquinol-10 were measured by high-performance liquid chromatography with spectrophotometric detection.[11] Serum homocysteine concentrations were determined by fluorescence polarization immunoassay (Abbott Laboratories, Abbott Park, Ill) after enzymatic conversion of total homocysteine to *S*-adenosyl-*L*-homocysteine. Serum LDL fraction for determinations of diene conjugation and total peroxyl radical trapping antioxidant potential was isolated with buffered heparin.[12] Oxidation of LDL was estimated by measuring spectrophotometrically the baseline level of diene conjugation in LDL particles.[13] The antioxidant potential of isolated LDL samples was measured luminometrically in vitro.[13]

Table 1. Baseline Characteristics and Circulating Antioxidants of Men Randomized to the Dietary Treatment and Habitual Diet Groups*

Characteristics	Mean (SD)		*P* Value
	Habitual Diet	Dietary Treatment	
Subjects, No.	60	60	
Age, y	48.4 (6.2)	48.0 (6.2)	.72
Weight, kg	81.4 (9.7)	82.4 (9.3)	.57
Body mass index, kg/m²	25.6 (2.5)	25.9 (2.1)	.40
Systolic blood pressure, mm Hg	126.5 (14.2)	124.3 (12.4)	.36
Diastolic blood pressure, mm Hg	82.7 (8.6)	81.9 (8.5)	.63
Current smoker, %	33.3	21.7	.15
Total cholesterol, mg/dL	259 (24)	250 (21)	.04
LDL-C, mg/dL	183 (23)	175 (22)	.05
HDL-C, mg/dL	49 (12)	52 (12)	.12
Triglycerides, mg/dL	150 (58)	135 (56)	.11
Apolipoprotein A1, mg/dL	135 (19)	137 (20)	.60
Apolipoprotein B, mg/dL	139 (21)	129 (17)	.01
Glucose, mg/dL	97 (9)	97 (11)	.73
Insulin, μU/mL	5.5 (2.4)	5.7 (2.6)	.61
α-Tocopherol, mg/dL	1.39 (0.24)	1.37 (0.26)	.53
β-Carotene, μg/dL	68 (53)	63 (40)	.74
Ascorbic acid, mg/dL	1.13 (0.32)	1.09 (0.27)	.46
Erythrocyte folate, ng/mL	357 (98)	385 (143)	.28
Homocysteine, mg/L	1.39 (0.28)	1.42 (0.46)	.90
Ubiquinol-10, μmol/L	1.46 (0.57)	1.48 (0.52)	.77
LDL TRAP, μmol/L†	104 (26)	101 (21)	.60
LDL TRAP, μmol/mmol‡	24.6 (5.1)	25.2 (4.8)	.42
LDL diene conjugation, μmol/L†	39.9 (10.8)	34.3 (11.6)	.003
LDL diene conjugation, μmol/mmol‡	9.5 (2.2)	8.5 (2.4)	.01

*LDL-C indicates low-density lipoprotein cholesterol; HDL-C, high-density lipoprotein cholesterol; LDL TRAP, total peroxyl radical trapping potential of LDL (antioxidant potential of isolated LDL particles); and LDL diene conjugation, oxidation products of fatty acids in LDL particles. To convert cholesterol to mmol/L, multiply values by 0.0259; to convert triglycerides to mmol/L, multiply values by 0.0113; to convert glucose to mmol/L, multiply values by 0.0555; and to convert insulin to pmol/L, multiply values by 6.945.
†Expressed per liter of original serum.
‡Expressed per millimole of isolated LDL-C.

Drugs

Capsules containing simvastatin or placebo were prepared in a local pharmacy according to the European Pharmacopoeia. Commercially available tablets containing 20 mg of simvastatin were powdered and mixed with wheat starch. One hundred tablets were used in each manufacturing run. The mixture was divided into 100 gelatin capsules with an accuracy of 10%. Placebo capsules indistinguishable from the active drug capsules contained wheat starch and additives to guarantee the blinding. Each patient received his or her own bottle containing the capsules for each study period. Compliance with the drug treatment was controlled by counting the number of returned capsules.

Dietary Treatment

The targeted composition of the weight-stable, modified, Mediterranean–type diet was the following: no more than 10% energy from saturated and *trans*-unsaturated fatty acids; cholesterol intake no more than 250 mg/d; omega-3 fatty acid intake of plant origin (α-linolenic acid) and marine origin of at least 4 g/d and the ratio of omega-6/omega-3 polyunsaturated fatty acids less than 4; and increased intakes of fruits, vegetables, and soluble fiber.

The subjects randomized to the dietary treatment were advised to use leaner meat products, low-fat cheese, skim milk, fat-free sour milk, and low-fat yogurt. Fish was recommended as a main meal once or twice a week. Rapeseed margarine was recommended as a replacement for butter, a mixture of butter and vegetable oils, and sunflower margarine. Rapeseed margarine and oil, oat bran (20 g/d), and frozen berries (blueberry, lingonberry, or black currant at 50 g/d) were supplied free to study subjects. The experimental diet was supervised by a nutritionist in 1 individual session and in 2 group counseling sessions at the beginning of the treatment and in 5 subsequent monthly group brush-up sessions during the dietary treatment.

The subjects randomized to the habitual diet group were advised to continue eating their usual diet during the study period.

Statistical Analysis

Baseline (end of the placebo run-in period) comparisons between the dietary treatment and habitual diet groups were made with a *t* test for continuous variables and by a χ^2 test for categorical variables to verify the success of the randomization. Analysis of variance for repeated measures of variance, with contrasts between baseline and simvastatin or placebo treatment periods, was used to test the significance of dietary changes within the dietary treatment and habitual diet groups. The analysis of variance model was fitted separately to the dietary treatment and habitual diet groups, where period and carryover effects were tested. Because no period or carryover effects were observed and baseline values affected the outcome, repeated analyses of covariance with baseline values as covariates, dietary treatment and habitual diet as intersubject factors, and placebo and simvastatin treatment as intrasubject factors were included in the final models. Validity of the models was evaluated with residual analysis. Normality of residuals was checked with the Shapiro-Wilk statistics and constancy of residuals by a graphic analysis. Log or square root transformations were applied if necessary. Because statistical inferences after transformation were unchanged, raw results are reported. The association between triglyceride and insulin was tested by repeated analysis of covariance with triglyceride as the variable covariate, baseline insulin as the fixed covariate, drug treatment (placebo or simvastatin) as the intrafactor, and dietary treatment (dietary treatment or habitual diet) as the interfactor. Polytomous response models were used to test changes in the frequency and intensity of leisure-time physical activity. The data are given as mean (SE) values with 95% confidence intervals for the mean changes. One subject with a nonfasting blood sample at baseline was not included in the analyses (Figure 1). We set .05 as the level of significance. All statistical analyses were conducted with SAS version 6.12 (SAS Institute, Cary, NC).

RESULTS

The baseline characteristics of subjects randomized to the dietary treatment or habitual diet groups are summarized in TABLE 1 and TABLE 2.

Table 2. Dietary Intake of Cholesterol, Fatty Acids, Fiber, Vitamin E, Ascorbic Acid, and β-Carotene*

Variable	Habitual Diet, Mean (SD)			Dietary Treatment, Mean (SD)		
	Baseline	12-Week Placebo	12-Week Simvastatin	Baseline	12-Week Placebo	12-Week Simvastatin
Cholesterol, mg/d	322 (102)	313 (101)	340 (111)	298 (92)	214 (82)†	225 (74)†
Fat intake, % of total energy intake	36.5 (4.7)	36.9 (4.6)	38.1 (5.8)‡	34.3 (4.8)	34.8 (5.6)	34.7 (4.5)
SAFA, % of total energy intake	14.4 (2.5)	14.8 (2.7)	15.2 (3.6)‡	13.5 (2.4)	9.3 (2.1)†	9.6 (2.0)†
MUFA, % of total energy intake	12.6 (2.1)	12.6 (1.9)	13.1 (2.4)	11.8 (1.7)	14.1 (3.0)†	13.9 (2.6)†
PUFA, % of total energy intake	5.8 (1.6)	5.9 (1.5)	5.9 (1.6)	5.7 (1.2)	8.1 (1.6)†	8.0 (1.4)†
Omega-6/omega-3 fatty acid ratio	4.9 (1.8)	4.9 (1.6)	4.8 (1.6)	4.5 (1.5)	2.9 (0.7)†	3.0 (1.0)†
Linolenic acid, g/d	1.18 (0.55)	1.11 (0.54)	1.23 (0.60)	1.01 (0.35)	3.57 (1.45)†	3.57 (1.53)†
Linolic acid, g/d	7.24 (2.44)	6.77 (2.37)	7.52 (2.88)	6.68 (2.53)	11.45 (3.59)†	11.46 (3.31)†
Linolenic/linolic acid ratio	0.17 (0.07)	0.17 (0.05)	0.17 (0.06)	0.16 (0.05)	0.31 (0.06)†	0.31 (0.08)†
Fiber, g/d	20.9 (6.2)	19.6 (6.1)	19.4 (6.4)‡	20.0 (6.7)	27.2 (7.8)†	27.6 (7.8)†
Ascorbic acid, mg/d	59 (33)	60 (42)	64 (45)	62 (38)	93 (46)†	98 (46)†
Vitamin E, mg/d	10.2 (3.3)	10.0 (3.2)	10.0 (2.9)	9.5 (2.9)	14.6 (4.0)†	14.9 (4.3)†
β-Carotene, μg/d	2453 (1580)	1693 (1100)†	1992 (1540)‡	2093 (1298)	2420 (1751)	2239 (1442)

*SAFA indicates saturated fatty acids; MUFA, monounsaturated fatty acids; and PUFA, polyunsaturated fatty acids.
†$P<.001$ compared with baseline.
‡$P<.05$ compared with baseline.

Compliance with the drug treatment was good: subjects in the dietary and habitual diet groups took 91% to 95% of the prescribed capsules during the placebo run-in period and the 12-week drug treatment.

In the dietary treatment group, mean (SD) body weight was 82.4 (9.3) kg at baseline and 82.7 (9.5) and 82.8 (9.2) kg after 12 weeks' treatment with placebo and simvastatin, respectively. In the habitual diet group, mean body weight was 81.2 (9.7), 82.1 (9.6), and 82.1 (9.8) kg at baseline and after placebo and simvastatin treatments, respectively. The small weight gain was not associated with simvastatin or dietary treatment (analysis of covariance).

On average, the dietary treatment group achieved the predetermined target values (Table 2). Daily intake of cholesterol fell to less than 250 mg. The proportion of fats in total energy intake remained unchanged. Energy derived from saturated fatty acids decreased to less than 10%. The percentages of energy from monounsaturated and polyunsaturated fatty acids increased, reflecting decreased saturated fatty acid intake and increased intake of rapeseed oil. The mean ratio of omega-6 to omega-3 polyunsaturated fatty acids fell to 3 or less. The intake of linolenic acid nearly quadrupled, and that of linolic acid nearly doubled, resulting in a 2-fold linolenic to linolic acid ratio. Dietary intake of fiber, ascorbic acid, and vitamin E increased because of increased daily intake of oat bran (17 g), bread (15 g), vegetables (6 g), fruits (1 g), and berries (46 g). In the habitual diet group, nutrient intake remained virtually unchanged.

In the habitual diet and dietary treatment groups, the frequency ($P=.42$) and intensity ($P=.58$) of physical activity did not change from baseline during placebo and simvastatin treatment.

Serum Lipids, Glucose, Insulin, and Blood Pressure

Dietary treatment decreased average serum cholesterol concentration by 7.6%, LDL-C by 10.8%, HDL-C by 4.9%, and apolipoprotein B by 5.7% (TABLE 3). The treatment also decreased insulin levels by 14.0% and insulin resistance by 15.1% (FIGURE 2). Serum triglyceride, apolipoprotein A1, and glucose levels remained unchanged.

Simvastatin treatment decreased average serum cholesterol concentration by 20.8%, triglyceride levels by 13.6%, and apolipoprotein B levels by 22.4%. The treatment increased HDL-C levels by 7.0% and apolipoprotein A1 levels by 2.4% (Table 3). It also increased insulin levels and insulin resistance and decreased LDL-C levels (Figure 2). Glucose levels remained unchanged.

The combined effect of diet and simvastatin on serum lipid, lipoprotein, glucose, and insulin levels was equal to the sum of the components (Table 3, Figure 2). For HDL-C, fasting serum insulin levels, and insulin resistance, the effects were opposite, with simvastatin increasing the levels and dietary treatment decreasing them. Dietary treatment ($P=.049$) and a decrease in triglyceride levels ($P<.001$) were associated with a decrease in serum insulin levels (analysis of covariance); simvastatin ($P<.001$) was associated with an increase.

Neither dietary treatment nor simvastatin affected blood pressure (data not shown).

Table 3. Effects of Dietary Treatment and Simvastatin*

Variable	Mean (SE) [95% CI]		P Value†		
	Dietary Treatment – Habitual Diet	Simvastatin – Placebo	Dietary Effect	Simvastatin Effect	Interaction
Metabolic Variables					
Cholesterol, mg/dL	−19 (3) [−26 to −12]	−53 (2) [−58 to −48]	<.001	<.001	.37
LDL-C, mg/dL	−19 (3) [−25 to −14]	−53 (2) [−57 to −49]	<.001	<.001	.57
HDL-C, mg/dL	−2 (1) [−4 to −0.4]	3 (0.4) [2 to 5]	.01	<.001	.55
Triglycerides, mg/dL	−1 (5) [−12 to 10]	−19 (4) [−27 to −12]	.90	<.001	.45
Apolipoprotein A1, mg/dL	−3 (2) [−7 to 0]	3 (1) [0 to 6]	.08	.007	.12
Apolipoprotein B, mg/dL	−8 (2) [−13 to −3]	−30 (1) [−33 to −27]	.003	<.001	.05
Fasting serum glucose, mg/dL	−0.5 (1) [−3 to 1]	1 (1) [0 to 2]	.52	.14	.95
Fasting plasma insulin, μU/mL	−0.78 (0.32) [−1.42 to −0.15]	0.74 (0.26) [0.22 to 1.26]	.02	.005	.36
HOMA IR	−0.20 (0.08) [−0.37 to −0.04]	0.19 (0.07) [0.06 to 0.33]	.02	.006	.32
Circulating Variables of Antioxidation					
α-Tocopherol, mg/dL	−0.05 (0.02) [−0.10 to −0.002]	−0.22 (0.02) [−0.26 to −0.19]	.04	<.001	.41
β-Carotene, μg/dL	5 (4) [−2 to 12]	−13 (2) [−16 to −9]	.16	<.001	.07
Ascorbic acid, mg/dL	0.01 (0.03) [−0.05 to 0.08]	0.03 (0.02) [−0.02 to 0.07]	.65	.24	.90
Erythrocyte folate, ng/mL	3 (14) [−30 to 25]	11 (8) [−5 to 27]	.84	.18	.80
Homocysteine, mg/L	0.04 (0.02) [−0.01 to 0.09]	−0.03 (0.02) [−0.06 to 0.01]	.14	.12	.66
Ubiquinol-10, μmol/L	0.02 (0.05) [−0.08 to 0.13]	−0.32 (0.04) [−0.40 to −0.25]	.69	<.001	.16
LDL TRAP, μmol/L‡	−5.0 (2.1) [−9.1 to −0.9]	−17.3 (1.4) [−20.1 to −14.5]	.02	<.001	.35
LDL TRAP, μmol/mmol§	2.2 (0.7) [0.8 to 3.5]	4.3 (0.5) [3.4 to 5.3]	.002	<.001	15
LDL diene conjugation, μmol/L‡	−3.1 (1.2) [−5.5 to −0.6]	−5.9 (0.6) [−7.2 to −4.7]	.01	<.001	.40
LDL diene conjugation, μmol/mmol§	0.1 (0.3) [−0.4 to 0.7]	1.2 (0.2) [0.8 to 1.6]	.64	<.001	.89

*CI indicates confidence interval; HOMA IR, homeostasis model assessment of insulin resistance; LDL, low-density lipoprotein; TRAP, total peroxyl radical trapping potential of LDL (antioxidant potential of isolated LDL particles); and LDL diene conjugation, oxidation products of fatty acids in LDL particles. See footnote to Table 1 for SI conversion equations.
†In analysis of covariance with baseline values as covariates.
‡Expressed per liter of original serum.
§Expressed per millimole of isolated LDL-C.

COMMENT

The separate effects of dietary treatment and simvastatin on plasma lipid and lipoprotein levels were consistent with published data.[1,14-16] An important finding was that their effects on levels of lipids, lipoproteins, glucose, insulin, α-tocopherol, ascorbic acid, β-carotene, homocysteine, ubiquinol-10, and erythrocyte folate and on the indicators of LDL oxidation (LDL TRAP and LDL diene conjugation) were independent and additive. For example, dietary treatment alone, simvastatin treatment alone, and the treatments combined lowered LDL-C levels by 11%, 30%, and 41%, respectively. The independent and additive effects of dietary treatment and simvastatin on lipoprotein levels agree with those in a previous article reporting 5%, 27%, and 32% decreases in LDL-C in patients treated with a National Cholesterol Education Program Step II diet alone, lovastatin (20 mg/d) alone, or a combination.[17] Unlike in our study, decreased cholesterol and saturated fatty acid intakes were accompanied by a decreased intake of monounsaturated fatty acids and a decreasing trend in the intake of polyunsaturated fatty acids. The authors concluded that the reduction in LDL-C was small, and its benefit was possibly offset by the observed reduction in HDL-C.

In this study, dietary treatment decreased average serum cholesterol concentration by 19 mg/dL (0.5 mmol/L). This effect resulted mainly from dietary replacement of saturated fat with monounsaturated and polyunsaturated fats. Our finding is supported by a meta-analysis in which replacement of 7% of energy from saturated fat with either monounsaturated or polyunsaturated fats decreased total cholesterol levels by roughly 25 mg/dL (0.65 mmol/L).[14] Dietary intake of cholesterol decreased by approximately 80 mg/d in the dietary treatment group, which would decrease serum cholesterol levels by only 2 mg/dL (0.05 mmol/L).[15] Also, increased fiber intake's contribution to reduced serum cholesterol concentration was probably limited. According to a recent meta-analysis, eating 20 g of oats daily (corresponding to 3.4 g of fiber and 0.7 g of soluble fiber) decreases total cholesterol concentration in serum by 1 mg/dL (0.03 mmol/L).[16] We observed an intake increase of 2.2 g of soluble fiber daily in the dietary treatment group, which would result in a decrease of approximately 3 mg/dL (0.09 mmol/L) in total cholesterol.

Another important finding was that simvastatin treatment decreased se-

Antioxidant Vitamins, Provitamins, and LDL Oxidation

Dietary treatment decreased serum α-tocopherol levels by 3.5%, total peroxyl radical trapping potential of serum LDL by 4.9%, and estimated actual level of oxidized LDL in circulation (LDL diene conjugation) by 8.3% (Table 3, Figure 2). Relative antioxidant power of LDL preparations (LDL-TRAP/mmol of LDL-C) increased by 8.8%. Ascorbic acid, β-carotene, homocysteine, ubiquinol-10, and erythrocyte folate levels and the relative level of oxidized LDL (LDL diene conjugation/mmol of LDL-C) remained unchanged.

Simvastatin treatment decreased serum α-tocopherol levels by 16.2%, β-carotene levels by 19.5%, ubiquinol-10 levels by 22.0%, and total peroxyl radical trapping potential of serum LDL by 16.9% (Table 3, Figure 2). Because of decreased serum LDL-C levels, relative antioxidant power of LDL preparations (LDL TRAP/mmol of LDL-C) increased by 17.4%. The estimated actual level of oxidized LDL in circulation (LDL diene conjugation) decreased by 16.0%, but the relative level of oxidized LDL (LDL diene conjugation/mmol of LDL-C) increased by 13.1%. Simvastatin treatment did not change serum ascorbic acid, homocysteine, or erythrocyte folate levels.

There were no interactions between the effects of diet and simvastatin on levels of serum α-tocopherol, ascorbic acid, β-carotene, homocysteine, ubiquinol-10, and erythrocyte folate and on serum LDL fraction for diene conjugation and antioxidant potential (Table 3).

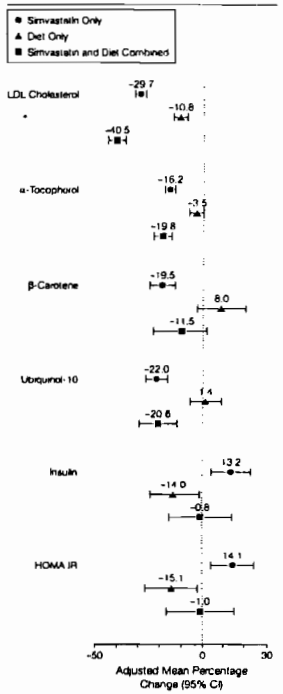


Figure 2. Separate and Combined Effects of Diet and Simvastatin on LDL Cholesterol, α-Tocopherol, β-Carotene, Ubiquinol-10, Insulin, and Insulin Resistance

Values are adjusted mean percentage changes with 95% confidence intervals (CIs). LDL indicates low-density lipoprotein; HOMA IR, homeostasis model assessment of insulin resistance.

rum concentrations of some antioxidant vitamins and provitamins. The concentrations of α-tocopherol, β-carotene, and ubiquinol-10 were lowered by 16% to 22%. Despite the increased dietary intake of α-tocopherol, cholesterol-lowering dietary treatment was associated with small decreases in serum α-tocopherol levels. Dietary treatment had no effects on serum β-carotene and ubiquinol-10 levels.

The decreased serum ubiquinol-10 concentration during simvastatin treatment agrees with findings of previous studies.[18-20] Ubiquinone is a by-product of cholesterol synthesis, and its decrease during simvastatin treatment may explain why the drug reduced serum ubiquinol levels, whereas dietary treatment did not.

In our study, LDL-C concentration decreased by 30% and HDL-C concentration increased by 7% during simvastatin treatment. Circulating α-tocopherol is bound to lipoproteins. In men, approximately 30% of α-tocopherol is bound to HDL-C; 60%, to LDL-C.[21] Thus, the observed changes in serum lipid concentrations are expected to result in a 16% decrease in serum α-tocopherol concentration, which also was the case.

Whether reduction in circulating concentrations of ubiquinone, α-tocopherol, and β-carotene would decrease their concentrations in human tissues is largely unknown. According to an uncontrolled study,[20] simvastatin (20 mg/d for 6 months) did not change ubiquinone levels in human skeletal muscle. Whether changes in serum α-tocopherol, β-carotene, and ubiquinone levels have any impact on platelet function, cell proliferation, immune responses, mitochondrial function, antioxidative processes other than LDL oxidation, and clinical outcomes has to be clarified in further studies.

In our study, reductions in serum LDL antioxidant potential during dietary and simvastatin treatments are in line with changes in serum concentrations of fat-soluble antioxidant vitamins and provitamins. However, the relative antioxidant potential of LDL increased during simvastatin and dietary treatment, mainly because of decreases in LDL concentrations.

The oxidized form of LDL may play a key role in atherogenesis. Most studies have regarded the susceptibility of isolated LDL to oxidation ex vivo as an indicator of LDL oxidation in vivo. We measured real end products of lipid peroxidation (formed diene conjugates of isolated LDL) in vivo to estimate oxidation of circulating LDL. In our study, both dietary and simvastatin treatments decreased serum concentrations of LDL diene conjugates. However, the formation of LDL diene conjugates relative to LDL-C increased during simvastatin treatment but remained unchanged during dietary treatment, suggesting qualitative deterioration of LDL by simvastatin but not by dietary treatment. Recently, an uncontrolled study[22] reported that simvastatin (20 mg/d) did not change LDL diene formation ex vivo when expressed per mole of LDL. Simvastatin increases the proportion of protein and decreases proportions of free cholesterol and cholesterol esters in LDL, which may result in a change not only in the amount but also in the composition of LDL.[23] Thus, differences in measurement techniques and expression of diene conjugation may explain the apparent differences in our data and those of a recently published study.[22] Simvastatin has been reported to possess antioxidant potential in vitro.[23] Our study does not support that this property would have any significant impact on circulating LDL, possibly because of decreased concentrations of circulating fat-soluble antioxidant vitamins and provitamins and possibly because of preferred hepatic uptake of native (nonoxidized) LDL compared with oxidized LDL.

Dietary intervention with reduced saturated fatty acid intake and increased monounsaturated and polyunsaturated fatty acid intake decreased, while simvastatin treatment increased, fasting serum insulin levels. The effects of the diet agree with previous data from cross-sectional and experimental studies. Increased fasting serum insulin levels and decreased insulin sensitivity have been associated with decreased concentrations of long-chain polyunsaturated fatty acids within muscle-membrane phospholipids[24] and with a decreased ratio of omega-6 polyunsaturated fatty acids to saturated fatty acids in serum phospholipids.[25] A diet low in saturated fat and rich in monounsaturated and polyunsaturated fats improves glucose tolerance in healthy women.[26] Fatty acid composition of cell membranes, reflecting fatty acid intake and metabolism, may modulate insulin binding and glucose transport.[27] Polyunsaturated fatty acids may also influence the action of insulin by acting as precursors for the generation of second messengers such as eicosanoids and diacylglycerols.[28]

Only a few randomized controlled studies, all in patients with type 2 diabetes mellitus, have examined the effects of simvastatin on fasting serum insulin levels or insulin sensitivity.[29-32] The results have been contradictory. Ohrvall and colleagues[29] found that simvastatin (10 mg/d for 4 months) increased fasting insulin concentrations by 21% and decreased insulin sensitivity by 28% but did not affect fasting triglyceride concentrations. In 2 small placebo-controlled studies, simvastatin produced nonsignificant changes in various determinants of insulin sensitivity.[30,31] In the most recent study, with 61 patients randomized to simvastatin and 67 to placebo, simvastatin decreased insulin resistance by 9%.[32] Changes in insulin levels were not shown. As in our study, a decrease in serum triglyceride level was a significant determinant for an increase in insulin sensitivity. In our study, simvastatin (20 mg/d for 12 weeks) increased fasting serum insulin levels of 120 nondiabetic hypercholesterolemic men by 13% and insulin resistance by 14%, despite concomitant favorable effects on serum triglyceride concentrations. Although we did not measure insulin sensitivity directly, the modest increase in fasting insulin levels together with completely unchanged glucose concentrations may indicate a slight decrease in insulin sensitivity after simvastatin treatment. On the other hand, the increase in serum insulin levels was fully counteracted by concomitant dietary treatment. Further studies are needed to explore the mechanisms by which simvastatin may result in increased serum insulin levels.

There were some important limitations in our study. We performed the dietary treatment single-blind, since it is impossible to accomplish this kind of dietary intervention in a double-blind fashion. To avoid the confounding influence of menstrual cycle, oral contraceptives, and estrogen replacement therapy, only men were included in this study, so the findings may be valid only for hypercholesterolemic men. Only a 12-week study period was long enough to show the effectiveness of dietary and statin treatments on the measured biochemical variables, but feasibility of the treatments should be evaluated in long-term studies. Further studies are needed to evaluate other potential cardioprotective effects of separate and combined dietary and simvastatin treatments not addressed in this study: effects on platelet aggregation, hemostasis, fibrinolysis, and endothelial function and the propensity for arrhythmia.

Both simvastatin treatment and a diet low in saturated fat intake but enriched in monounsaturated and polyunsaturated fatty acids, including α-linolenic acid, cereals, fruits, berries, and vegetables, significantly affected levels of serum lipids, insulin, and antioxidants. Both simvastatin and the diet reduced total cholesterol and LDL-C concentrations, with the effect of simvastatin being 3-fold that of dietary treatment. Simvastatin decreased the concentrations of 3 important antioxidant vitamins or provitamins, α-tocopherol, β-carotene, and ubiquinol-10, in serum by 16% to 22%, whereas dietary treatment decreased α-tocopherol concentration only slightly. Dietary treatment and simvastatin had opposite effects on fasting serum insulin levels, which were increased by simvastatin.

To conclude, the combination of a modified Mediterranean-type diet and statin treatment of hypercholesterolemia in nondiabetic men not only results in a beneficial additive effect on lowering serum total cholesterol and LDL-C concentration but also counteracts the elevating influence on fasting insulin level associated with simvastatin treatment. The combination is clinically sound, and the importance of diet as an integral part of statin treatment of hypercholesterolemic patients should be emphasized.

Author Contributions: *Study concept and design:* Jula, Marniemi, Huupponen, Rönnemaa.
Acquisition of data: Jula, Rastas.
Analysis and interpretation of data: Jula, Marniemi, Huupponen, Virtanen, Rönnemaa.
Drafting of the manuscript: Jula.
Critical revision of the manuscript for important intellectual content: Jula, Marniemi, Huupponen, Virtanen, Rastas, Rönnemaa.
Statistical expertise: Virtanen.
Study supervision: Jula, Marniemi, Huupponen, Rönnemaa.
Dietary counseling: Rastas.

REFERENCES

1. Scandinavian Simvastatin Survival Study Group. Randomized trial of cholesterol lowering in 4444 patients with coronary heart disease: the Scandinavian Simvastatin Survival Study (4S). *Lancet*. 1994;344:1383-1389.
2. Shepherd J, Cobbe SM, Ford I, et al. Prevention of coronary heart disease with pravastatin in men with hypercholesterolemia. *N Engl J Med*. 1995;333:1301-1307.
3. Burr ML, Fehily AM, Gilbert JF, et al. Effects of changes in fat, fish, and fibre intakes on death and myocardial reinfarction: Diet and Reinfarction Trial (DART). *Lancet*. 1989;2:757-761.
4. De Lorgeril M, Renaud S, Mamelle N, et al. Mediterranean alpha-linoleic acid-rich diet in secondary prevention of coronary heart disease. *Lancet*. 1994;343:1454-1459.
5. Singh RB, Rastogi SS, Verma R, et al. Randomised controlled trial of cardioprotective diet in patients with recent acute myocardial infarction: results of one year follow up. *BMJ*. 1992;304:1015-1019.
6. Hakala P, Marniemi J, Knuts LR, et al. Calculated vs analysed nutrient composition of weight reduction diets. *Food Chem*. 1996;57:71-75.
7. Beaumont JL, Carlson LA, Cooper GR, et al. Classification of hyperlipidemias and hyperlipoproteinaemias. *Bull World Health Organ*. 1970;43:891-915.
8. Haffner SM, Miettinen H, Stern MP. The homeostasis model in the San Antonio Heart Study. *Diabetes Care*. 1997;20:1087-1092.
9. Milne DB, Botnen J. Retinol, alpha-tocopherol, lycopene, and alpha- and beta-carotene simultaneously determined in plasma by isocratic liquid chromatography. *Clin Chem*. 1986;32:874-876.
10. Day BR, Williams DR, Marsh CA. A rapid manual method for routine assay of ascorbic acid in serum and plasma. *Clin Biochem*. 1979;12:22-26.
11. Takada M, Ikenoya S, Yuzuriha T, et al. Simultaneous determination of reduced and oxidized ubiquinones. *Methods Enzymol*. 1985;105:147-155.
12. Ahotupa M, Ruutu M, Mantyla E. Simple methods of quantifying oxidation products and antioxidant potential of low density lipoproteins. *Clin Biochem*. 1996;29:139-144.
13. Wieland H, Seidel D. A simple specific method for precipitation of low density lipoproteins. *Clin Biochem J Lipid Res*. 1983;24:904-909.
14. Gardner CD, Kraemer HC. Monounsaturated versus polyunsaturated dietary fat and serum lipids: a meta-analysis. *Arterioscler Thromb Vasc Biol*. 1995;15:1917-1927.
15. Howell WH, McNamara DJ, Tosca MA, et al. Plasma lipid and lipoprotein responses to dietary fat and cholesterol: a meta-analysis. *Am J Clin Nutr*. 1997;65:1747-1764.
16. Brown L, Rosner B, Willett WW, Sacks FM. Cholesterol-lowering effects of dietary fiber: a meta-analysis. *Am J Clin Nutr*. 1999;69:30-42.
17. Hunninghake DB, Stein EA, Dujovne CA, et al. The efficacy of intensive dietary therapy alone or combined with lovastatin in outpatients with hypercholesterolemia. *N Engl J Med*. 1993;328:1213-1219.
18. Human JA, Ubbink JB, Jerling JJ, et al. The effect of simvastatin on the plasma antioxidant concentrations in patients with hypercholesterolaemia. *Clin Chim Acta*. 1997;263:67-77.
19. Watts GF, Castelluccio C, Rice-Evans C, et al. Plasma coenzyme Q (ubiquinone) concentrations in patients treated with simvastatin. *J Clin Pathol*. 1993;46:1055-1057.
20. Laaksonen R, Jokelainen K, Laakso J, et al. The effect of simvastatin treatment on natural antioxidants in low-density lipoproteins and high-energy phosphates and ubiquinone in skeletal muscle. *Am J Cardiol*. 1996;77:851-854.
21. Behrens WA, Thompson JN, Madère R. Distribution of α-tocopherol in human plasma lipoproteins. *Am J Clin Nutr*. 1982;35:691-696.
22. Girona J, La Ville AE, Solà R, et al. Simvastatin decreases aldehyde production derived from lipoprotein oxidation. *Am J Cardiol*. 1999;83:846-851.
23. Kleinveld HA, Demacker PNM, de Haan AFJ, et al. Decreased in vitro oxidizability of low density lipoprotein in hypercholesterolaemic patients treated with 3-hydroxy-3-methylglutaryl-CoA reductase inhibitors. *Eur J Clin Invest*. 1993;23:289-295.
24. Borkman M, Storlien LH, Pan DA, et al. The relation between insulin sensitivity and the fatty-acid composition of skeletal-muscle phospholipids. *N Engl J Med*. 1993;328:238-244.
25. Pelikanova T, Kohout M, Valek J, et al. Insulin secretion and insulin action related to the serum phospholipid fatty acid pattern in healthy men. *Metabolism*. 1989;38:188-192.
26. Uusitupa M, Schwab U, Makimattila S, et al. Effects of two high-fat diets with different fatty acid compositions on glucose and lipid metabolism in healthy young women. *Am J Clin Nutr*. 1994;59:1310-1316.
27. Ginsberg BH, Brown TJ, Simon I, Spector AA. Effect of the membrane lipid environment on the properties of insulin receptor. *Diabetes*. 1981;30:773-780.
28. Stralfors P. Insulin stimulation of glucose uptake can be mediated by diacylglycerol in adipocytes. *Nature*. 1988;335:554-556.
29. Ohrvall M, Lithell H, Johansson J, et al. A comparison between the effects of gemfibrozil and simvastatin on insulin sensitivity in patients with non-insulin dependent diabetes mellitus and hyperlipoproteinemia. *Metabolism*. 1995;44:212-217.
30. Stewart MW, Dyer RG, Alberti KG, et al. The effects of lipid lowering drugs on metabolic control and lipoprotein composition in type 2 diabetic patients with mild hyperlipidemia. *Diabet Med*. 1995;12:250-257.
31. Hwu CM, Kwok CF, Chen HS, et al. Lack of effects of simvastatin on insulin sensitivity in type 2 diabetic patients with hypercholesterolemia: results from a double-blind, randomized, placebo-controlled crossover study. *Diabet Med*. 1999;16:749-754.
32. Paolisso G, Barbagallo M, Petrella G, et al. Effects of simvastatin and atorvastatin administration on insulin resistance and respiratory quotient in aged dyslipidemic non-insulin dependent diabetic patients. *Atherosclerosis*. 2000;150:121-127.

ENTRAINMENT OF FREE-RUNNING CIRCADIAN RHYTHMS BY MELATONIN IN BLIND PEOPLE

ROBERT L. SACK, M.D., RICHARD W. BRANDES, B.S., ADAM R. KENDALL, B.S., AND ALFRED J. LEWY, M.D., PH.D.

ABSTRACT

Background Most totally blind people have circadian rhythms that are "free-running" (i.e., that are not synchronized to environmental tima cues and that oscillate on a cycle slightly longer than 24 hours). This condition causes recurrent insomnia and daytime sleepiness when the rhythms drift out of phase with the normal 24-hour cycle. We investigated whether a daily dose of melatonin could entrain their circadian rhythms to a normal 24-hour cycle.

Methods We performed a crossover study involving seven totally blind subjects who had free-running circadian rhythms. The subjects were given 10 mg of melatonin or placebo daily, one hour before their preferred bedtime, for three to nine weeks. They were then given the other treatment. The timing of the production of endogenous melatonin was measured as a marker of the circadian time (phase), and sleep was monitored by polysomnography.

Results At base line, the subjects had free-running circadian rhythms with distinct and predictable cycles averaging 24.5 hours (range, 24.2 to 24.9). These rhythms were unaffected by the administration of placebo. In six of the seven subjects the rhythm was entrained to a 24.0-hour cycle during melatonin treatment ($P<0.001$). After entrainment, the subjects spent less time awake after the initial onset of sleep ($P=0.05$) and the efficiency of sleep was higher ($P=0.06$). Three subjects subsequently participated in a trial in which a 10-mg dose of melatonin was given daily until entrainment was achieved. The dose was than reduced to 0.5 mg per day over a period of three months; the entrainment persisted, even at the lowest dose.

Conclusions Administration of melatonin can entrain circadian rhythms in most blind people who have free-running rhythms. (N Engl J Med 2000;343:1070-7.)

THE endogenous circadian pacemaker oscillates with a period that is slightly longer than 24 hours and that therefore requires synchronization, or entrainment, to the 24-hour day. Entrainment involves regular adjustments of the circadian pacemaker, known as phase shifts, that depend on exposure to environmental time cues, particularly the daily light–dark cycle.[1] Light cues necessary for entrainment are conveyed from the retina to the suprachiasmatic nucleus (the locus of the mammalian circadian pacemaker) by way of the retinohypothalamic tract, a neural pathway that is separate from the visual and oculomotor pathways.[2] In totally blind people, light cues are unavailable, and disturbances of circadian rhythms are common.[3,9]

Among these disturbances are "free-running" rhythms, which reflect the intrinsic oscillation of the circadian pacemaker when it is not influenced by environmental time cues. Free-running rhythms are characterized by a consistent delay in the timing of the circadian cycle by as much as 60 to 70 minutes per day and can be detected by measurement at regular intervals of a marker rhythm, such as the daily rise in the plasma melatonin concentration. In blind people who have free-running rhythms, periodic symptoms of insomnia and daytime sleepiness commonly occur when the circadian pacemaker and, therefore, the circadian rhythm of sleepiness drift out of phase with the desired time for sleeping.[8] These symptoms vary considerably but can be among the most burdensome aspects of blindness. We evaluated the daily administration of melatonin as a method of entraining the circadian rhythms of totally blind people with free-running rhythms.

METHODS

Study Design

We studied seven subjects who were totally blind, as determined by ophthalmologic examination. They had free-running circadian rhythms, indicated by a predictable shift in the time of the cyclic rise in the plasma melatonin concentration, measured on three occasions about two weeks apart. At the time of a screening assessment, the subjects were in good general health and were not taking any medications that might affect plasma melatonin concentrations or sleep. Information about the study was provided to the subjects in print, in Braille, and on an audiotape; all the subjects gave written informed consent. The institutional review board of the Oregon Health Sciences University approved the protocol and the consent forms.

This study had a crossover design, balanced according to the order of treatment (melatonin first or placebo first). The subjects took 10 mg of oral melatonin or placebo nightly, approximately one hour before their preferred bedtime. We selected the 10-mg dose of melatonin because we were not able to document unequivocally the occurrence of entrainment in a previous, three-week trial of 5 mg.[10] The timing of the circadian cycle was assessed near the beginning, middle, and end of each trial by determining the time of day at which the endogenous melatonin concentration rose above 10 pg per milliliter (43 pmol per liter). This event has been found to be a reliable marker of the phase of the endogenous circadian cycle.[11]

The optimal timing of melatonin administration was determined with use of the melatonin phase-response curve, which describes the relation between the time in the circadian cycle that melatonin is given and its effects on the circadian rhythm. Treatment was initiated when the subject's free-running rhythm was approaching a normal phase (defined as a cycle in which the rise in the plasma melatonin concentration to 10 pg per milliliter occurred at about 9 p.m.[12]) and was continued until this rise was projected to occur at about 9 a.m. — that is, about 12 hours out of phase (assuming that treatment had no effect on the phase). Consequently, the planned duration of treatment varied among the subjects from three to nine weeks according to each subject's free-running circadian period. The subjects were asked to maintain consistent sleep schedules according to their preferred times for sleeping. Only the principal investigator and the project manager, and not the subjects, nurses, or research assistants, were aware of the treatment being given.

Approximately three months after the initial treatment, three of the subjects were treated a second time with melatonin at a dose of 10 mg per day. After their circadian rhythms were again entrained, the dose was gradually reduced to 0.5 mg per day, with the aim of determining the minimal effective dose of melatonin. In two of these three subjects, after treatment was discontinued, the time at which the plasma melatonin concentration rose above 10 pg per milliliter was determined every day or every other day for one week to explore the possibility that melatonin treatment might produce effects on the circadian pacemaker that persist after discontinuation of treatment (aftereffects).[13]

Analytic-grade melatonin (administered under Investigational New Drug application 26,318) was obtained from Regis Chemical (Morton Grove, Ill.) and was formulated under a pharmacist's supervision in gelarin capsules with a lactose filler. The placebo capsules contained only lactose. The pill containers were coded and were labeled for the subjects in both print and Braille.

Outcome Measures

The timing of the increase in endogenous melatonin production was determined as a marker of the circadian phase. To measure plasma melatonin on a given day, the subjects were admitted to the General Clinical Research Center of Oregon Health Sciences University, and blood samples were obtained every hour for 24 hours. To ensure that there was no interference from exogenous melatonin, no study capsules were taken on the day of sampling or on the preceding one or two days. Plasma melatonin concentrations were measured in the core laboratory by radioimmunoassay with an antibody raised in the laboratory of Kennaway et al.[14] and with reagents supplied by American Laboratory Products (Windham, N.H.). The lower limit of sensitivity of this assay is 1.0 pg per milliliter (4.3 pmol per liter); the interassay coefficient of variation is 10.4 percent at a concentration of 10 pg per milliliter (43 pmol per liter). This assay has been validated by gas chromatography–mass spectrometry.[15] Each 24-hour set of plasma melatonin measurements generated the time of day at which the plasma melatonin concentration exceeded the threshold of 10 pg per milliliter. The circadian period was then determined at base line and for the melatonin and placebo trials by fitting a linear regression line to the times of this increase on successive days.

The effects of treatment on total time asleep, sleep latency (the interval between the beginning of the opportunity for sleep and the onset of sleep), sleep efficiency (the total time asleep divided by the time allowed as an opportunity for sleep), and time spent awake after the onset of sleep were assessed by polysomnography performed in a sleep laboratory. The subjects were allowed to sleep from 10 p.m. until 6 a.m. Seven polysomnograms were recorded for each subject; the first (obtained during the screening assessment) was used to rule out any primary sleep disorders and to accustom the subject to the sleep laboratory. Polysomnograms were also recorded at the beginning, middle, and end of the melatonin and placebo trials, within several days before or after each assessment of circadian phase. The time at which the plasma melatonin concentration was predicted to reach the threshold of 10 pg per milliliter on a specific night of polysomnographic recording was extrapolated from the linear regression line fitted to the measured time points.

Statistical Analysis

Unless otherwise stated, data are expressed as means ±SD. Differences in circadian period were tested for statistical significance by two-sided t-tests for repeated measures. Differences in polysomnographic variables according to treatment (melatonin or placebo), the stage of the trial (beginning, middle, or end), and interaction (a differential effect of treatment depending on the stage of the trial) were tested for statistical significance by analysis of variance and by post hoc two-sided t-tests for repeated measures.

RESULTS

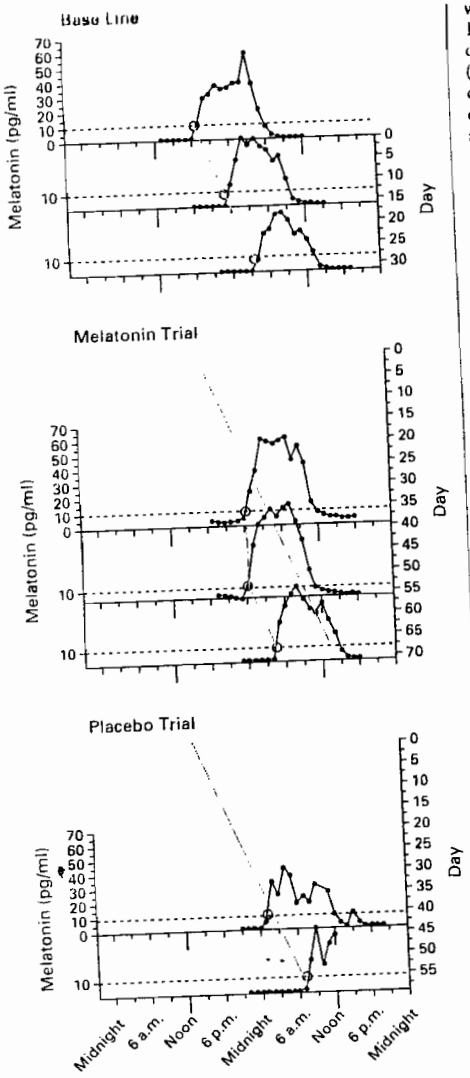
The characteristics of the seven subjects are shown in Table 1. The plasma melatonin concentrations in a representative subject (Subject 2) are shown in Figure 1. A regression line fitted to the three time points at

TABLE 1. CHARACTERISTICS OF SEVEN BLIND SUBJECTS WITH FREE-RUNNING CIRCADIAN RHYTHMS AT BASE LINE AND DURING THE ADMINISTRATION OF PLACEBO OR MELATONIN.

SUBJECT NO.	SEX	AGE	CAUSE OF BLINDNESS	AGE AT ONSET OF TOTAL BLINDNESS	STATUS OF EYES	CIRCADIAN PERIOD		
						BASE LINE	PLACEBO	MELATONIN
		yr		yr			hr	
1	F	42	Congenital glaucoma	6	Bilateral enucleation	24.2	24.2	24.0
2	M	46	Trauma	31	Eyes present	24.3	24.3	24.0
3	M	47	Congenital glaucoma	36	Unilateral enucleation	24.4	24.3	24.0
4*	F	50	Retinopathy of prematurity	Birth	Eyes present	24.5	—	24.0
5	F	45	Retinopathy of prematurity	Birth	Bilateral enucleation	24.4	24.3	24.0
6	M	57	Trauma	26	Bilateral enucleation	24.6	24.5	24.0
7	M	44	Trauma	12	Partial bilateral enucleation (prostheses)	24.9	24.8	24.3
Mean ±SD		48±5				24.5±0.2	24.4±0.2	24.0±0.1

*Complete data for the placebo trial were not obtained for Subject 4, and hence this subject was excluded from the statistical analysis.

From the Sleep and Mood Disorders Laboratory, Department of Psychiatry, School of Medicine, Oregon Health Sciences University, Portland. Address reprint requests to Dr. Sack at Mail Code L-469, Oregon Health Sciences University, Portland, OR 97201, or at sackr@ohsu.edu.

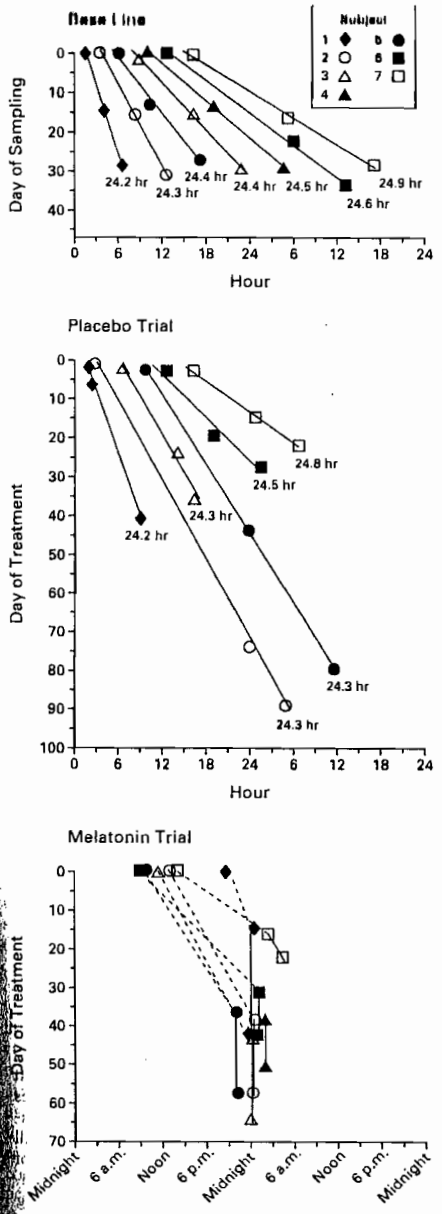


which the plasma melatonin concentration rose above 10 pg per milliliter before treatment (at base line) indicates a free-running circadian period of 24.3 hours (i.e., a delay in the rise in the plasma melatonin concentration to >10 pg per milliliter by 0.3 hour per day). The rhythm was regular, as indicated by a standard error in the slope of the regression line of 0.005 hour; consequently, the rise in this subject's plasma melatonin concentration could be accurately projected (for example, the standard error of prediction three weeks after the last measurement of plasma melatonin was 14 minutes). During the trial of melatonin, Subject 2 took melatonin daily at 11 p.m., except on days 38 and 57, when his endogenous melatonin profiles were assessed. On day 38, the plasma melatonin concentration rose above 10 pg per milliliter 3.6 hours earlier than the time predicted for a free-running rhythm with a period of 24.3 hours, and on day 57 it was 9.4 hours earlier (Fig. 1). Furthermore, the time of the rise in plasma melatonin was similar on days 38 and 57, indicating that the rhythm was effectively entrained to a 24-hour cycle. During the placebo trial, the circadian period was 24.3 hours, indistinguishable from that at base line.

The times of day at which the plasma melatonin concentration rose above 10 pg per milliliter in each of the seven subjects are shown in Figure 2. Complete data for the placebo trial were not obtained for Subject 4, and she was consequently excluded from the statistical analysis. The circadian periods at base line and during the placebo trial were highly correlated ($r=0.95$, $P=0.003$ for the correlation). The mean circadian period during the trial of melatonin was significantly different from the mean circadian period during the placebo trial (24.0±0.1 hours [95 percent confidence interval, 23.9 to 24.1] vs. 24.4± 0.2 hours [95 percent confidence interval, 24.2 to 24.6]; $P<0.001$), but it was not significantly different from 24.0 hours ($P=0.12$).

Figure 1. Plasma Melatonin Profiles in Subject 2 at Base Line and during the Placebo and Malatonin Trials.

The times of day at which the plasma melatonin concentration rose above a threshold of 10 pg per milliliter are shown as open circles on the dashed lines. The slopes of the regression lines drawn between these circles are an indication of the circadian rhythm: a slanting line indicates a free-running rhythm and a vertical line an entrained rhythm. The assessment at base line showed that this subject had a free-running rhythm (circadian period, 24.3 hours). During treatment with melatonin (10 mg taken daily at 11 p.m.), the rhythm was entrained (circadian period, 24.0 hours); it reverted to a free-running rhythm (circadian period, 24.3 hours) after melatonin treatment was stopped (on day 57). The rhythm was not affected by placebo taken daily at 11 p.m. (circadian period, 24.3 hours). All the regression lines have the same slope, indicating a circadian period of 24.3 hours, except for that during melatonin treatment, which indicated a circadian period of 24.0 hours. To convert the values for plasma melatonin to picomoles per liter, multiply by 4.3.



In Subject 7, there was a cumulative phase advance of 3 hours during the trial of melatonin, but the rhythm clearly failed to entrain to a 24-hour cycle. Subject 7 had the longest circadian period at base line (24.9 hours) of any of the subjects. Because the short duration of melatonin administration in this subject (18 days) may have accounted for the lack of entrainment, he was subsequently given melatonin on an open-label basis (10 mg daily at bedtime) for approximately 3 months. His plasma melatonin concentration rose above 10 pg per milliliter at 1:30 a.m. on day 45 of this treatment and at 7:30 a.m. on day 86, indicating lack of entrainment.

Polysomnographic data are shown in Table 2. The stage of each trial (beginning, middle, or end) had a significant effect on the efficiency of sleep and on the amount of time spent awake after the onset of sleep, but not on total time asleep or sleep latency. Because the subjects had free-running rhythms during placebo administration, the circadian phase (as reflected by the time of the increase in melatonin production) was progressively later as the trial proceeded. As expected, sleep efficiency was higher ($P=0.05$) and the amount of time spent awake after the onset of sleep was lower ($P=0.02$) at the beginning of the placebo trial (when subjects' rhythms were relatively close to normal phase) than at the end of the placebo trial (when subjects' rhythms were 12 hours out of phase). However, there was no effect on total time asleep or on sleep latency.

At the end of the melatonin trial the average time at which the plasma melatonin concentration rose above 10 pg per milliliter was at 12:18 a.m. (SD, 1.5 hours), close to that for a normal time but delayed as compared with 8:48 p.m. (SD, 1.3 hours), the average time in our study of people with normal sight.[12] The polysomnograms obtained at the end of each trial showed that less time was spent awake after the onset of sleep during the melatonin trial than

Figure 2. Circadian Rhythms in Seven Blind Subjects with Free-Running Circadian Rhythms at Base Line and during the Melatonin and Placebo Trials.

Each data point represents an assessment of circadian phese as determined by the time that endogenous plasma melatonin concentrations rose above the threshold of 10 pg per milliliter. The slopes of the fitted regression lines are indicative of the subjects' circadian period (shown in hours below the regression lines for the base line and placebo conditions). Treatment with melatonin or placebo was begun on day 1. In the top and middle panels, the regression lines are arranged on a relative time scale in ascending order so that they can be easily compared. In the bottom panel, the time scale is absolute and shows the assessments of circadian phase and fitted regression lines for all seven subjects before (dashed lines) and after (solid lines) the melatonin trial. Treatment with melatonin resulted in entrainment (a circadian period of 24.0 hours) in all but one subject (Subject 7); on average, the rise in plasma melatonin after entrainment occurred at 12:18 a.m.

TABLE 2. CIRCADIAN PHASE AND POLYSOMNOGRAPHIC RESULTS IN SEVEN BLIND SUBJECTS AT EACH STAGE OF THE PLACEBO AND MELATONIN TRIALS.*

VARIABLE	PLACEBO				MELATONIN				P VALUE‡		
	BEGINNING	MIDDLE	END	P VALUE†	BEGINNING	MIDDLE	END	P VALUE§	TREATMENT	STAGE OF TRIAL	INTERACTION
Time of rise in plasma melatonin to >10 pg/ml											
Mean	9:18 p.m.	2:06 a.m.	5:48 a.m.		8:18 p.m.	11:48 p.m.	12:18 a.m.				
SD (hr)	1.7	2.4	3.4		1.5	0.8	1.5		0.12	0.12	0.70
Total time asleep (min)	361.0±68.1	370.8±83.0	309.4±91.6		399.8±58.3	404.7±54.1	382.6±60.0		0.16	0.32	0.48
Sleep latency (min)	7.7±6.9	22.2±29.6	13.7±11.0		4.6±3.4	7.3±6.4	10.5±6.6		0.06	0.04	0.66
Sleep efficiency (%)	76.2±15.2	76.3±15.7	62.8±16.7	0.05	87.1±9.3	84.7±10.8	79.5±12.5	0.06	0.06	0.003	0.21
Time spent awake after the onset of sleep (min)	63.0±32.1	87.4±72.5	165.9±71.8	0.02	55.7±43.6	61.9±46.1	88.4±61.2	0.05			

*Plus–minus values are means ±SD. Where P values are not shown, the analysis of variance did not indicate significance, and post hoc analyses were not performed.

†P values are for the comparison between the result at the beginning of the placebo trial and the result at the end of the placebo trial.

‡P values were calculated by analysis of variance and are for the comparisons according to treatment (melatonin or placebo), stage of the trial (beginning, middle, or end), and interaction (a differential effect of treatment depending on the stage of the trial).

§P values are for the comparison between the result at the end of the placebo trial and the result at the end of the melatonin trial.

during the placebo trial (P=0.05) and that sleep efficiency was greater with melatonin than with placebo (P=0.06).

During the trial of melatonin taken in gradually reduced doses (step-down protocol), the rise in the plasma melatonin concentration to more than 10 pg per milliliter in the three participating subjects occurred consistently at about midnight (12:47 a.m. [SD, 0.6 hour] in Subject 1, 12:23 a.m. [SD, 0.5 hour] in Subject 2, and 11:50 p.m. [SD, 0.4 hour] in Subject 5), even at the lowest dose (0.5 mg per day), indicating stable entrainment, for approximately 120 days (Fig. 3). In Subject 5, about a month passed after melatonin was discontinued before the circadian rhythm reverted to the base-line period of 24.4 hours. In contrast, in Subject 2, a free-running rhythm resumed and the period returned to base line within several days after melatonin was discontinued (Fig. 3).

DISCUSSION

Although melatonin can induce phase advances in circadian rhythms, questions have been raised regarding its potency.[15] Our results indicate that the phase-advancing effects of melatonin are of sufficient magnitude to entrain free-running circadian rhythms in most blind persons who have such rhythms. The average daily phase advance required for entrainment was equal to the circadian period minus 24.0. In all seven subjects in this study, melatonin induced an average daily phase advance of up to 0.6 hour, but this was insufficient to entrain the free-running rhythm in Subject 7, whose base-line circadian period was 24.9 hours. The time of day at which the plasma melatonin concentration rose above 10 pg per milliliter after entrainment with melatonin was somewhat later than that reported in sighted persons.[12] An abnormally late circadian phase may be corrected by giving melatonin at an earlier time of day, thereby achieving a more normal relation between the rise in the plasma melatonin concentration and the desired sleeping schedule.

Three subjects underwent a second, open-label trial of treatment with melatonin. After entrainment was achieved at a dose of 10 mg of melatonin per day, the dose was gradually reduced. Entrainment was maintained at progressively lower doses for a period of four months, suggesting that a long-term benefit is likely with continuing treatment. The lowest dose tested in this protocol, 0.5 mg daily, resulted in plasma melatonin concentrations that were close to the physiologic range and can therefore be presumed to be very safe. According to these preliminary findings, it appears that treatment with a high dose of melatonin (10 mg per day) may be used to "capture" (initially entrain) a free-running rhythm, but that the dose can be gradually reduced without loss of effect. In some subjects, especially those with a circadian period close to 24.0 hours, a lower dose may be effective as the initial treatment.

Although we studied only seven subjects, the phase-shifting effects of melatonin treatment on circadian rhythms were clear and were consistent with previous data on the resetting of the circadian rhythm in sighted people[16-18] and with findings in case studies in blind people.[6,7,10,19-24] Because of the variability inherent in polysomnographic data, more subjects will need to be studied to document fully the effects of melatonin treatment on sleep. However, previous studies have shown that blind subjects with free-running

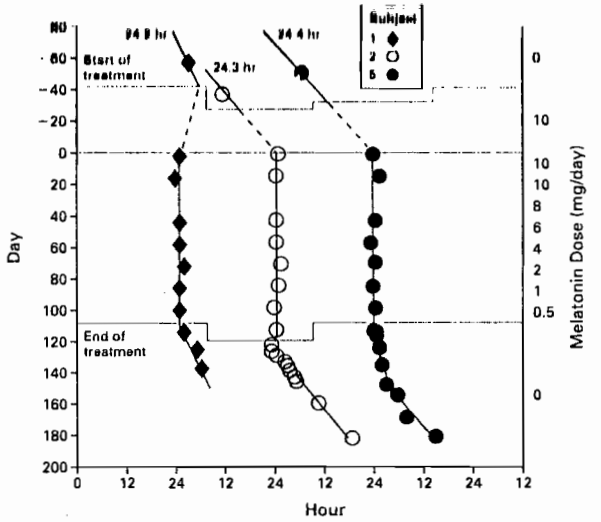


Figure 3. Assessments of Circadian Phase in Three Subjects Who Received Decreasing Doses of Melatonin in an Open-Label Trial.

Each data point represents an assessment of circadian phase as indicated by the time of day that endogenous plasma melatonin concentrations rose above the threshold of 10 pg per milliliter. For clarity, assessments of circadian phase are aligned with the dose-reduction schedule shown on the right. Treatment began (as indicated by the top horizontal line) at 42 days (Subject 1), 28 days (Subject 2), and 33 days (Subject 5) before the assessment of circadian phase on day 0. Free-running circadian periods before the beginning of treatment are shown (in hours) for each subject. After entrainment had been confirmed by a phase assessment on day 28, the dose was reduced every two weeks. The rise in plasma melatonin concentration during treatment occurred consistently at about midnight for 120 days, indicating entrainment, even at a dose as low as 0.5 mg per day. In Subject 1, there was no evidence of aftereffects; however, this subject had only three circadian-phase assessments, at relatively infrequent intervals, after the discontinuation of treatment. In Subject 2, the circadian rhythm became free running within days after treatment was discontinued. In Subject 5, about a month passed after the end of treatment before the circadian period returned to its base-line value, suggesting that entrainment for 120 days had a persistent effect on the circadian pacemaker. Circadian periods before the beginning of treatment are shown (in hours) for each subject.

rhythms sleep better when their circadian rhythms of sleepiness are more in phase with their desired times for sleeping.[9,25,26]

Why was melatonin effective in entraining the circadian rhythms of these subjects, whereas some previous attempts were not reliably successful?[10,27] Perhaps the longer duration of treatment used in the current study was important for successful entrainment. In addition, a 10-mg daily dose of melatonin may be more effective than a 5-mg daily dose, the dose used in our previous study.[10] On the other hand, in three of the subjects in our current study, much lower doses (which mimic endogenous plasma melatonin concentrations and induce a phase shift in sighted people[16-18]) maintained entrainment. Finally, there may be substantial individual variations in the response of the circadian rhythm to melatonin.

Determination of the optimal timing of melatonin administration requires the use of the melatonin phase–response curve, which describes the relation between the time in the circadian cycle that melatonin is given and its effects on the circadian rhythm. Administration of melatonin during the phase-advance portion of the curve (between 8 hours before and 4 hours after the increase in endogenous plasma melatonin production) results in shifts in the cycle to an earlier time of day; administration during the phase-delay portion of the curve (4 to 16 hours after the increase in endogenous plasma melatonin production) shifts the cycle to a later time of day.[16-18] With-

out knowledge of a person's circadian rhythm, it may be difficult to know what day to begin treatment so that the administration of melatonin at bedtime coincides with the phase-advance portion of the melatonin phase–response curve. This is the timing that was attempted in this study. Consequently, several weeks or even months of treatment may be required for the optimal phase relation to develop so that entrainment can be achieved and recurrent sleep problems can be resolved. We recommend that, if possible, treatment be initiated on a day when the plasma melatonin is predicted to rise above 10 pg per milliliter a few hours before the time the drug would be administered.

The benefits with respect to sleep in these subjects may be related not only to entrainment but also to a direct, sleep-promoting action of melatonin. Exogenous melatonin appears to have direct soporific effects, especially if it is ingested at a time when there is no endogenous production of melatonin (for example, during daytime hours in sighted people who are sleeping at conventional times).[28] We recently reported that a 10-mg dose of melatonin given once daily, at bedtime, to blind persons with free-running circadian rhythms improved sleep (without causing substantial phase shifts) when administered during a period when the circadian rhythm was "inverted" — that is, shifted 12 hours out of the normal phase.[29] Furthermore, we found no significant difference between the effects of melatonin and those of placebo when the subjects were tested on nights when their circadian rhythms were congruent with their preferred, conventional times for sleeping. We have speculated that melatonin does not generate sleep but that it can facilitate expression of the need to sleep that accumulates when one is awake.[30]

Melatonin has been widely used as a nutritional supplement in the United States for several years, with no reports of serious adverse effects. Nevertheless, long-term administration of 10 mg per day should be supervised by a physician. This dose can probably be gradually reduced without loss of efficacy.

There are approximately 1 million blind people in the United States, of whom about 20 percent are totally blind.[31] Extrapolating from our previous series,[9] at least half of this 20 percent (about 100,000 people) probably have free-running circadian rhythms, with a high proportion having circadian sleep–wake disorders. Melatonin may prove to be a safe and effective treatment for many of these people.

The phase-shifting effects of melatonin observed in this study of circadian rhythms in blind people may be relevant to the treatment of sighted people as well. People who fly across multiple time zones or who work nighttime or early-morning shifts routinely have symptoms of disordered sleep as a result of circadian disturbances. Similar pathophysiologic mechanisms have been proposed for advanced and delayed sleep-phase syndromes as well as for winter depression.[32] Administration should be timed according to the melatonin phase–response curve, since adverse effects may result if melatonin is given at times that would produce an antidromic (contrary) phase shift.[16,18]

In conclusion, free-running circadian rhythms in blind people can be entrained to a 24-hour cycle with a daily dose of melatonin, thereby preventing a burdensome sleep disorder.

Supported by grants from the Public Health Service (R01 MH 56874, to Dr. Sack; R01 MH55703 and R01 AG15140, to Dr. Lewy; and MO1 RR00334, to the General Clinical Research Center of Oregon Health Sciences University).

We are indebted to the nursing staff of the General Clinical Research Center; to Vance Bauer, Aaron Clemons, Neil Anderson, Victoria Chamberlin, and Lisa deJongh for technical assistance; to Gary Sexton, Ph.D., for statistical advice; and to Keith Parrott, Pharm.D., for the formulation of the melatonin capsules.

REFERENCES

1. Moore-Ede M, Sulzman FM, Fuller CA. The clocks that time us: physiology of the circadian timing system. Cambridge, Mass.: Harvard University Press, 1982.
2. Klein DC, Moore RY, Reppert SM. Suprachiasmatic nucleus: the mind's clock. New York: Oxford University Press, 1991.
3. Miles LEM, Raynal DM, Wilson MA. Blind man living in normal society has circadian rhythms of 24.9 hours. Science 1977;198:421-3.
4. Orth DN, Besser GM, King PH, Nicholson WE. Free-running circadian plasma cortisol rhythm in a blind human subject. Clin Endocrinol (Oxf) 1979;10:603-17.
5. Lewy AJ, Newsome DA. Different types of melatonin circadian secretory rhythms in some blind subjects. J Clin Endocrinol Metab 1983;56: 1103-7.
6. Sack RL, Lewy AJ, Hoban TM. Free-running melatonin rhythms in blind people: phase shifts with melatonin and triazolam administration. In: Rensing L, an der Heiden U, Mackey MC, eds. Temporal disorder in human oscillatory systems. Berlin, Germany: Springer-Verlag, 1987:219-24.
7. Sack RL, Stevenson J, Lewy AJ. Entrainment of a previously free-running blind human with melatonin administration. Sleep Res 1990;19:404.
8. Nakagawa H, Sack RL, Lewy AJ. Sleep propensity free-runs with the temperature, melatonin and cortisol rhythms in a totally blind person. Sleep 1992;15:330-6.
9. Sack RL, Lewy AJ, Blood ML, Keith LD, Nakagawa H. Circadian rhythm abnormalities in totally blind people: incidence and clinical significance. J Clin Endocrinol Metab 1992;75:127-34.
10. Sack RL, Lewy AJ, Blood ML, Stevenson J, Keith LD. Melatonin administration to blind people: phase advances and entrainment. J Biol Rhythms 1991;6:249-61.
11. Lewy AJ, Sack RL, Boney RS, et al. Assays for measuring the dim light melatonin onset (DLMO) in human plasma. Sleep Res 1997;26:733.
12. Lewy AJ, Bauer VK, Cutler NL, et al. Morning vs evening light treatment of patients with winter depression. Arch Gen Psychiatry 1998;55: 890-6.
13. Pittendrigh CS, Daan S. A functional analysis of circadian pacemakers in nocturnal rodents. I. The stability and lability of spontaneous frequency. J Comp Physiol 1976;106:223-52.
14. Kennaway DJ, Frith RG, Phillipou G, Matthews DC, Seamark RF. A specific radioimmunoassay for melatonin in biological tissue and fluids and its validation by gas chromatography-mass spectrometry. Endocrinology 1977;101:119-27.
15. Czeisler CA. Evidence for melatonin as a circadian phase-shifting agent. J Biol Rhythms 1997;12:618-23.
16. Lewy AJ, Ahmed S, Jackson JML, Sack RL. Melatonin shifts human circadian rhythms according to a phase-response curve. Chronobiol Int 1992;9:380-92.
17. Zaidan R, Geoffriau M, Brun J, et al. Melatonin is able to influence its secretion in humans: description of a phase-response curve. Neuroendocrinology 1994;60:105-12.
18. Lewy AJ, Bauer VK, Ahmed S, et al. The human phase response curve (PRC) to melatonin is about 12 hours out of phase with the PRC to light. Chronobiol Int 1998;15:71-83.
19. Palm L, Blennow G, Wetterberg L. Correction of non-24-hour sleep/wake cycle by melatonin in a blind retarded boy. Ann Neurol 1991;29:336-9.

20. Lewy AJ, Sack RL, Miller LS, Hoban TM. Antidepressant and circadian phase-shifting effects of light. Science 1987;235:352-4.
21. Lewy AJ, Bauer VK, Cutler NL, Sack RL, Ahmed S, Thomas KH, et al. Morning vs evening light treatment of winter depression: evidence that the therapeutic effects of light are mediated by circadian phase shifting. Chronobiol Int 1998;15:71-83.
22. Sack RL, Brandes RW, Lewy AJ. Totally blind people with free-running circadian rhythms can be normally entrained with melatonin. Sleep Res Online 1999;2:Suppl:S624. abstract. (See http://www.sro.org/cfiemplate/wfarscongress/indiv.cfm?ID=19998624.)
23. Skene DJ, Lockley SW, Arendt J. Melatonin entrains the free-running circadian rhythms of some blind subjects. Sleep Res Online 1999;2:Suppl: S725. abstract.
24. Lockley SW, Skene DJ, James K, Thapan K, Wright J, Arendt J. Melatonin administration can entrain the free-running circadian system of blind subjects. J Endocrinol 2000;164:R1-R6.
25. Klein T, Martens H, Dijk DJ, Kronauer RE, Seely EW, Czeisler CA. Circadian sleep regulation in the absence of light perception: chronic non-24-hour circadian rhythm sleep disorder in a blind man with a regular 24-hour sleep-wake schedule. Sleep 1993;16:333-43.
26. Lockley SW, Skene DJ, Arendt J. Changes in sleep in relation to circadian phase in the blind. In: Touitou Y, ed. Biologic clocks: mechanisms and applications. Amsterdam: Elsevier, 1998:447-53.
27. Pollard K, Arendt J, Akilonis M, Bennett T. Melatonin stabilizes sleep onset time in a blind man without entrainment of cortisol or temperature rhythms. Neurosci Lett 1990;113:193-8.
28. Sack RL, Lewy AJ. Melatonin as a chronobiotic: treatment of circadian desynchrony in night workers and the blind. J Biol Rhythms 1997;12:595-603.
29. Sack RL, Brandes RW, deJongh E, Pen SD, Nordstrom SL, Lewy AJ. The acute effects of melatonin administration on sleep in a blind man whose circadian rhythms were inverted. Sleep 1999;22:Suppl 1:S140. abstract. (See http://www.websciences.org/cfiemplate/APSS99/indiv99.cfm?ID=19999490.)
30. Sack RL, Hughes RJ, Edgar DM, Lewy AJ. Sleep-promoting effects of melatonin: at what dose, in whom, under what conditions, and by what mechanisms? Sleep 1997;20:908-15.
31. Vision research — a national plan: 1999-2003. Executive summary. Bethesda, Md.: National Institutes of Health, 1998. (NIH publication no. 98-4288.)
32. Lewy AJ, Bauer VK, Cutler NL, Sack RL. Melatonin treatment of winter depression: a pilot study. Psychiatry Res 1998;77:57-61.

Santorini

AGNES T. FRANK

BENEFICIAL EFFECTS OF HIGH DIETARY FIBER INTAKE IN PATIENTS WITH TYPE 2 DIABETES MELLITUS

MANISHA CHANDALIA, M.D., ABHIMANYU GARG, M.D., DIETER LUTJOHANN, PH.D., KLAUS VON BERGMANN, M.D., SCOTT M. GRUNOY, M.D., PH.D., AND LINDA J. BRINKLEY, R.D.

ABSTRACT

Background The effect of increasing the intake of dietary fiber on glycemic control in patients with type 2 diabetes mellitus is controversial.

Methods In a randomized, crossover study, we assigned 13 patients with type 2 diabetes mellitus to follow two diets, each for six weeks: a diet containing moderate amounts of fiber (total, 24 g; 8 g of soluble fiber and 16 g of insoluble fiber), as recommended by the American Diabetes Association (ADA), and a high-fiber diet (total, 50 g; 25 g of soluble fiber and 25 g of insoluble fiber) containing foods not fortified with fiber (unfortified foods). Both diets, prepared in a research kitchen, had the same macronutrient and energy content. We compared the effects of the two diets on glycemic control and plasma lipid concentrations.

Results Compliance with the diets was excellent. During the sixth week of the high-fiber diet, as compared with the sixth week of the ADA diet, mean daily preprandial plasma glucose concentrations were 13 mg per deciliter (0.7 mmol per liter) lower (95 percent confidence interval, 1 to 24 mg per deciliter [0.1 to 1.3 mmol per liter]; P=0.04) and mean daily urinary glucose excretion was 1.3 g lower (median difference, 0.23 g; 95 percent confidence interval, 0.03 to 1.83; P=0.008). The high-fiber diet also lowered the area under the curve for 24-hour plasma glucose and insulin concentrations, which were measured every two hours, by 10 percent (P=0.02) and 12 percent (P=0.05), respectively. The high-fiber diet reduced plasma total cholesterol concentrations by 6.7 percent (P=0.02), triglyceride concentrations by 10.2 percent (P=0.02), and very-low-density lipoprotein cholesterol concentrations by 12.5 percent (P=0.01).

Conclusions A high intake of dietary fiber, particularly of the soluble type, above the level recommended by the ADA, improves glycemic control, decreases hyperinsulinemia, and lowers plasma lipid concentrations in patients with type 2 diabetes. (N Engl J Med 2000;342:1392-8.)

DIETARY guidelines for patients with diabetes mellitus were revised by the American Diabetes Association (ADA) earlier this year.[1] The ADA recommends that the composition of the diet be individualized on the basis of a nutritional assessment and the outcomes desired. Consistent with the previous recommendations of the ADA,[2] the new guidelines advise replacing saturated fat with carbohydrates. However, on the basis of previous studies,[3-10] an alternative approach of replacing saturated fat with cis monounsaturated fat was also included in the recommendations.[1] This new approach is further supported by epidemiologic studies that have shown the healthful effects of diets rich in cis monounsaturated fat in Mediterranean countries.[11,12]

Another, less strongly emphasized aspect of Mediterranean diets is the high intake of fruits, vegetables, and grains that are rich sources of dietary fiber.[13,14] The ADA recommended a moderate increase in the intake of dietary fiber to 20 to 35 g per day because of the cholesterol-lowering effects of soluble fiber. However, the effects of dietary fiber on glycemic control were considered inconsequential.[1] Furthermore, the expert panel of the ADA considered it difficult to achieve a high dietary intake of soluble fiber without consuming foods or supplements fortified with fiber.[1] We therefore designed the present study to determine the effects on glycemic control and plasma lipid concentrations of increasing the intake of dietary fiber in patients with type 2 diabetes exclusively through the consumption of foods not fortified with fiber (unfortified foods) to a level beyond that recommended by the ADA. In addition, we studied the effects of such an intervention on the intestinal absorption of cholesterol and the fecal excretion of sterols in an attempt to uncover the mechanisms by which a high-fiber diet lowers plasma cholesterol.

From the Department of Internal Medicine (M.C., A.G., S.M.G., L.J.B.) and the Center for Human Nutrition (A.G., S.M.G.), University of Texas Southwestern Medical Center, Dallas; the Department of Veterans Affairs Medical Center, Dallas (M.C., A.G., S.M.G.); and the Department of Clinical Pharmacology, Rheinische Friedrich-Wilhelms-Universität, Bonn, Germany (D.L., K.B.). Address reprint requests to Dr. Garg at the Center for Human Nutrition, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd., Dallas, TX 75390.

METHODS

Patients

We studied 12 men and 1 woman (9 non-Hispanic whites and 4 blacks) with type 2 diabetes at the general clinical research center of the University of Texas Southwestern Medical Center at Dallas. The protocol for the study was approved by the institutional review board of the medical center, and each patient gave written informed consent. In all patients the onset of diabetes was insidious; the disease developed in most of the patients after 40 years of age. Their mean (±SD) age was 61±9 years (range, 45 to 70). Their mean body weight was 93.5±12.7 kg, and the mean body-mass index (the weight in kilograms divided by the square of the height in meters) was 32.3±3.9. Three patients were treated with diet alone, and the other 10 patients were treated with 2.5 to 20 mg of glyburide daily in addition to diet. The dose of glyburide was not changed during the study.

On entry into the study, the patients' plasma cholesterol and triglyceride concentrations after an overnight fast ranged from 181 to 324 mg per deciliter (4.90 to 8.38 mmol per liter) and 67 to 390 mg per deciliter (0.76 to 4.40 mmol per liter), respectively, and their fasting plasma glucose concentrations were less than 200 mg per deciliter (11.1 mmol per liter). Their glycosylated hemoglobin values ranged from 6.0 to 9.8 percent. Two patients had a history of coronary heart disease, but none had recently had a myocardial infarction, unstable angina pectoris, or congestive heart failure. Also, none had thyroid, renal, or hepatic disease. None of the patients were receiving lipid-lowering therapy.

Design of the Study

All the patients were first admitted to the general clinical research center for five days (the base-line period), during which a detailed history was taken, a physical examination was performed, and laboratory tests were performed. After the base-line period, all the patients received both the ADA diet and the high-fiber diet, each diet for a period of six weeks. Six patients received the high-fiber diet first, and the other seven received the ADA diet first. There was a median interval of seven days between the two study periods, during which the patients were instructed to consume an isocaloric diet. During the last week of each dietary period (days 36 to 42), the patients were hospitalized for evaluation.

On weekdays, all the patients ate at least one meal (breakfast, lunch, or dinner) at the general clinical research center. Other meals were supplied in packages so that they could be consumed at home. The dietitian monitored compliance by interviewing the patients. The patients were instructed to bring back any unconsumed food and to maintain a constant level of physical activity throughout the study.

Blood for lipid analyses was drawn, after an overnight fast, daily for two days before the institution of the study diet and daily on days 38 through 42 during both dietary periods. Plasma glucose was measured at 7 and 11 a.m. and at 4 and 8 p.m. each day during the base-line period and on days 38 through 42 of both dietary periods. Glycosylated hemoglobin was measured during the base-line period and at the end of each dietary period. On the last day of each dietary period, blood samples were obtained every two hours for measurements of plasma glucose and insulin. On days 38 through 42, patients collected 24-hour urine specimens for quantitative determination of glucose.

To permit us to determine fecal sterol balance and the percentage of cholesterol absorption, each patient took a capsule containing 30 mg of sitostanol, 3 mg of [26,26,26,27,27,27-$^2H_6$]-cholesterol, and 3 mg of [5,6,22,23-$^2H_4$]-sitostanol (Medical Isoropes, Pelham, N.H.) three times a day on days 36 through 42. Fecal samples were collected on day 35 or 36 and on the last three days of each dietary period. Fecal samples were frozen within 12 hours after collection and were pooled for analysis of small aliquots.

Diets

The composition of the study diets is shown in Table 1. The composition of the diets was calculated by means of a software program based on the Department of Agriculture Handbook Series 8 (Nutriplanner, Practocare, San Diego, Calif.).[15] The content of total as well as soluble and insoluble dietary fiber was estimated according to the data provided in the *CRC Handbook of Dietary Fiber in Human Nutrition*.[16] Both diets consisted of unfortified foods. The patients were allowed some choices of food items. Both diets provided 15 percent of the total energy as protein, 55 percent as carbohydrate, and 30 percent as fat; saturated, cis monounsaturated, and polyunsaturated fats accounted for 7 percent, 17 percent, and 6 percent of the total energy, respectively.

The high-fiber diet provided 50 g of total fiber per day; soluble and insoluble fiber content provided 25 g each. The ADA diet contained 24 g of total fiber per day, with 8 g as soluble fiber and 16 g as insoluble fiber. Unfortified foods, particularly those rich

TABLE 1. COMPOSITION OF THE STUDY DIETS.

CONSTITUENT	ADA DIET*	HIGH-FIBER DIET
Carbohydrate (% of total energy)	55	55
Protein (% of total energy)	15	15
Fat (% of total energy)	30	30
Saturated	7	7
Cis monounsaturated	17	17
Polyunsaturated	6	6
Cholesterol (mg/day)	300	297
Fiber (g/day)		
Total	24	50
Soluble	8	25
Insoluble	16	25

*ADA denotes American Diabetes Association.

TABLE 2. SAMPLE MENUS OF THE STUDY DIETS.*

ADA DIET		HIGH-FIBER DIET	
FOOD	WEIGHT	FOOD	WEIGHT
	grams		grams
Breakfast			
Orange juice	220	Orange sections	300
White grits	50	Oatmeal	50
Egg substitute	40	Scrambled egg	37
Olive oil	10	Olive oil	10
Decaffeinated coffee	2	Decaffeinated coffee	2
Lunch			
Ham (5% fat)	50	Ham (5% fat)	52
Mayonnaise	6	Mayonnaise	12
Iceberg lettuce	15	Iceberg lettuce	10
Fresh tomato	30	Fresh tomato	15
Low-sodium bread	60	Whole-wheat bread	60
Corn (canned)	140	Corn (canned)	40
Cider vinegar	5	Green peas (canned)	110
Dehydrated onion	2	Dehydrated onion	2
Olive oil	10	Olive oil	10
Fresh green pepper	10	Fresh green pepper	10
Fresh celery	15	Fresh celery	15
Fruit cocktail (canned)	105	Fresh papaya	250
Instant tea	2	Instant tea	2
Oatmeal raisin cookie	20		
Dinner			
Chicken breast (skinned)	90	Chicken breast (skinned)	90
Bran flakes	10	Bran flakes	10
Low-sodium bread	20	Oat bran	5
Parmesan cheese	1	Parmesan cheese	1
Whole egg	1	Egg substitute	10
Tomato (canned)	120	Tomato (canned)	105
Low-fat cheese	11	Low-fat cheese	19
Spaghetti	45	Spaghetti	19
Green beans	75	Zucchini	195
Olive oil	17	Olive oil	19
Whole-wheat bread	21	Whole-wheat bread	30
Graham crackers	21	Fresh peaches	300
Instant tea	2	Instant tea	2
Bedtime snack			
Mozzarella cheese	30	Fruit cocktail (canned)	200
Low-sodium bread	30	Cherries (canned)	100
Pineapple juice	190	Granola	15

*Each menu provided 2308 kcal per day. ADA denotes American Diabetes Association.

to edible fiber, such as cantaloupe, grapefruit, oranges, papaya, tomatoes, lima beans, okra, sweet potato, winter squash, zucchini, granola, oat bran, and oatmeal, were used to achieve high fiber intake. No fiber supplements were used. Sample menus of both the study diets are shown in Table 2. The individual foods were weighed daily during meal preparation in the research kitchen of the general clinical research center.

Biochemical Analyses

Fasting plasma total cholesterol, lipoprotein cholesterol, and triglycerides were measured according to the procedures of the Lipid Research Clinics.[17] Cholesterol and triglycerides were measured enzymatically with the use of kits (Boehringer Mannheim, Indianapolis). Very-low-density lipoproteins (VLDLs) (density, less than 1.006 g per milliliter) were removed by ultracentrifugation, and cholesterol was measured in the VLDL fraction and the infranatant. High-density-lipoprotein (HDL) cholesterol was measured enzymatically after lipoproteins containing apolipoprotein B had been precipitated with heparin–manganese. Cholesterol in the low-density lipoprotein (LDL) fraction was estimated to be the difference between the cholesterol content of the infranatant and that of the HDL fraction.

Plasma and urinary glucose were measured by the glucose oxidase method (Beckman Glucose Analyzer, Beckman Instruments, Fullerton, Calif.). Glycosylated hemoglobin was measured with ion-exchange high-performance liquid chromatography (Bio-Rad Laboratories, Hercules, Calif.). Plasma insulin was measured by radioimmunoassay.[18,19]

Pooled fecal samples collected within the last week of each dietary period were prepared for analysis of neutral and acidic fecal sterols as described previously.[20] Gas–liquid chromatography of neutral and acidic fecal sterols was performed on a gas chromatograph (model HP5890, Hewlett–Packard, Palo Alto, Calif.) equipped with an automatic sample injector. Cholesterol absorption was measured during the same period from fecal samples by gas–liquid chromatography and mass spectrometry.[21]

Statistical Analysis

To compare the two study periods and to assess the effect of the sequence in which the patients received the high-fiber and ADA diets, we used repeated measures analysis of variance.[22] No skewed data, we used the Wilcoxon signed-rank test to compare the two dietary periods.[23]

RESULTS

The compliance with both the study diets was excellent, according to interviews with the patients and estimates of the energy content of any leftover foods. Three patients reported consuming extra food on one day during the study, two while eating the high-fiber diet and one the ADA diet. The patients commented about the larger quantities of food in the high-fiber diet, but they consumed all the food given to them. The results are presented irrespective of the order of the diets, because the sequence of the diets had no effect on the results.

During the last week of each study period, the patients in both groups had similar daily energy intakes and body weights and received a similar dose of glyburide (Table 3). The mean plasma glucose concentration was lower (by 13 mg per deciliter [0.7 mmol per liter], or 8.9 percent) when patients completed the high-fiber diet than when they completed the ADA diet (P=0.04), and mean daily urinary glucose excretion was 1.3 g lower (P=0.008). Daily plasma glucose concentrations were 10 percent lower with the high-fiber diet than with the ADA diet (values for the area under the curve, 3743±944 vs. 3365±1003 mg·hour per deciliter [207.8±52.4 vs. 186.8±55.7 mmol·hour per liter]; P=0.02), and plasma insulin concentrations were 12 percent lower (values for the area under the curve, 1107±650 vs. 971±491 μU·hour per milliliter [6642±3900 vs. 5826±2946 pmol·hour per liter]; P=0.05) (Fig. 1). Glycosylated hemoglobin values were slightly lower after the high-fiber diet (P=0.09).

As compared with the ADA diet, the high-fiber diet resulted in a lower fasting plasma total cholesterol concentration (by 6.7 percent, P=0.02), a lower plasma triglyceride concentration (by 10.2 percent, P=0.02), and a lower plasma VLDL cholesterol concentration (by 12.5 percent, P=0.01) (Table 4). The fasting plasma LDL cholesterol concentration was 6.3 percent lower with the high-fiber diet (P=0.11). There were no significant differences between the two diets in terms of the fasting plasma HDL cholesterol concentration.

As compared with the ADA diet, the high-fiber diet decreased gastrointestinal absorption of cholesterol by 10 percent (48.5±9.6 vs. 43.7±7.4 percent; 95 percent confidence interval for the decrease, 0.6 to 9.0 percent; P=0.03) and increased fecal acidic sterol excretion by 41 percent (895±301 vs. 1258±458 mg per day; 95 percent confidence interval for the increase, 137 to 589 mg per day; P=0.005), but did not significantly affect the excretion of neutral

TABLE 3. METABOLIC VARIABLES DURING THE LAST WEEK OF THE STUDY PERIODS (DAYS 38 THROUGH 42).*

VARIABLE	ADA DIET	HIGH-FIBER DIET	DIFFERENCE BETWEEN DIETS (95% CI)	P VALUE†
Energy intake (kcal/day)	2308±236	2308±236	—	1.00
Weight (kg)	90.7±13.3	90.5±12.7	−0.2 (−1.1 to 0.6)	0.60
Dose of glyburide (mg/day)	10.0±8.7	10.0±8.7	—	1.00
Plasma glucose (mg/deciliter)‡	142±36	130±38	−13 (−24 to −1)	0.04
Urinary glucose (g/day)				
Mean	2.3±4.3	1.0±1.9	—	—
Median§	0.76	0.0	−0.23 (−1.83 to −0.03)	0.008
Glycosylated hemoglobin (%)	7.2±1.3	6.9±1.2	−0.3 (−0.6 to 0.1)	0.09

*Plus–minus values are means ±SD. ADA denotes American Diabetes Association, and CI confidence interval.

†An analysis of variance was used to compare the two diets, except for urinary glucose, for which the Wilcoxon signed-rank test was used.

‡The values are averages of plasma glucose concentrations measured at 7 and 11 a.m. and at 4 and 8 p.m. each day for five days during hospitalization. To convert values for plasma glucose to millimoles per liter, multiply by 0.056.

§The values are averages of five daily urine collections during hospitalization.

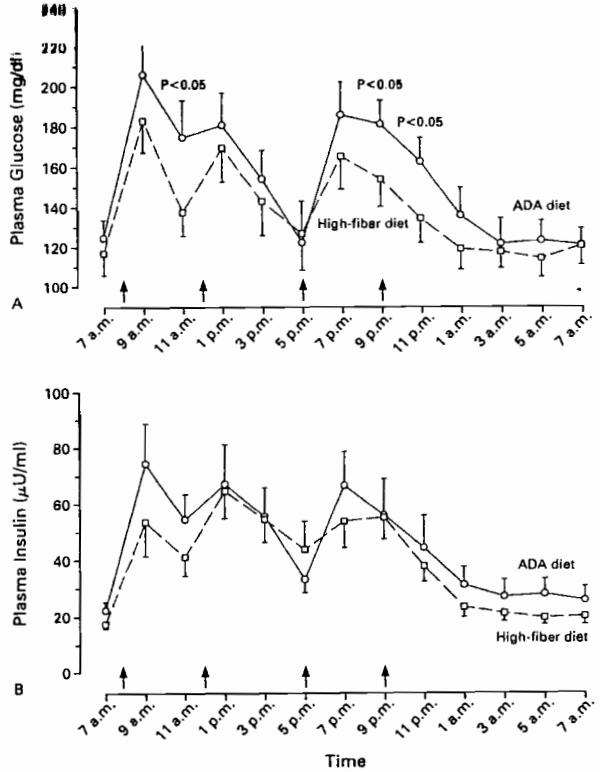


Figure 1. Mean (±SE) 24-Hour Profile of Plasma Glucose Concentrations (Panel A) and Insulin Concentrations (Panel B) during the Last Day of the American Diabetes Association (ADA) Diet and the Last Day of the High-Fiber Diet in 13 Patients with Type 2 Diabetes Mellitus.

The arrows indicate the times at which the main meals and a snack were consumed during the day. To convert values for glucose to millimoles per liter, multiply by 0.056. To convert values for insulin to picomoles per liter, multiply by 6.

TABLE 4. FASTING PLASMA LIPID AND LIPOPROTEIN CONCENTRATIONS DURING THE LAST WEEK OF THE STUDY PERIODS (DAYS 38 THROUGH 42).*

VARIABLE	ADA DIET	HIGH-FIBER DIET	DIFFERENCE BETWEEN DIETS (95% CI)	P VALUE†
	mg/dl			
Plasma total cholesterol	210±33	196±31	−14 (−27 to −2)	0.02
Plasma triglycerides	205±95	184±76	−21 (−37 to −4)	0.02
Plasma VLDL cholesterol	40±19	35±16	−5 (−9 to −1)	0.01
Plasma LDL cholesterol	142±29	133±29	−9 (−22 to 3)	0.11
Plasma HDL cholesterol	29±7	28±4	−1 (−4 to 3)	0.80

*Plus-minus values are means ±SD. To convert values for cholesterol and triglycerides to millimoles per liter, multiply by 0.026 and 0.011, respectively. ADA denotes American Diabetes Association, CI confidence interval, VLDL very-low density lipoprotein, LDL low density lipoprotein, and HDL high density lipoprotein.

†An analysis of variance was used to compare the two diets.

sterols (1052±375 vs. 1122±565 mg per day; 95 percent confidence interval for the difference, −194 to 334 mg per day; P=0.60).

DISCUSSION

The intake of dietary fiber among people living in Western countries remains low, and according to the Third National Health and Nutrition Examination Survey (NHANES), it averages 17 g per day in the United States.[24] Although patients with diabetes are advised to increase their intake of dietary fiber, in the NHANES study, their average daily intake was found to be only 16 g.[24] Why the intake of dietary fiber in patients with diabetes remains low — despite its well-documented effect of lowering plasma cholesterol concentrations — remains unexplained. It is possible that the controversy about whether there are beneficial effects of dietary fiber on glycemic control reduces the enthusiasm of physicians and dietitians for recommending high-fiber diets. The main purpose of our study was to investigate the effects on glycemic control of increasing the intake of dietary fiber. To avoid the confounding effects of concomitant changes in energy and macronutrients, the two study diets were isocaloric and the macronutrient composition of the diets was identical. Furthermore, unfortified foods were used as the source of dietary fiber.

Most important, we found that the high-fiber diet improved glycemic control, as evidenced by decreases in the mean daily preprandial and 24-hour plasma glucose concentrations. Urinary glucose excretion was also lowered by the high-fiber diet. The high-fiber diet lowered glycosylated hemoglobin values slightly but not significantly. The high-fiber diet also lowered 24-hour plasma insulin concentrations.

The results of previous studies that evaluated the role of dietary fiber on glycemic control in patients with type 2 diabetes were inconsistent. In some of the studies, the lack of control for concomitant changes in the intake of macronutrients makes the data difficult to interpret. For example, in the study by Kiehm et al.[25] and in that by Simpson et al.,[26] the high-fiber diet had a lower fat and higher carbohydrate content than the low-fiber diet. In other studies, the interpretation of the results was confounded by the short duration of the dietary intervention,[27-29] the lack of random assignment of the sequence of the high-fiber and low-fiber diets,[27,29] and unexplained weight loss during the high-fiber diet.[29]

Only a few well-controlled studies have evaluated the glycemic effects of increasing the intake of dietary fiber with the use of either preparations of refined concentrated fiber or unfortified food, and the results have been inconsistent.[1,30] For example, diets that included 15 to 21 g of guar-gum fiber or oat-bran concentrate per day had no effect on glycemic control[31,32] or resulted in only a slight improvement.[33,34] In randomized, crossover trials of six weeks' duration in which the intake of dietary fiber was increased by 16 g per 1000 kcal through the consumption of foods prepared in a research kitchen or by 14 g per day through dietary instruction, there was no improvement in glycemic control.[35,36] In contrast, increasing dietary fiber by 23 g for three weeks and by 30 g for six weeks resulted in decreased fasting and postprandial plasma glucose concentrations.[37,38] We found that an increase in the intake of total dietary fiber, which consisted predominantly of soluble fiber, significantly improved glycemic control and decreased the degree of hyperinsulinemia in patients with type 2 diabetes.

Our study also demonstrates the feasibility of achieving a high intake of dietary soluble fiber by consuming unfortified foods. Our patients accepted the high-fiber diet well and had few side effects; therefore, we recommend that patients with diabetes be encouraged to use unfortified foods instead of less palatable purified-fiber preparations and supplements to increase their intake of dietary fiber.

The mechanisms of the improved glycemic control associated with high fiber intake remain undefined. Whether this effect is due to an increase in soluble fiber, insoluble fiber, or both is unclear. Besides causing increased fecal excretion of bile acids, dietary fiber may cause malabsorption of fat.[39] However, in our study, the patients' weight did not change with the high-fiber diet, which suggests that the degree of reduction in the absorption of fat was insignificant. Another possibility is that dietary fiber improves glycemic control by reducing or delaying the absorption of carbohydrates.

As expected, the high-fiber diet reduced plasma total cholesterol concentrations by 6.7 percent, a finding consistent with the results of previous reports of the cholesterol-reducing effects of soluble but not insoluble fiber.[40,41] Therefore, the lowering of cholesterol can be attributed primarily to an average increase of 17 g in the intake of soluble fiber. Previous studies in normal subjects have reported no effects of the amount of dietary fiber on plasma triglyceride concentrations.[42] In our study, the decrease in plasma triglyceride and VLDL cholesterol concentrations during the high-fiber diet could have been due to the improvement in glycemic control.

The mechanisms of the reduction in plasma cholesterol concentrations induced by the increased dietary fiber intake are controversial, however. The increase in bile-acid excretion probably explains most of the reduction, and the reduction in cholesterol absorption may also have contributed to this finding. Previous studies have also reported a variable increase in bile-acid excretion resulting from the consumption of pectin,[39,43] oat bran,[44,45] bagasse,[46] and diets with a mixture of soluble fiber and insoluble fiber,[47] but not psyllium.[48] In contrast, Kesaniemi et al.[47] reported that a high-fiber diet did not change cholesterol absorption in normal subjects. However, the high-fiber diet they used included 26 g of fiber, and it did not lower plasma cholesterol concentrations.[47]

In conclusion, an increase in the intake of dietary fiber, predominantly of the soluble type, by patients with type 2 diabetes mellitus improved glycemic control and decreased hyperinsulinemia in addition to the expected lowering of plasma lipid concentrations. Therefore, dietary guidelines for patients with diabetes should emphasize an overall increase in dietary fiber through the consumption of unfortified foods, rather than the use of fiber supplements.

Supported in part by grants (M01-RR00633 and HL-29252) from the National Institutes of Health and by research grants from the Bundesministerium für Bildung, Forschung, Wissenschaft und Technologie (01EC9402) and the Deutsche Forschungsgemeinschaft (BE 1673/1-1).

We are indebted to Angela Osborn, Travis Petricek, and the nursing and dietetic service of the General Clinical Research Center of the University of Texas Southwestern Medical Center, Dallas, for their excellent technical support and to Beverley Adams-Huet, M.S., for statistical analysis.

REFERENCES

1. Nutrition recommendations and principles for people with diabetes mellitus. Diabetes Care 2000;23:S43-S46.
2. American Diabetes Association. Nutritional recommendations and principles for individuals with diabetes mellitus: 1986. Diabetes Care 1987;10: 126-32.
3. Garg A, Bonanome A, Grundy SM, Zhang Z-J, Unger RH. Comparison of a high-carbohydrate diet with high-monounsaturated-fat diet in patients with non-insulin-dependent diabetes mellitus. N Engl J Med 1988; 319:829-34.
4. Rivellese AA, Giacco R, Genovese S, et al. Effects of changing amount of carbohydrate in diet on plasma lipoproteins and apolipoproteins in type II diabetic patients. Diabetes Care 1990;13:446-8.
5. Parillo M, Rivellese AA, Ciardullo AV, et al. A high-monounsaturated-fat/low-carbohydrate diet improves peripheral insulin sensitivity in non-insulin-dependent diabetic patients. Metabolism 1992;41:1373-8.
6. Rasmussen OW, Thomsen C, Hansen KW, Vesterlund M, Winther E, Hermansen K. Effects on blood pressure, glucose and lipid levels of a high monounsaturated fat diet compared with a high carbohydrate diet in NIDDM subjects. Diabetes Care 1994;17:1269-72.
7. Campbell LV, Marmot PR, Dyer JA, Borkman M, Storlien LH. The high monounsaturated fat diet as a practical alternative for NIDDM. Diabetes Care 1994;17:177-82.
8. Lerman-Garber I, Ichazo-Cerro S, Zamora-Gonzalez J, Cardoso-Saldana G, Posadas-Romero C. Effect of a high monounsaturated fat diet enriched with avocado in NIDDM patients. Diabetes Care 1994;17:311-5.
9. Garg A, Grundy SM, Unger RH. Comparison of effects of high and low carbohydrate diets on plasma lipoproteins and insulin sensitivity in patients with mild NIDDM. Diabetes 1992;41:1278-85.
10. Garg A, Bantle JP, Henry RR, et al. Effects of varying carbohydrate content of diet in patients with non-insulin-dependent diabetes mellitus. JAMA 1994;271:1421-8.
11. Keys A, Menotti A, Karvonen MJ, et al. The diet and 15-year death rate in the Seven Countries Study. Am J Epidemiol 1986;124:903-15.
12. Menotti A, Keys A, Aravanis C, et al. Seven Countries Study: first 20-year mortality data in 12 cohorts of six countries. Ann Med 1989;21:175-9.
13. Nestle M. Mediterranean diets: historical and research overview. Am J Clin Nutr 1995;61:Suppl:1313S-1320S.
14. Kromhout D, Keys A, Aravanis C, et al. Food consumption patterns in the 1960s in seven countries. Am J Clin Nutr 1989;49:889-94.
15. Department of Agriculture, Agricultural Research Service. Composition of foods. Agriculture handbook No. 8. Series 8-1 to 8-16. Washington, D.C.: Government Printing Office, 1976-1987.
16. Schakel S, Sievert YA, Buzzard IM. Dietary fiber values for common foods. In: Spiller GA, ed. CRC handbook of dietary fiber in human nutrition. 2nd ed. Boca Raton, Fla.: CRC Press, 1993:567-93.
17. Public Health Service. Manual of laboratory operations: lipid research clinics program: lipid and lipoprotein analysis. 2nd ed. Washington, D.C.: Government Printing Office, 1982.
18. Herbert V, Lau KS, Gottlieb CW, Bleicher SJ. Coated charcoal immunoassay of insulin. J Clin Endocrinol Metab 1965;25:1375-84.
19. Yalow RS, Berson SA. Immunoassay of endogenous plasma insulin in man. J Clin Invest 1960;39:1157-75.
20. Czubayko F, Beumers B, Lammsfuss S, Lutjohann D, von Bergmann K. A simplified micro-method for quantification of fecal excretion of neutral and acidic sterols for outpatient studies in humans. J Lipid Res 1991; 32:1861-7.
21. Lutjohann D, Meese CO, Crouse JR III, von Bergmann K. Evaluation of deuterated cholesterol and deuterated sitostanol for measurement of cholesterol absorption in humans. J Lipid Res 1993;34:1039-46.
22. Jones B, Kenward MG. Design and analysis of crossover trials. London: Chapman & Hall, 1989.
23. Conover WJ. Practical nonparametric statistics. 2nd ed. New York: John Wiley, 1980:288-92.
24. National Health and Nutrition Examination Survey III, 1988-94. NCHS CD-ROM series 11. No. 2A. ASCII version. Hyattsville, Md.: National Center for Health Statistics, April 1998.
25. Kiehm TG, Anderson JW, Ward K. Beneficial effects of a high carbohydrate, high fiber diet on hyperglycemic diabetic men. Am J Clin Nutr 1976;29:895-9.
26. Simpson HCR, Simpson RW, Lousley S, et al. A high carbohydrate leguminous fibre diet improves all aspects of diabetic control. Lancet 1981; 1:1-5.
27. Rivellese A, Riccardi G, Giacco A, et al. Effect of dietary fibre on glucose control and serum lipoproteins in diabetic patients. Lancet 1980;2: 447-50.
28. Riccardi G, Rivellese A, Pacioni D, Genovese S, Mastranzo P, Mancini M. Separate influence of dietary carbohydrate and fibre on the metabolic control in diabetes. Diabetologia 1984;26:116-21.
29. O'Dea K, Traianedes K, Ireland P, et al. The effects of diet differing in fat, carbohydrate, and fiber on carbohydrate and lipid metabolism in type II diabetes. J Am Diet Assoc 1989;89:1076-86.
30. Nuttall FQ. Dietary fiber in the management of diabetes. Diabetes 1993;42:503-8.
31. Holman RR, Steemson J, Darling P, Turner RC. No glycemic benefit from guar administration in NIDDM. Diabetes Care 1987;10:68-71.
32. Uusitupa M, Siitonen O, Savolainen K, Silvasti M, Penttila I, Parviainen M. Metabolic and nutritional effects of long-term use of guar gum in the treatment of noninsulin-dependent diabetes of poor metabolic control. Am J Clin Nutr 1989;49:345-51.
33. Aro A, Uusitupa M, Voutilainen E, Hersio K, Korhonen T, Siitonen O. Improved diabetic control and hypocholesterolaemic effect induced by long-rerm dietary supplementation with guar gum in type 2 (insulin-independent) diabetes. Diabetologia 1981;21:29-33.
34. Pick ME, Hawrysh ZJ, Gee MI, Toth E, Garg ML, Hardin RT. Oat

(189)
[131]

Severe Hypertriglyceridemia with Insulin Resistance Is Associated with Systemic Inflammation: Reversal with Bezafibrate Therapy in a Randomized Controlled Trial

Iris J. Jonkers, MD, Martina F. Mohrschladt, MD, Rudi G. Westendorp, MD, PhD, Arnoud van der Laarse, PhD, Augustinus H. Smelt, MD, PhD

PURPOSE: To determine whether hypertriglyceridemia is associated with systemic inflammation, which may contribute to the increased cardiovascular risk in patients who have hypertriglyceridemia. In addition, we investigated whether fibrates reverse this inflammatory state.

PATIENTS AND METHODS: Serum lipid levels, body mass index, insulin resistance, and inflammatory parameters were compared between 18 patients who had severe hypertriglyceridemia without cardiovascular disease and 20 normolipidemic controls. We measured the ex vivo production capacity of tumor necrosis factor (TNF)-α and interleukin (IL)-6 after whole-blood stimulation with lipopolysaccharide, as well as circulating levels of C-reactive protein and fibrinogen. A randomized controlled trial was conducted to determine whether bezafibrate (400 mg administered daily for 6 weeks) affected these parameters in hypertriglyceridemic patients.

RESULTS: When compared with normolipidemic controls, hypertriglyceridemic patients had significantly lower high-density lipoprotein (HDL) cholesterol and higher triglyceride levels, body mass index, and insulin resistance. In addition, hypertriglyceridemic patients had a significantly higher production capacity of TNF-α (mean difference, 11 700 pg/mL; 95% confidence interval [CI]: 7800 to 15 700 pg/mL]) and IL-6 (mean difference, 20 400 pg/mL; 95% CI: 7800 to 32 900 pg/mL), and higher levels of C-reactive protein (mean difference, 0.8 mg/L; 95% CI: 0.1 to 2.4 mg/L) and fibrinogen (mean difference, 0.8 g/dL; 95% CI: 0.3 to 1.3 g/dL). Bezafibrate therapy significantly increased HDL cholesterol levels, reduced triglyceride and insulin resistance levels, and reduced production capacity of TNF-α and IL-6, as well as levels of C-reactive protein and fibrinogen.

CONCLUSION: Systemic inflammation is present in patients who have the clinical phenotype that is associated with severe hypertriglyceridemia, and may contribute to the increased risk of cardiovascular disease in these patients. Bezafibrate has anti-inflammatory effects in these patients. Am J Med. 2002;112: 275–280. ©2002 by Excerpta Medica, Inc.

Inflammation plays an important role in the etiology of cardiovascular disease (1). Elevated plasma levels of acute phase proteins, such as C-reactive protein (2–4) and fibrinogen (5), are associated with cardiovascular disease (2–5). Hypertriglyceridemia, per se, may be a risk factor for cardiovascular disease (6), and is associated with elevated levels of fibrinogen (7,8).

We have demonstrated that subjects with hypertriglyceridemia have elevated tumor necrosis factor (TNF)-α production capacity (9). Since TNF-α is an early mediator of the acute phase response (10), we hypothesized that an increased production of proinflammatory cytokines in patients with hypertriglyceridemia induces systemic inflammation, as represented by elevated plasma levels of C-reactive protein and fibrinogen. If this is true, systemic inflammation might contribute to the risk of cardiovascular disease in patients with hypertriglyceridemia. Therefore, the first aim of this study was to investigate whether hypertriglyceridemia is associated with an inflammatory state, by comparing the ex vivo production capacity of TNF-α and interleukin (IL)-6, as well as plasma levels of C-reactive protein and fibrinogen between patients with endogenous severe hypertriglyceridemia without clinically manifest cardiovascular disease and normolipidemic controls.

Fibrates, which are peroxisome proliferator–activated receptor (PPAR)-α activators, are used for the treatment of patients with severe hypertriglyceridemia. Besides their lipid-lowering effects, they have been shown to reduce the risk of cardiovascular events (11,12). The beneficial effects of fibrates are far greater than expected on the basis of change in lipoprotein concentrations. Data from the Veterans Affairs High-Density Lipoprotein Cholesterol Intervention Trial showed that only 23% of the gemfibrozil-associated treatment benefit could be attributed to lower lipid levels (13). Recently, it was shown that absence of PPAR-α expression in mice augmented the inflammatory response (14). Thus, the antiatherogenic effects of fibrates may be mediated by a reduction in the inflammatory state. Therefore, the second aim of this study was to investigate the effects of bezafibrate therapy on inflammatory markers in patients with endogenous severe hypertriglyceridemia.

From the Departments of General Internal Medicine (IJJ, MFM, RGW, AHS) and Cardiology (AVDL), Leiden University Medical Center, Leiden, The Netherlands.

Supported by the Praeventiefonds (#28-1642-4), The Hague, The Netherlands, and by an unrestricted research grant from Boehringer Mannheim, Mannheim, Germany.

Requests for reprints should be addressed to Augustinus H. Smelt, MD, PhD, Department of General Internal Medicine, B3-Q, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands.

Manuscript submitted July 26, 2001, and accepted in revised form November 7, 2001.

0002-9343/02/$–see front matter 275
PII S0002-9343(01)01123-8

METHODS

The study sample consisted of 18 unrelated patients with endogenous severe hypertriglyceridemia who were recruited from the lipid clinic of the Leiden University Medical Center, The Netherlands. All patients received personal dietary advice. The diagnosis was based on the mean of two fasting blood samples obtained after a prudent diet of at least 8 weeks. Diagnostic criteria for endogenous hypertriglyceridemia (15) were total serum triglyceride level >354 mg/dL (4.0 mmol/L), very low-density lipoprotein (VLDL) cholesterol level >39 mg/dL (1.0 mmol/L), and low-density lipoprotein (LDL) cholesterol <174 mg/dL (4.5 mmol/L). Exclusion criteria were the presence of cardiovascular disease, diabetes mellitus, current infection, homozygosity for apolipoprotein E2, or secondary hyperlipidemia (renal, liver, or thyroid disease; fasting glucose level >126 mg/dL [7.0 mmol/L]; alcohol intake >40 g/d), and the use of lipid-lowering, anti-inflammatory, or glucose-lowering drugs. The presence of cardiovascular disease was excluded by history, physical examination, and the presence of a normal electrocardiogram. Twenty normolipidemic, age- and sex-matched control subjects were recruited via a newspaper advertisement.

Study Design

The patients were randomly assigned to receive, in a double-blind crossover fashion, bezafibrate (400 mg administered once daily) or placebo for 6 weeks. The two treatment periods were separated by a 6-week washout period. Before and at the end of each treatment period, fasting venous blood samples were obtained to determine levels of lipids, insulin, glucose, C-reactive protein, and fibrinogen. Insulin resistance was assessed using the homeostasis model approximation, which correlates well with results from both hyperinsulinemic euglycemic clamp and the intravenous glucose tolerance test, by the following formula: insulin resistance = insulin/($22.5e^{-\ln(\text{glucose})}$) (16). From the control subjects, fasting blood samples were obtained at baseline. In a subset of the sample (13 patients and 19 controls), production capacity of IL-6 and TNF-α was determined in fasting venous blood. Tumor necrosis factor-α production capacity in this sample has been described (9). Informed consent was obtained from each participant, and the protocol was approved by the Medical Ethics Committee of our institution.

Laboratory Measurements

Ultracentrifugation was performed to determine serum VLDL, LDL, and high-density lipoprotein (HDL) cholesterol levels (15). Triglyceride and total cholesterol levels were measured using commercially available kits. Apolipoprotein B levels were determined using an enzyme-linked immunosorbent assay. The contribution of apolipoprotein B48 to the total amount of apolipoprotein B was less than 10% in samples from hypertriglyceridemic patients, indicating a limited contribution of chylomicron remnants to the isolated VLDL fraction. Insulin was measured with a radioimmunoassay (MedGenix, Brussels, Belgium). Glucose was measured with a Hitachi 747 analyzer (Roche Diagnostics, Mannheim, Germany). Plasma concentrations of C-reactive protein were measured in a single batch using a highly sensitive sandwich immunoassay (Kordia Laboratory Supplies, Leiden, The Netherlands). For determination of fibrinogen level, venous blood was collected in siliconized Vacutainer tubes (Becton Dickinson, Milan, Italy) in 0.1-volume 0.13-M trisodium citrate. Fibrinogen was assayed according to Von Clauss (17).

Cytokine Production

Blood samples were kept in endotoxin-free heparin tubes (4.0-mL Chromogenix endotube, Mölndal, Sweden) with 120 IU of sodium heparin (18). Briefly, whole-blood samples were diluted 1:1 with RPMI 1640 (Gibco Life Technologies, Paisley, United Kingdom) and divided in 1.0-mL aliquots in polystyrene microtiter plates. Serial 10-fold dilutions were prepared once daily from 10 μL of 250-μL aliquots of lipopolysaccharide (*Escherichia coli* 0111; B4, Boivin method, Difco Laboratories, Detroit, Michigan), which were added to the diluted whole-blood aliquots, to obtain final lipopolysaccharide concentrations of 10 ng/mL, 100 ng/mL, and 1000 ng/mL in each well. Whole-blood samples were incubated at 37°C under 5% carbon dioxide with increasing concentrations of lipopolysaccharide, for 4 hours to determine the production capacity of TNF-α, and for 24 hours to determine the production capacity of IL-6. Both cytokines are derived mainly from monocytes (19). Stimulation of 1:1 diluted whole-blood without addition of lipopolysaccharide was performed as a negative control. After stimulation, microtiter plates were centrifuged twice at 600 g and the supernatants were stored at −70°C until determination of TNF-α and IL-6. Enzyme-linked immunosorbent assays for cytokine measurement were performed according to the manufacturer's guidelines (TNF-α, Central Laboratory of the Blood Transfusion Service, Amsterdam, The Netherlands; IL-6, BioSource, Fleurus, Belgium). All supernatants were analyzed in a single batch. For both cytokines, the detection limit was 4 pg/mL of whole blood. Intra- and interassay coefficients of varia-

tion were 6.8% and 8.1% for TNF-α and 7.6% and 8.5% for IL-6.

Data Analysis

Baseline comparisons. Data are expressed as means ± SD, except for C-reactive protein and triglyceride levels, which are expressed as medians (interquartile ranges) because of their skewed distribution. Differences in baseline C-reactive protein and triglyceride levels between patients and controls were compared using the Mann-Whitney *U* test. Differences in parametric variables between these groups were compared using the *t* test, except for differences in cytokine profiles between patients and controls, which were evaluated with analysis of variance for the three different lipopolysaccharide conditions simultaneously. Difference in sex and smoking habits between patients and controls were compared using the Fisher exact test. Ninety-five percent confidence intervals were calculated for the mean differences, except for C-reactive protein and triglyceride levels, for which 95% confidence intervals were calculated for the median differences.

Effects of bezafibrate therapy. Differences between placebo and bezafibrate therapy were evaluated using a paired sample *t* test, except for levels of C-reactive protein and triglyceride, which were compared using the Wilcoxon signed rank test. Treatment with placebo did not affect serum lipid levels in hypertriglyceridemic patients; therefore, the data for placebo were assigned as baseline data for the hypertriglyceridemic patients. The effect of bezafibrate therapy on cytokine profiles was analyzed using general linear model analysis, taking into account the full set of observations in each patient. The bezafibrate-induced difference for each parameter was calculated by subtracting the value on placebo from that obtained upon bezafibrate therapy. Associations between the paired differences in systemic inflammatory markers and the difference in lipids and insulin resistance were studied using the Spearman rank correlation analysis (n = 13). Ninety-five percent confidence intervals were calculated for the mean differences, except for C-reactive protein and triglyceride levels, for which 95% confidence intervals were calculated for the median differences.

RESULTS

There were no differences between hypertriglyceridemic patients and controls in age, sex, and smoking status, but body mass index was significantly higher in patients with hypertriglyceridemia. Significant differences between both groups were observed for insulin and glucose levels, and insulin resistance. Apolipoprotein B levels did not differ between hypertriglyceridemic patients and controls. Total cholesterol, total triglyceride, and VLDL cholesterol levels were significantly higher, whereas HDL and LDL cholesterol levels were lower in hypertriglyceridemic patients compared with controls (Table 1).

Compared with controls, hypertriglyceridemic patients had higher TNF-α and IL-6 production capacity upon whole-blood stimulation, and higher C-reactive protein and fibrinogen levels (Table 1).

Effects of Bezafibrate Therapy in Hypertriglyceridemic Patients

All patients completed the study without reporting any adverse effects. No significant changes occurred in body mass index (mean [± SD] body mass index, 28.0 ± 2.8 kg/m^2 on placebo vs. 28.4 ± 2.7 kg/m^2 on bezafibrate). Bezafibrate reduced insulin levels and insulin resistance significantly without affecting glucose levels. Bezafibrate reduced serum total cholesterol, total triglyceride, and VLDL cholesterol levels significantly, and increased serum HDL and LDL cholesterol levels significantly (Table 2).

Bezafibrate reduced the production capacity of TNF-α and IL-6, and plasma C-reactive protein and fibrinogen levels (Table 2). One patient had a C-reactive protein level >10 mg/L, a commonly used indicator of current infection. When this patient was excluded, the differences in C-reactive protein levels between hypertriglyceridemic patients and controls and between hypertriglyceridemic patients on placebo and bezafibrate were similar.

Associations of Paired Differences in Systemic Inflammatory Parameters with Bezafibrate Therapy

The bezafibrate-induced reduction in C-reactive protein level correlated significantly with the reduction in fibrinogen level ($r = 0.56$, $P = 0.02$). Changes in C-reactive protein and fibrinogen levels were not correlated with changes in lipid, insulin, or glucose levels, or insulin resistance (all $P > 0.3$, Table 3).

DISCUSSION

We found that endogenous severe hypertriglyceridemia is associated with an inflammatory state, as represented by higher ex vivo production capacity of TNF-α and IL-6 and elevated levels of C-reactive protein and fibrinogen. Six-week treatment with bezafibrate reduced these markers of inflammation significantly in patients with severe hypertriglyceridemia. The observed increase in IL-6 production and the elevated C-reactive protein levels in patients with severe hypertriglyceridemia without known cardiovascular disease add to the earlier findings of increased TNF-α production capacity (9) and elevated fibrinogen levels in hypertriglyceridemia (7,8). These findings may be of clinical importance, since C-reactive protein (2–4) and fibrinogen (5) are thought to be independent risk factors for cardiovascular disease.

Bezafibrate reduced the ex vivo production capacity of TNF-α and IL-6 significantly, as has been seen for fenofibrate in mildly hyperlipidemic men (20,21). In addition, C-reactive protein and fibrinogen levels were reduced significantly by bezafibrate. Two studies have investigated the effects of fibrates on C-reactive protein

Table 1. Baseline Characteristics of Controls and Hypertriglyceridemic Patients

Characteristic (Units)	Controls (n = 20)	Hypertriglyceridemic (n = 18)	Mean Difference (95% Confidence Interval)*
	Number (%), Mean ± SD, or Median (Interquartile Range)		
Male sex	18 (90)	16 (89)	. . .
Current smoking	2 (10)	7 (39)	. . .
Age (years)	47.9 ± 7.4	48.5 ± 8.8	0.7 (−4.7 to 6.0)
Body mass index (kg/m^2)	24.2 ± 3.3	28.0 ± 2.8	4.2 (1.9 to 5.9)
Insulin (mU/L)	11.1 ± 6.8	55.6 ± 26.3	44.5 (32.2 to 56.9)
Glucose (mg/dL)	93 ± 6	107 ± 14	14 (7 to 21)
Insulin resistance (HOMA index)	2.6 ± 1.7	14.6 ± 7.0	12.1 (8.8 to 15.3)
Apolipoprotein B (mg/dL)	106 ± 22	111 ± 19	5 (−8 to 19)
Total cholesterol (mg/dL)	197 ± 36	299 ± 91	102 (57 to 147)
Total triglycerides (mg/dL)	82 (52–97)	958 (556–1141)	859 (599 to 1019)
VLDL cholesterol (mg/dL)	10 ± 7	169 ± 78	159 (123 to 195)
LDL cholesterol (mg/dL)	136 ± 34	103 ± 25	−33 (−53 to −13)
HDL cholesterol (mg/dL)	51 ± 11	28 ± 5	−23 (−29 to −17)
TNF-α production (pg/mL)†	23 200 ± 8700	35 000 ± 10 000	11 700 (7800 to 15 700)
IL-6 production (pg/mL)†	61 900 ± 27 800	82 200 ± 31 900	20 400 (7800 to 32 900)
C-reactive protein (mg/L)	0.9 (0.4–2.4)	2.2 (1.0–5.8)	0.8 (0.1 to 2.4)
Fibrinogen (g/dL)	2.8 ± 0.5	3.6 ± 1.0	0.8 (0.3 to 1.3)

* Median difference for C-reactive protein and triglycerides.

† As obtained from stimulation with lipopolysaccharide at concentrations of 10 ng/mL, 100 ng/mL, and 1000 ng/mL; data are based on measurements in 19 controls and 13 hypertriglyceridemic patients.

HDL = high-density lipoprotein; HOMA = homeostasis model assessment; IL-6 = interleukin-6; LDL = low-density lipoprotein; TNF-α = tumor necrosis factor-α; VLDL = very low-density lipoprotein.

Table 2. Effects of Placebo and Bezafibrate on Serum Lipids, Insulin, Glucose, Insulin Resistance, and Inflammatory Parameters in 18 Hypertriglyceridemic Patients

Parameter (Units)	Placebo	Bezafibrate	Mean Difference (95% Confidence Interval)*
	Mean ± SD or Median (Interquartile Range)		
Insulin (mU/L)	55.6 ± 26.3	25.5 ± 15.7	−30.1 (−42.4 to −17.7)
Glucose (mg/dL)	107 ± 14	104 ± 11	−3 (−5 to 1)
Insulin resistance (HOMA index)	14.6 ± 7.0	6.7 ± 4.7	−7.9 (−11.0 to −4.8)
Total cholesterol (mg/dL)	299 ± 92	228 ± 48	−71 (−106 to −36)
Total triglycerides (mg/dL)	958 (556–1141)	359 (276–432)	−564 (−1054 to −416)
VLDL cholesterol (mg/dL)	169 ± 78	55 ± 24	−114 (−147 to −80)
LDL cholesterol (mg/dL)	103 ± 25	138 ± 33	35 (21 to 50)
HDL cholesterol (mg/dL)	28 ± 5	35 ± 5	7 (6 to 9)
TNF-α production (pg/mL)†	35 000 ± 6500	31 400 ± 6500	−3600 (−6500 to −600)
IL-6 production (pg/mL)†	82 200 ± 23 600	69 000 ± 23 900	−13 200 (−23 900 to −2500)
C-reactive protein (mg/L)	2.2 (1.0–5.8)	1.2 (0.8–2.5)	−0.8 (−3.2 to 0.0)
Fibrinogen (g/dL)	3.6 ± 1.0	2.8 ± 0.5	−0.8 (−1.2 to −0.4)

* Median difference for C-reactive protein and triglycerides.

† In 13 subjects.

HDL = high-density lipoprotein; HOMA = homeostasis model assessment; IL-6 = interleukin-6; LDL = low-density lipoprotein; TNF-α = tumor necrosis factor-α; VLDL = very low-density lipoprotein.

Table 3. Correlations between Differences (Placebo-Bezafibrate) in Inflammatory Parameters, Lipoproteins, Glucose, and Insulin in Hypertriglyceridemic Patients

	Δ C-Reactive Protein	Δ Fibrinogen
	Correlation Coefficient (r)	
Δ Fibrinogen	0.56*	...
Δ Triglycerides	−0.04	0.15
Δ Total Cholesterol	0.07	0.16
Δ HDL Cholesterol	0.03	0.22
Δ LDL Cholesterol	−0.01	−0.15
Δ VLDL Cholesterol	−0.06	0.08
Δ Insulin	−0.12	0.18
Δ Glucose	−0.10	0.02
Δ Insulin resistance	−0.10	0.19

* $P = 0.02$. All other correlations are not significant.
Δ = change in; HDL = high-density lipoprotein; LDL = low-density lipoprotein; VLDL = very low-density lipoprotein.

levels. Staels et al. observed that 4 weeks of fenofibrate therapy reduced C-reactive protein levels significantly in mildly hyperlipidemic patients (21), whereas bezafibrate did not affect C-reactive protein levels in patients who had type 2b hyperlipidemia and cardiovascular disease (22).

Hypertriglyceridemia is part of a complex metabolic disorder, characterized by high serum triglyceride levels and low HDL cholesterol levels. It is also associated with insulin resistance and obesity (23). This characteristic phenotype is likely to contribute to the inflammatory state in severe hypertriglyceridemia. Several reports have described associations between insulin resistance and obesity, and systemic inflammation (24–26). Furthermore, HDL cholesterol has anti-inflammatory effects, as demonstrated by its capacity to clear inflammatory mediators (27), to inhibit the expression of adhesion molecules on the endothelial wall (28), and to scavenge oxidation products (29). Thus, the inflammatory state in subjects with severe hypertriglyceridemia may be linked to one or more of these phenotypic features. Identification of the metabolic pathway that leads to systemic inflammation requires further experiments because of the interrelation between inflammation, lipids, obesity, and insulin resistance (10,30–32).

Our observations of the effects of fibrates on C-reactive protein level (33,34) and production capacity of TNF-α and IL-6 (35) are similar to those seen with statins. In the studies of statins (33,34), as well as in the current study, no relation was observed between therapy-induced changes in C-reactive protein and lipid levels. In addition, we could not demonstrate relations between the change in systemic inflammatory parameters and insulin resistance.

The mechanism of the anti-inflammatory properties of statins is not known. Statins reduce the expression of messenger ribonucleic acid, as well as lower protein levels of proinflammatory cytokines and PPAR-α in human endothelial cells and hepatocytes (36). Statin-induced inhibition of the Rho signal pathway activates PPAR-α (37). In addition, PPAR-α mediates the anti-inflammatory effects of fibrates by negative regulation of the nuclear factor-κB and activated protein-1 signaling pathways (38), which are responsible for activation of inflammatory response genes, such as TNF-α and IL-6. Thus, although fibrates, synthetic PPAR-α agonists, and statins activate PPAR-α by different pathways, their anti-inflammatory effects may be mediated through a common mechanism.

The anti-inflammatory effects of fibrates may be of clinical relevance, since C-reactive protein and fibrinogen levels are associated with cardiovascular disease (2–5). Unfortunately, no data are available on the effects of fibrates on inflammatory parameters in previous large-scale trials. Whether the anti-inflammatory effects of bezafibrate in patients with severe hypertriglyceridemia occur in other patients and contribute to the fibrate-associated cardiovascular risk reduction in patients with mild-to-moderate hypertriglyceridemia needs to be determined.

In conclusion, we found that severe hypertriglyceridemia with insulin resistance is associated with systemic inflammation, which may contribute to the increased risk of cardiovascular disease in these patients. Bezafibrate therapy has anti-inflammatory effects, as represented by reductions in the ex vivo production capacity of TNF-α and IL-6 and in the plasma levels of C-reactive protein and fibrinogen.

REFERENCES

1. Ross R. Atherosclerosis—an inflammatory disease. *N Engl J Med.* 1999;340:115–126.
2. Ridker PM, Rifai N, Pfeffer MA, et al. Inflammation, pravastatin, and the risk of coronary events after myocardial infarction in patients with average cholesterol levels. Cholesterol and Recurrent Events (CARE) Investigators. *Circulation.* 1998;98:839–844.
3. Thompson SG, Kienast J, Pyke SD, et al. Hemostatic factors and the risk of myocardial infarction or sudden death in patients with angina pectoris. European Concerted Action on Thrombosis and Disabilities Angina Pectoris Study Group. *N Engl J Med.* 1995;332:635–641.
4. Ridker PM, Cushman M, Stampfer MJ, et al. Inflammation, aspirin, and the risk of cardiovascular disease in apparently healthy men. *N Engl J Med.* 1997;336:973–979.
5. Maresca G, Di Blasio A, Marchioli R, Di Minno G. Measuring plasma fibrinogen to predict stroke and myocardial infarction: an update. *Arterioscler Thromb Vasc Biol.* 1999;19:1368–1377.
6. Austin MA, Hokanson JE, Edwards KL. Hypertriglyceridemia as a cardiovascular risk factor. *Am J Cardiol.* 1998;81:7B–12B.
7. Simpson HC, Mann JI, Meade TW, et al. Hypertriglyceridaemia and hypercoagulability. *Lancet.* 1983;1:786–790.
8. Avellone G, Di Garbo V, Cordova R, et al. Fibrinolysis in hypertriglyceridaemic subjects in response to venous occlusion. *Blood Coagul Fibrinolysis.* 1993;4:429–433.

9. Mohrschladt MF, Weverling-Rijnsburger AW, de Man FH, et al. Hyperlipoproteinemia affects cytokine production in whole blood samples ex vivo. The influence of lipid-lowering therapy. *Atherosclerosis.* 2000;148:413–419.
10. Richards C, Gauldie J. Role of cytokines in the acute phase response. Aggarwal BB, Puri RK, eds. *Human Cytokines: Their Roles in Disease Therapy.* Cambridge, Massachusetts: Blackwell Science; 1995:253–269.
11. Rubins HB, Robins SJ, Collins D, et al. Gemfibrozil for the secondary prevention of coronary heart disease in men with low levels of high-density lipoprotein cholesterol. Veterans Affairs High-Density Lipoprotein Cholesterol Intervention Trial Study Group. *N Engl J Med.* 1999;341:410–418.
12. Ruotolo G, Ericsson CG, Tettamanti C, et al. Treatment effects on serum lipoprotein lipids, apolipoproteins and low density lipoprotein particle size and relationships of lipoprotein variables to progression of coronary artery disease in the Bezafibrate Coronary Atherosclerosis Intervention Trial (BECAIT). *J Am Coll Cardiol.* 1998;32:1648–1656.
13. Robins SJ, Collins D, Wittes JT, et al. Relation of gemfibrozil treatment and lipid levels with major coronary events: VA-HIT: a randomized controlled trial. *JAMA.* 2001;285:1585–1591.
14. Devchand PR, Keller H, Peters JM, et al. The PPAR alpha-leukotriene B4 pathway to inflammation control. *Nature.* 1996;384:39–43.
15. de Man FH, Weverling-Rijnsburger AW, van der Laarse A, et al. Not acute but chronic hypertriglyceridemia is associated with impaired endothelium-dependent vasodilation: reversal after lipid-lowering therapy by atorvastatin. *Arterioscler Thromb Vasc Biol.* 2000;20:744–750.
16. Matthews DR, Hosker JP, Rudenski AS, et al. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. *Diabetologia.* 1985;28:412–419.
17. Von Clauss A. Gerinnungsphysiologische schnellmethode zur beatimmung des fibrinogens. *Acta Huematol.* 1957;17:237–246.
18. Van der Linden MW, Huizinga TWJ, Stoeken DJ, et al. Determination of tumor necrosis factor α and interleukin 10 production in a whole blood stimulation system: assessment of laboratory error and individual variation. *J Immunol Methods.* 1998;218:63–71.
19. De Grootte D, Zangerle PF, Gevaert Y, et al. Direct stimulation of cytokines (IL-1β, TNF-α, IL-6, IL-2, INF-γ and GM-CSF) in whole blood. I. Comparison with isolated PMBC stimulation. *Cytokine.* 1992;4:239–248.
20. Madej A, Okopien B, Kowalski J, et al. Effects of fenofibrate on plasma cytokine concentrations in patients with atherosclerosis and hyperlipoproteinemia IIb. *Int J Clin Pharmacol Ther.* 1998;36:345–349.
21. Staels B, Koenig W, Habib A, et al. Activation of human aortic smooth-muscle cells is inhibited by PPARalpha but not by PPAR-gamma activators. *Nature.* 1998;393:790–793.
22. Cortellaro M, Cofrancesco E, Boschetti C, et al. Effects of fluvastatin and bezafibrate combination on plasma fibrinogen, t-plasminogen activator inhibitor and C reactive protein levels in coronary artery disease patients with mixed hyperlipidaemia (FACT study). Fluvastatin Alone and in Combination Treatment. *Thromb Haemost.* 2000;83:549–553.
23. Grundy SM. Hypertriglyceridemia, insulin resistance, and the metabolic syndrome. *Am J Cardiol.* 1998;83:25F–29F.
24. Festa A, D'Agostino RJ, Howard G, et al. Chronic subclinical inflammation as part of the insulin resistance syndrome: the Insulin Resistance Atherosclerosis Study (IRAS). *Circulation.* 2000;102:42–47.
25. Yudkin JS, Stehouwer CD, Emeis JJ, Coppack SW. C-reactive protein in healthy subjects: associations with obesity, insulin resistance, and endothelial dysfunction: a potential role for cytokines originating from adipose tissue? *Arterioscler Thromb Vasc Biol.* 1999;19:972–978.
26. Lemieux I, Pascot A, Prud'homme D, et al. Elevated C-reactive protein: another component of the atherothrombotic profile of abdominal obesity. *Arterioscler Thromb Vasc Biol.* 2001;21:961–967.
27. Levine DM, Parker TS, Donnelly TM, et al. In vivo protection against endotoxin by plasma high density lipoprotein. *Proc Natl Acad Sci USA.* 1993;90:12040–12044.
28. Cockerill GW, Rye KA, Gamble JR, et al. High-density lipoproteins inhibit cytokine-induced expression of endothelial cell adhesion molecules. *Arterioscler Thromb Vasc Biol.* 1995;15:1987–1994.
29. Mackness MI, Durrington PN. HDL, its enzymes and its potential to influence lipid peroxidation. *Atherosclerosis.* 1995;115:243–253.
30. Dichtl W, Nilsson L, Goncalves I, et al. Very low-density lipoprotein activates nuclear factor-kappaB in endothelial cells. *Circ Res.* 1999;84:1085–1094.
31. Knudsen P, Eriksson J, Lahdenpera S, et al. Changes of lipolytic enzymes cluster with insulin resistance syndrome. Botnia Study Group. *Diabetologia.* 1995;38:344–350.
32. Steiner G. Altering triglyceride concentrations changes insulin-glucose relationships in hypertriglyceridemic patients. Double-blind study with gemfibrozil with implications for atherosclerosis. *Diabetes Care.* 1991;14:1077–1081.
33. Ridker PM, Rifai N, Clearfield M, et al. Measurement of C-reactive protein for the targeting of statin therapy in the primary prevention of acute coronary events. *N Engl J Med.* 2001;344:1959–1965.
34. Albert MA, Danielson E, Rifai N, Ridker PM. Effect of statin therapy on C-reactive protein levels. *JAMA.* 2001;286:64–70.
35. Rosenson RS, Tangney CC, Casey LC. Inhibition of proinflammatory cytokine production by pravastatin. *Lancet.* 1999;353:983–984.
36. Inoue I, Goto S, Mizotani K, et al. Lipophilic HMG-CoA reductase inhibitor has an anti-inflammatory effect: reduction of mRNA levels for interleukin-1beta, interleukin-6, cyclooxygenase-2, and p22phox by regulation of peroxisome proliferator-activated receptor alpha (PPARalpha) in primary endothelial cells. *Life Sci.* 2000;67:863–876.
37. Martin G, Duez H, Blanquart C, et al. Statin-induced inhibition of the Rho-signaling pathway activates PPARalpha and induces HDL apoA-I. *J Clin Invest.* 2001;107:1423–1432.
38. Delerive P, De Bosscher K, Besnard S, et al. Peroxisome proliferator-activated receptor α negatively regulates the vascular inflammatory gene response by negative cross-talk with transcription factors NF-κB and AP-1. *J Biol Chem.* 1999;45:32048–32054.

440) 1997]

Substituting Walnuts for Monounsaturated Fat Improves the Serum Lipid Profile of Hypercholesterolemic Men and Women

A Randomized Crossover Trial

Daniel Zambón, MD; Joan Sabaté, MD, DrPH; Sonia Muñoz, PhD; Betina Campero, RD; Elena Casals, MD; Manuel Merlos, PhD; Juan C. Laguna, PhD; and Emilio Ros, MD

Background: It has been reported that walnuts reduce serum cholesterol levels in normal young men.

Objective: To assess the acceptability of walnuts and their effects on serum lipid levels and low-density lipoprotein (LDL) oxidizability in free-living hypercholesterolemic persons.

Design: Randomized, crossover feeding trial.

Setting: Lipid clinic at a university hospital.

Patients: 55 men and women (mean age, 56 years) with polygenic hypercholesterolemia.

Intervention: A cholesterol-lowering Mediterranean diet and a diet of similar energy and fat content in which walnuts replaced approximately 35% of the energy obtained from monounsaturated fat. Patients followed each diet for 6 weeks.

Measurements: Low-density lipoprotein fatty acids (to assess compliance), serum lipid levels, lipoprotein(a) levels, and LDL resistance to in vitro oxidative stress.

Results: 49 persons completed the trial. The walnut diet was well tolerated. Planned and observed diets were closely matched. Compared with the Mediterranean diet, the walnut diet produced mean changes of −4.1% in total cholesterol level, −5.9% in LDL cholesterol level, and −6.2% in lipoprotein(a) level. The mean differences in the changes in serum lipid levels were −0.28 mmol/L (95% CI, −0.43 to −0.12 mmol/L) (−10.8 mg/dL [−16.8 to −4.8 mg/dL]) ($P < 0.001$) for total cholesterol level, −0.29 mmol/L (CI, −0.41 to −0.15 mmol/L) (−11.2 mg/dL [−16.3 to −6.1 mg/dL]) ($P < 0.001$) for LDL cholesterol level, and −0.021 g/L (CI, −0.042 to −0.001 g/L) ($P = 0.042$) for lipoprotein(a) level. Lipid changes were similar in men and women except for lipoprotein(a) levels, which decreased only in men. Low-density lipoprotein particles were enriched with polyunsaturated fatty acids from walnuts, but their resistance to oxidation was preserved.

Conclusion: Substituting walnuts for part of the monounsaturated fat in a cholesterol-lowering Mediterranean diet further reduced total and LDL cholesterol levels in men and women with hypercholesterolemia.

Ann Intern Med. 2000;132:538-546.

For author affiliations, current addresses, and contributions, see end of text.

Lifestyle modification is the cornerstone of population-based strategies for prevention of coronary heart disease and is the first line of therapy in patients with hypercholesterolemia. Diets low in saturated fatty acids and cholesterol have long been recommended to decrease low-density lipoprotein (LDL) cholesterol levels and reduce cardiovascular risk (1). Ample evidence suggests that polyunsaturated fatty acids and monounsaturated fatty acids have a similar cholesterol-lowering effect when substituted for saturated fatty acids (2–4). However, most studies of fatty acids and blood lipids have been done with fats and oils, rarely with whole fatty foods. Because people usually buy and consume whole food products, it is desirable to know the effects of specific foods on risk factors for coronary heart disease.

Recent reports suggest that the regular consumption of nuts might reduce cardiovascular risk (5). Walnuts are particularly rich in polyunsaturated fatty acids (6), and epidemiologic evidence suggests that frequent walnut consumption protects against coronary heart disease (7). In a controlled feeding trial by Sabaté and colleagues (8), a diet in which walnuts represented 55% of the energy from fat reduced blood cholesterol levels in normal young men when compared with a standard low-fat diet. However, the results cannot easily be extrapolated to the population at risk for coronary heart disease because women, older age groups, and hypercholesterolemic persons were not studied (8). In addition, because meals were served at a metabolic kitchen, the study did not address the question of whether free-living persons would incorporate substantial quantities of walnuts into their diets. Because oxidized LDL plays a key role in atherogenesis (9) and oxidative damage involves peroxidation of polyunsaturated fatty acids in LDL lipids (10), there is concern that walnut intake may promote LDL oxidation. Therefore, we designed a dietary intervention study in free-living adult men and women with polygenic hypercholesterolemia to compare the effects of a walnut-rich diet with those of a cholesterol-lowering Mediterranean diet on serum lipid levels, lipoprotein levels, and LDL resistance to oxidation.

Methods

Patients

Adult men and women with polygenic hypercholesterolemia attending the Lipid Clinic of the Hospital Clínic of Barcelona were eligible if they had serum LDL cholesterol concentrations greater than 3.36 mmol/L (130 mg/dL) and triglyceride concentrations less than 2.82 mmol/L (250 mg/dL); no evidence of alcohol, tobacco, or drug abuse; absence of diabetes mellitus and liver, kidney, thyroid, or other endocrine diseases, as assessed by medical history, a complete physical examination, and laboratory tests; no intake in the previous 8 weeks of medications known to affect lipid metabolism, including hypolipidemic agents and estrogen compounds in women; infrequent consumption of nuts and no known history of allergy to them; and no use of multivitamin or vitamin E supplements. Because the target population had common (polygenic) hypercholesterolemia, we excluded persons whose elevated blood cholesterol levels had a strong genetic basis (such as heterozygous familial hypercholesterolemia or familial combined hyperlipidemia), as established by standard criteria. On admission to the Lipid Clinic, all patients were advised to follow a Mediterranean-type hypolipidemic diet (11).

For a crossover design, statistical power calculations indicated that to detect mean differences of 0.39 mmol/L (15 mg/dL), 34 patients would need to complete the two treatment periods (α statistic, 0.05; power > 0.8). From a computerized register of clinical records, 75 hypercholesterolemic patients (35 women and 40 men) who initially met the eligibility criteria were selected for screening and were asked to participate in the study. They were offered free walnuts but no monetary compensation.

Study Design

A crossover design was used. Patients were randomly assigned to the two diet sequences by using a computer-generated random-number table, with stratification by sex. Because patients followed each diet for 6 weeks and lipoprotein changes due to dietary intervention stabilize in less than 4 weeks (12), we did not incorporate a washout period between diets. In their crossover feeding study with walnuts, Sabaté and colleagues (8) did not observe a carryover effect. In the week before the trial began, patients received expert dietary counseling individually and in a group class. Twice during the pretrial week and on weeks 5 and 6 of each one of the two dietary periods, patients came to the clinic for a medical visit, an interview with the dietitian, anthropometric measurements, and blood extraction. The main outcomes of the study were changes in serum levels of total and LDL cholesterol from the control diet period to the walnut diet period. Secondary outcomes were changes in other lipid variables and oxidizability of LDL particles. The study protocol was approved by the institutional review board of the Hospital Clínic of Barcelona, and all patients gave informed consent.

Diets

The experimental diets were individually prescribed and were based on estimated energy requirements. Because participants ate on their own, detailed dietary information was provided to them and, if appropriate, to their partners. Diets were calculated in increments of 200 kcal to cover the range from 1600 to 2200 kcal. The control diet was Mediterranean and was composed of natural foodstuffs. Red meat and eggs were limited, vegetable products and fish were emphasized, olive oil was indicated for culinary use, and no nuts were allowed. The walnut diet was similar to the control diet, but walnuts partially replaced olive oil and other fatty foods. Prepackaged daily allowances of raw, shelled walnuts were provided daily in amounts varying from 41 g to 56 g (the equivalent of 8 to 11 walnuts), according to the participants' total energy intake. Walnuts were consumed as snacks or with meals in desserts or salads. In the walnut diet, walnuts contributed approximately 18% of the total energy and 35% of the total fat. To improve compliance, each family unit was given 1000-g packs of walnuts at the beginning of the walnut diet period.

Adherence to the study diets was carefully monitored. Unannounced 24-hour diet recalls were performed weekly by telephone during the two dietary periods, for a total of 12 recalls per patient. This method allows reliable estimations of food intake (13). The nutrient composition of the diets was calculated with Food Processor Plus software, version 5.0 (ESHA Research, Salem, Oregon), which was adapted to nutrient databases of specific Mediterranean foods when appropriate. We defined *noncompliance* as at least 20% deviation from dietary instructions regarding walnut or nutrient intake. Compliance during the walnut diet was also assessed at each clinic visit by a count of the empty walnut packages. The fatty acid content of LDL lipids was analyzed as a biological measure of adherence to the prescribed diets.

Laboratory Measurements

Blood samples were obtained after an overnight fast, and serum and EDTA plasma were collected and processed immediately. Serum lipid and apolipoprotein levels were determined as described else-

where (11). In brief, serum cholesterol and triglyceride levels were measured by using enzymatic reagents (Trinder, Bayer Diagnostics, Tarrytown, New York) adapted to a Cobas Mira automated analyzer (Hoffmann-LaRoche, Basel, Switzerland). High-density lipoprotein (HDL) cholesterol was quantified after precipitation with phosphotungstic acid and magnesium chloride. Levels of apolipoprotein A-I and apolipoprotein B were determined by an immunoturbidimetric method (Unimate 3, Roche Diagnostic Systems, Basel, Switzerland). Measurements of the cholesterol content of very-low-density lipoprotein particles and LDL particles were based on preparative ultracentrifugation (11). Lipoprotein(a) levels were measured by using an immunoturbidimetric method (Lipoprotein(a) SPQ II Test Kit, DiaSorin, Stillwater, Minnesota).

To obtain the LDL fraction for fatty acid and oxidizability studies, lipoproteins were fractionated by sequential centrifugation adjusting with NaBr to separate very-low-density lipoprotein particles (d = 1.006 g/mL) and to obtain LDL particles (d = 1.063 g/mL), as described elsewhere (14). Low-density lipoprotein fatty acids were measured by capillary gas chromatography in the cholesteryl ester, phospholipid, and triglyceride lipid fractions (14). Low-density lipoprotein susceptibility to oxidation was determined by measuring the α-tocopherol content of LDL particles and conjugated diene kinetics after incubation of 50 μg of LDL protein with 5 μmol of copper sulfate at 37 °C (15, 16).

Statistical Analysis

The two measurements obtained at baseline and at the end of each dietary period were averaged. Means and SDs are presented for each measurement. With methods described by Fleiss (17), two-tailed *t*-tests were used to compare changes in outcome variables in response to dietary treatment and diet period and carryover effects for the two-period crossover design. Differences between the walnut and control diets were also tested by analysis of covariance using general linear models; baseline values or sex were used as covariates. Analyses were performed by using SAS software (SAS, Inc., Cary, North Carolina) (18).

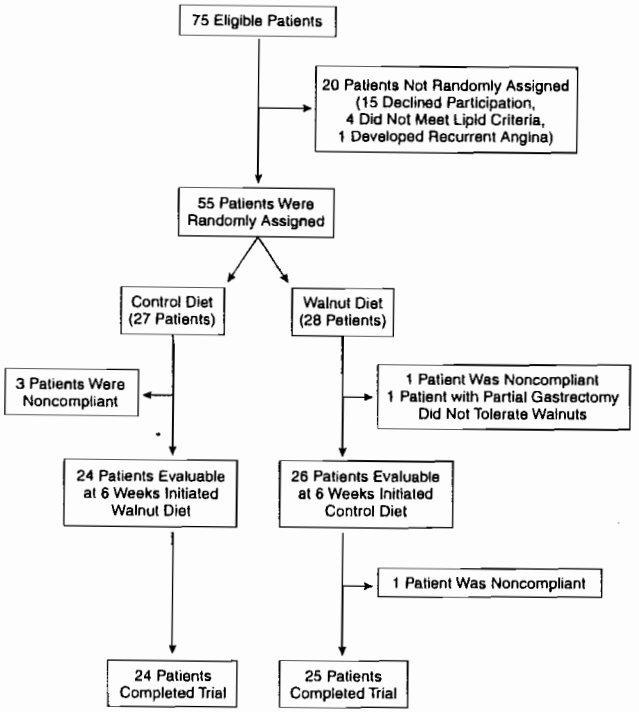


Figure 1. Flow of patients.

Table 1. Baseline Characteristics of Study Patients

Variable	Women (n = 27)		Men (n = 28)		All Patients (n = 55)
	Mean ± SD	Range	Mean ± SD	Range	Mean ± SD
Age, *y*	59 ± 8	42–68	53 ± 12	28–72	56 ± 11
Weight, *kg*	65.0 ± 8.9	47–79	75.2 ± 12.7	55–101	70.6 ± 12.1
Body mass index, *kg/m²*	26.3 ± 3.4	19–34	27.5 ± 2.9	20–33	27.0 ± 3.1
Waist-to-hip ratio	0.88 ± 0.09	0.75–1.05	0.95 ± 0.04	0.85–1.01	0.91 ± 0.07
Blood pressure, *mm Hg*					
Systolic	124 ± 20	95–160	123 ± 16	90–160	123 ± 18
Diastolic	81 ± 12	60–100	77 ± 9.2	55–95	79 ± 11
Serum lipids, *mmol/L (mg/dL)*					
Total cholesterol level	7.42 ± 0.75 (287 ± 29)	6.05–8.79 (234–340)	6.98 ± 0.88 (270 ± 34)	5.74–8.61 (222–333)	7.21 ± 0.83 (279 ± 32)
Low-density lipoprotein cholesterol level	5.22 ± 0.70 (202 ± 27)	4.09–6.47 (158–250)	4.78 ± 0.85 (185 ± 33)	3.36–6.52 (130–252)	5.05 ± 0.75 (195 ± 29)
High-density lipoprotein cholesterol level	1.60 ± 0.34 (62 ± 13)	1.01–2.22 (39–86)	1.32 ± 0.28 (51 ± 11)	0.67–1.91 (26–74)	1.44 ± 0.31 (56 ± 12)
Triglyceride level	1.42 ± 0.51 (126 ± 45)	0.73–2.66 (65–236)	1.56 ± 0.41 (138 ± 36)	0.89–2.81 (79–249)	1.51 ± 0.43 (134 ± 38)

Role of the Funding Sources

The California Walnut Commission provided funding and walnuts. Research grants were also obtained from national and local nonprofit agencies. The funding sources were not involved in the design of the study and had no role in the collection, analysis, or interpretation of the data or in the decision to submit the manuscript for publication. None of the authors has any financial interest in the nut food industry.

Results

Patient Characteristics

Of the 75 eligible patients, 20 left the study before randomization for various reasons (**Figure 1**). **Table 1** shows the baseline characteristics of the 55 patients (27 women and 28 men) who met all of the eligibility requirements, entered the study, and were randomly assigned to one of two dietary intervention sequences. Twenty-two women were postmenopausal. Nine patients (6 men and 3 women) had high blood pressure that was controlled with antihypertensive medication. Seven patients (5 men and 2 women) had coronary heart disease.

As shown in **Figure 1**, 6 patients withdrew before completing the two phases of the study. The baseline serum lipid profiles of the patients who were excluded for noncompliance were similar to the mean values of the overall group. Forty-nine participants (23 women and 26 men) completed both phases of the study. Subsequent data refer only to this group.

Dietary Compliance and Body Weight

The nutrient content of the self-reported diets was close to that of the planned diets (**Table 2**). The fatty acid composition of the control diet reflected the high monounsaturated fatty acid content of olive oil, and the fatty acid composition of the walnut diet mirrored the constitution of walnut fat, which is particularly rich in polyunsaturated fatty acids. Other nutrients had differences that were small but statistically significant. The significance can be explained by the high statistical power of the study, the use of calorie-adjusted nutrient values, and the participants' close adherence to the prescribed diets; all of these factors resulted in small SDs. According to participants' reports and to recounts of empty packages, compliance with walnut ingestion was 100%. The analysis of LDL cholesteryl ester fatty acids during the two dietary periods confirmed that participants had closely adhered to the prescribed diets. When the treatment effect of the walnut diet was compared with that of the control diet, molar percentages of oleic acid decreased 3.9 (CI, −4.6 to −3.2), molar percentages of linoleic acid increased 6.6 (CI, 8.0 to 5.0), and molar percentages of α-linolenic acid increased 0.34 (CI, 0.44 to 0.25) ($P <$ 0.001 for all comparisons). The respective mean changes of molar percentages of oleic acid, linoleic acid, and α-linolenic acid were −20%, 14%, and 83%. Similar changes occurred in LDL phospholipids and triglycerides.

Body weight was stable throughout the two intervention diet periods (**Table 3**). Daily walnut consumption was well tolerated by most patients. Twenty-five patients reported softening of the stools associated with walnut consumption. After walnut ingestion, 3 patient described mild symptoms of postprandial heaviness and bloating; however, these symptoms did not lead to withdrawal.

Table 2. Composition of the Prescribed and Observed Study Diets

Variable	Control Diet		Walnut Diet		*P* Value for Comparisons between Actual Diets
	Prescribed	Actual*	Prescribed	Actual*	
Energy, *kcal/d*	1600–2200	1771 ± 152	1600–2200	1824 ± 178	0.116
Energy derived from fat, %	30.2	31.2 ± 1.2	32.7	33.2 ± 1.3	<0.001
Saturated fatty acids	5.8	6.9 ± 0.7	5.0	6.0 ± 0.7	>0.2
Monounsaturated fatty acids	17.6	17.5 ± 1.1	13.2	13.5 ± 0.6	<0.001
Polyunsaturated fatty acids	4.2	4.8 ± 0.5	11.8	11.7 ± 0.8	<0.001
Linoleic acid (C18:2n-6)	3.3	3.8 ± 0.4	9.6	9.5 ± 0.6	<0.001
α-Linolenic acid (C18:3n-3)	0.4	0.4 ± 0.1	1.9	1.8 ± 0.1	<0.001
Energy derived from protein, %	18.1	19.0 ± 1.1	16.6	17.9 ± 1.1	<0.001
Energy derived from carbohydrates, %	51.7	49.8 ± 2.1	50.7	48.0 ± 1.9	<0.001
Cholesterol, *mg/1000 kcal*	103.9	124.8 ± 24.2	77.8	90.8 ± 17.0	<0.001
Soluble fiber, *g/1000 kcal*	4.8	4.7 ± 0.8	5.0	4.7 ± 0.8	>0.2
Vitamin E (total), *mg/1000 kcal*	8.2	7.1 ± 0.8	8.9	9.9 ± 0.8	<0.001
α-Tocopherol, *mg/1000 kcal*	5.8	5.8 ± 0.6	5.1	4.7 ± 0.6	<0.001
Vitamin C, *mg/1000 kcal*	76.2	97.8 ± 28.8	76.3	83.0 ± 23.9	0.007

* Values are the mean ± SD and were estimated from one 3-day food record and six 24-hour diet recalls during each diet period.

Effects on Serum Lipids and Lipoproteins

Figure 2 shows the changes from baseline values in serum lipids, lipoproteins, and apolipoproteins A-I and B. The actual values at baseline and at the end of each dietary period as well as the differences of effect between dietary interventions are presented in **Table 3**. No carryover effect was seen between the periods. The mean total cholesterol level decreased by 9.0% (0.65 mmol/L [25 mg/dL]) during the walnut diet and by 5.0% (0.36 mmol/L [14 mg/dL]) during the control diet. Similarly, the mean LDL cholesterol level decreased by 11.2% (0.57 mmol/L [22 mg/dL]) during the walnut diet and by 5.6% (0.28 mmol/L [11 mg/dL]) during the control diet. The two diets did not differ with respect to their effects on levels of HDL cholesterol, very-low-density lipoprotein cholesterol, triglycerides, or apolipoprotein A-I. Apolipoprotein B levels decreased after the two diets in parallel with LDL cholesterol levels (**Table 3**). The mean ratio of LDL cholesterol to HDL cholesterol did not change during the control diet and decreased by 8% during the walnut diet. The differences between the effects of the diets on the lipid profile did not change materially when they were adjusted for baseline values or sex by analysis of covariance.

Table 3 and **Figure 2** show the effects of the two diets on lipoprotein(a) levels. Lipoprotein(a) levels decreased by 9.1% (0.033 g/L) during the walnut diet and 3.4% (0.012 g/L) during the control diet. The difference in lipoprotein(a) reduction between the two diets was statistically significant in men ($P = 0.041$) but not in women ($P > 0.2$). In addition, this difference was statistically significant in patients with baseline lipoprotein(a) levels less than or equal to 0.3 g/L ($P = 0.042$) but not in those with baseline levels greater than 0.3 g/L ($P > 0.2$).

Low-Density Lipoprotein Oxidation

The α-tocopherol content of the LDL particles and the lag time of conjugate diene formation during copper-induced LDL oxidation were similar during the control and walnut diets (**Table 3**).

Table 3. Serum Lipid and Lipoprotein Levels, Analytes Related to Low-Density Lipoprotein Oxidation, and Body Weight at the End of Each Diet Period*

Variable	Mean Baseline Measurements ± SD	Mean Measurements during Control Diet ± SD	Mean Measurements during Walnut Diet ± SD
Total cholesterol level, *mmol/L (mg/dL)*	7.16 ± 0.85 (278 ± 33)	6.81 ± 0.79 (2.64 ± 31)	6.52 ± 0.90 (253 ± 35)
LDL cholesterol level, *mmol/L (mg/dL)*	5.05 ± 0.77 (196 ± 30)	4.77 ± 0.64 (185 ± 25)	4.48 ± 0.77 (174 ± 30)
HDL cholesterol level, *mmol/L (mg/dL)*	1.44 ± 0.33 (56 ± 13)	1.37 ± 0.31 (53 ± 12)	1.42 ± 0.36 (55 ± 14)
VLDL cholesterol level, *mmol/L (mg/dL)*	0.67 ± 0.33 (26 ± 13)	0.64 ± 0.36 (25 ± 14)	0.59 ± 0.33 (23 ± 13)
Triglyceride level, *mmol/L (mg/dL)*	1.54 ± 0.48 (136 ± 43)	1.51 ± 0.50 (134 ± 44)	1.42 ± 0.50 (126 ± 44)
Apolipoprotein A-I level, *g/L*	1.70 ± 0.24	1.61 ± 0.21	1.62 ± 0.23
Apolipoprotein B level, *g/L*	1.65 ± 0.23	1.52 ± 0.21	1.44 ± 0.22
LDL HDL ratio	3.6 ± 1.1	3.7 ± 1.0	3.4 ± 1.0
Lipoprotein(a) level, *g/L*	0.35 ± 0.24	0.34 ± 0.24	0.32 ± 0.22
α-Tocopherol level, *μg/mg of LDL protein*	9.2 ± 1.5	7.2 ± 2.2	7.5 ± 2.4
Lag time of conjugated diene production, *min*	41.9 ± 6.5	42.0 ± 6.8	40.6 ± 6.0
Body weight, *kg*	70.6 ± 12.1	70.1 ± 12.3	69.9 ± 12.5

* HDL = high-density lipoprotein, LDL = low-density lipoprotein, VLDL = very-low-density lipoprotein.
† Average differences between control and walnut diets were calculated by the method described by Fleiss (17).
‡ Walnut diet compared with control diet.

Discussion

In this 12-week crossover dietary intervention trial in 49 free-living, hypercholesterolemic men and women, we found that substituting walnuts for approximately 35% of the energy from fat without changing saturated fatty acid intake in a cholesterol-lowering Mediterranean diet further decreased total cholesterol and LDL cholesterol levels. Use of walnuts did not affect HDL cholesterol levels and thereby improved the ratio of LDL cholesterol to HDL cholesterol. Our findings are consistent with those of an earlier study in normal young men (8) and extend the results of that study to older persons of either sex with elevated blood cholesterol levels. In addition, apolipoprotein B levels decreased after the walnut diet in parallel with LDL cholesterol levels, and lipoprotein(a) levels decreased in men and in patients whose baseline levels were greater than 0.3 g/L. Furthermore, LDL particles were enriched with polyunsaturated fatty acids from walnuts but preserved their resistance to oxidation.

The design of the study—an outpatient crossover feeding trial—presented difficulties in ensuring compliance. However, these difficulties were partially offset by detailed dietary instructions, regular reinforcement, and frequent 24-hour diet recalls performed weekly by telephone throughout the trial. Manageable daily allowances of walnuts and their distribution into different recipes also facilitated compliance during the walnut phase of the study. In fact, compliance was very good and the actual diets consumed closely matched the prescribed diets. The observed changes in the fatty acid composition of LDL lipids are congruent with this assertion. Besides the predictable differences in monounsaturated fatty acid and polyunsaturated fatty acid content, the caloric value and nutrient composition of the two diets were similar (**Table 2**). The patients consumed slightly less cholesterol during the walnut diet than during the control diet, a small difference that only partially explains the extent of the walnut diet's cholesterol-lowering effect (19). Because body weight was not modified, our findings cannot readily be attributed to changes in body weight or caloric intake. In a recent 8-week intervention study with a lower degree of dietary control, a diet containing substantial quantities of walnuts had no hypolipidemic effect when compared with a standard low-fat diet in 21 hypercholesterolemic men (20).

Substantial epidemiologic and experimental evidence indicate that a diet high in saturated fatty acids from animal fat is associated with elevated levels of blood cholesterol, which in turn are related to high incidence of coronary heart disease (1). It is also widely acknowledged that cholesterol levels can be reduced if saturated fat in the diet is replaced by unsaturated fatty acids from vegetable oils and fats; this effect can be predicted by the amounts of fatty acid classes exchanged (2–4). Recent meta-analyses of studies comparing dietary monounsaturated and polyunsaturated fatty acids suggest that their effects on serum lipids are similar when 4% to 10% of the energy of each fatty acid class is substituted for the other (2–4). To ascertain the extent to which the different fat content of the two diets could explain the observed reduction in LDL cholesterol during the walnut diet, we applied the recently developed predictive models that include the regression coefficients for percentage energy changes in dietary saturated and polyunsaturated fatty acids, monounsaturated fatty acids (2, 3), and dietary cholesterol (4). These equations predicted decreases ranging from 0.13 to 0.18 mmol/L (5 to 7 mg/dL), which are smaller than the average observed decrease of 0.29 mmol/L (11 mg/dL) but close to the lower 95% CI of the actual change (**Table 3**). Therefore, the hypolipidemic effect of the walnut diet can be explained in part by its fat content. Most comparative studies of the effects of unsaturated fats on serum lipid levels have used plant and vegetable oils as sources of monounsaturated fatty acids (oleic acid) and polyunsaturated fatty acids (mainly linoleic acid). However, walnuts are solid, complex foods that may influence blood lipids by the nonfat components of their matrix; this could be investigated in a feeding trial that compares the lipid effects of whole walnuts with those of walnut oil.

Lipoprotein(a) is an apolipoprotein B–containing lipoprotein that has been found to be a determinant of the risk for coronary heart disease in clinical and epidemiologic studies (21–24). Evidence suggests that it is also an independent predictor of nonfatal myocardial infarction and coronary death in hypercholesterolemic men (25). Almost all cross-sectional and retrospective studies have shown an increase in

Table 3—*Continued*

Treatment Effect (95% CI)†	Percentage Change‡	*P* Value for Comparisons between Diets
−0.28 (−0.43 to −0.12) (−10.8 [−16.8 to −4.8])	−4.1	<0.001
−0.29 (−0.41 to −0.15) (−11.2 [−16.3 to −6.1])	−5.9	<0.001
0.04 (0.10 to −0.01) (1.6 [3.9 to −0.5])	3.2	0.134
−0.05 (−0.12 to 0.03) (−1.8 [−4.6 to 1.0])	−7.2	>0.2
−0.09 (−0.20 to 0.02) (−8 [−18 to −2])	−6.1	0.103
0.01 (0.05 to −0.03)	0.6	0.103
−0.07 (−0.11 to −0.03)	−4.7	<0.001
−0.3 (−0.5 to −0.1)	−8.1	<0.001
−0.02 (−0.042 to −0.001)	−6.2	0.042
0.3 (−1.2 to 1.8)	4.0	>0.2
−1.54 (−3.40 to 0.32)	−3.3	0.101
−0.2 (−0.4 to 0.0)	−0.3	0.07

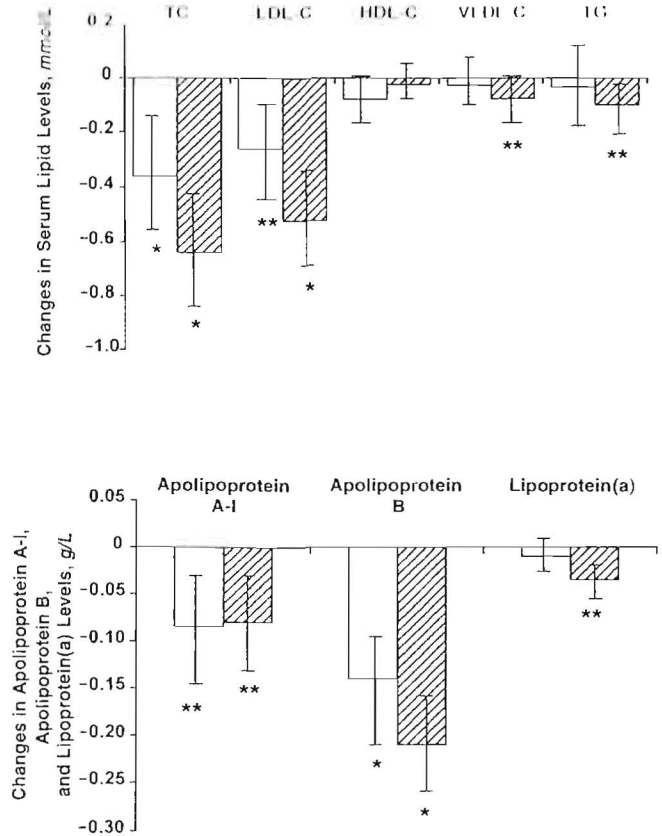


Figure 2. Changes from baseline values in levels of serum lipids, lipoproteins, and apolipoproteins in all patients who completed the study. **Top.** Mean changes from baseline in serum lipid levels. **Bottom.** Mean changes from baseline in levels of apolipoprotein A-I, apolipoprotein B, and lipoprotein(a). White bars indicate the control diet; striped bars indicate the walnut diet. HDL-C = high-density lipoprotein cholesterol; LDL-C = low-density lipoprotein cholesterol; TC = total cholesterol; TG = triglyceride; VLDL-C = very-low-density lipoprotein cholesterol. Error bars represent 95% CIs. To convert triglyceride values to mg/dL, divide by 0.01129, to convert high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, total cholesterol, and very-low-density lipoprotein cholesterol values to mg/dL, divide by 0.02586. Asterisks indicate significant differences from baseline by the two-tailed *t*-test. *$P < 0.001$, **$P < 0.05$.

cardiovascular risk associated with plasma levels of lipoprotein(a) greater than the 80th percentile (>0.3 g/L) (21). Blood concentrations are primarily controlled by genetics, and it is generally believed that dietary factors have a negligible effect. The walnut diet modestly but significantly reduced lipoprotein(a) levels in men (but not in women as a group) and in patients with baseline serum levels greater than 0.3 g/L. This is interesting in view of the observation that pharmacologic doses of n-3 polyunsaturated fatty acids from fish oils might have a beneficial effect on lipoprotein(a) levels (26–28). The effect of n-3 fatty acids from plant sources on lipoprotein(a) requires further study.

Dietary fatty acid composition largely determines the fatty acid composition of serum lipoproteins, which in turn influences the rate and extent of their oxidation (10). Unlike saturated and monounsaturated fatty acids, polyunsaturated fatty acids are susceptible to oxidation (10). Nevertheless, enrichment of LDL particles with polyunsaturated fatty acids during the walnut diet did not alter their resistance to oxidative damage, as assessed by α-tocopherol content (the principal protection of LDL particles against oxidation) and the lag time of conjugated diene formation (the only measure of LDL oxidation that has been associated with coronary heart disease in clinical studies [29, 30]). Other components of walnuts that have antioxidant potential, such as α-tocopherol and other phytochemicals, may be responsible for the lack of change in LDL oxidizability. Nonetheless, the fact that walnut intake is associated with a reduced risk for coronary heart disease (7) argues against an atherogenic effect caused by altered susceptibility of lipoproteins to oxidative damage.

Besides improving the serum lipid profile, walnut consumption may offer additional cardiovascular protection. Among naturally occurring foods, walnuts are one of the highest sources of the n-3 fatty acid α-linolenic acid; they contain approximately 7 g of α-linolenic acid per 100 g of edible matter (6). In recent reports from large prospective studies, dietary intake of α-linolenic acid has been inversely associated with risk for fatal coronary heart disease (31–33). In a secondary prevention trial, a Mediterranean diet enriched with α-linolenic acid had a striking beneficial effect on coronary heart disease morbidity and mortality when compared with a prudent western-type diet (34, 35). The low ratio of linoleic acid to α-linolenic acid in that study (4.5:1) was similar to that seen during the walnut diet in our study (5:1). Dietary α-linolenic acid may reduce the risk for death from coronary heart disease because of its antiarrhythmic properties (36) and other antiatherogenic effects (37, 38).

Our investigation adds further weight to the accumulating evidence that regular intake of nuts has a cholesterol-lowering effect (5). Since the first epidemiologic evidence that frequent nut consumption reduces the risk for coronary heart disease was reported (7), two large prospective studies have confirmed the observation (39, 40). It is reasonable to assume that cardiovascular protection is due, at least in part, to improvement of the lipid profile. The proportion of saturated fat in western diets is generally higher than that in Mediterranean diets. If the lipid profile is improved by partially substituting walnuts for typical Mediterranean foods and oils, which are rich in monounsaturated fatty acids and low in saturated fat, greater benefits might be obtained by partially substituting walnuts for traditional western dietary fats.

From Hospital Clínic and School of Pharmacy, University of Barcelona, Barcelona, Spain; and Loma Linda University, Loma Linda, California.

Acknowledgments: The authors thank Nahyr Schinca, RD, and Ana Pérez-Heras, RD, for dietary counseling of the participants in the study. Catherine Bouchet, Antonia Codinach, and Ana Asensio provided skillful technical assistance with the laboratory measurements.

Grant Support: By the California Walnut Commission, Fundació Privada Catalana de Nutrició i Lípids, Fondo de Investigaciónes Sanitarias de la Seguridad Social (FIS 94/0077), and Comisión Interministerial de Ciencia y Tecnología (CICYT, SAF 97/0215, and OLI 96/2132).

Requests for Single Reprints: Emilio Ros, MD, Lipid Clinic, Nutrition & Dietetics Service, Hospital Clínic i Provincial, Villarroel 170, E-08036 Barcelona, Spain.

Requests To Purchase Bulk Reprints (minimum, 100 copies): Barbara Hudson, Reprints Coordinator; phone, 215-351-2657; e-mail, bhudson@mail.acponline.org.

Current Author Addresses: Drs. Zambón and Ros and Ms. Campero: Lipid Clinic, Nutrition & Dietetics Service, Hospital Clínic i Provincial, Villarroel 170, E-08036 Barcelona, Spain.
Dr. Sabaté: Department of Nutrition, School of Public Health, Loma Linda University, Loma Linda, CA 92350.
Drs. Muñoz, Merlos, and Laguna: Pharmacology Unit, School of Pharmacy, University of Barcelona, E-08028 Barcelona, Spain.
Dr. Casals: Department of Clinical Biochemistry, Hospital Clínic i Provincial, Villarroel 170, E-08036 Barcelona, Spain.

Author Contributions: Conception and design: D. Zambón, J. Sabaté, E. Ros.
Analysis and interpretation of the data: D. Zambón, J. Sabaté, J.C. Laguna, E. Ros.
Drafting of the article: D. Zambón, J.C. Laguna, E. Ros.
Critical revision of the article for important intellectual content: D. Zambón, J. Sabaté, J.C. Laguna, E. Ros.
Final approval of the article: D. Zambón, J. Sabaté, S. Muñoz, B. Campero, E. Casals, M. Merlos, J.C. Laguna, E. Ros.
Provision of study materials or patients: D. Zambón, B. Campero.
Statistical expertise: D. Zambón, J. Sabaté.
Obtaining of funding: D. Zambón, J. Sabaté, J.C. Laguna, E. Ros.
Administrative, technical, or logistic support: S. Muñoz, E. Casals.
Collection and assembly of data: S. Muñoz, B. Campero, E. Casals, M. Merlos.

References

1. **Ulbricht TL, Southgate DA.** Coronary heart disease: seven dietary factors. Lancet. 1991;338:985-92.
2. **Mensink RP, Katan MB.** Effect of dietary fatty acids on serum lipids and lipoproteins. A meta-analysis of 27 trials. Arterioscler Thromb. 1992;12:911-9.
3. **Yu S, Derr J, Etherton TD, Kris-Etherton PM.** Plasma cholesterol-predictive equations demonstrate that stearic acid is neutral and monounsaturated fatty acids are hypocholesterolemic. Am J Clin Nutr. 1995;61:1129-39.
4. **Clarke R, Frost C, Collins R, Appleby P, Peto R.** Dietary lipids and blood cholesterol: quantitative meta-analysis of metabolic ward studies. BMJ. 1997;314:112-7.
5. **Sabaté J, Fraser GE.** Nuts: a new protective food against coronary heart disease. Curr Opin Lipidol. 1994;5:11-6.
6. **Exler J, Weihrauch JL.** Provisional table on the content of omega-3 fatty acids and other fat components in selected foods. Washington, DC: U.S. Department of Agriculture, 1986 (publication HNIS/P1-103).
7. **Fraser GE, Sabaté J, Beeson WL, Strahan TM.** A possible protective effect of nut consumption on risk of coronary heart disease. The Adventist Health Study. Arch Intern Med. 1992;152:1416-24.
8. **Sabaté J, Fraser GE, Burke K, Knutsen SF, Bennett H, Lindsted KD.** Effects of walnuts on serum lipid levels and blood pressure in normal men. N Engl J Med. 1993;328:603-7.
9. **Steinberg D, Parthasarathy S, Carew TE, Khoo JC, Witztum JL.** Beyond cholesterol. Modifications of low-density lipoprotein that increase its atherogenicity. N Engl J Med. 1989;320:915-24.
10. **Tsimikas S, Reaven PD.** The role of dietary fatty acids in lipoprotein oxidation and atherosclerosis. Curr Opin Lipidol. 1998;9:301-7.
11. **Zambón D, Ros E, Casals E, Sanllehy C, Bertomeu A, Campero I.** Effect of apolipoprotein E polymorphism on the serum lipid response to a hypolipidemic diet rich in monounsaturated fatty acids in patients with hypercholesterolemia and combined hyperlipidemia. Am J Clin Nutr. 1995;61:141-8.
12. **Kris-Etherton PM, Dietschy J.** Design criteria for studies examining individual fatty acid effects on cardiovascular disease risk factors: human and animal studies. Am J Clin Nutr. 1997;65(Suppl):1595S-6S.
13. **Buzzard J.** 24-hour dietary recall and food record methods. In: Willett W, ed. Nutritional Epidemiology. 2d ed. New York: Oxford Univ Pr; 1998:50-73.
14. **Vázquez M, Alegret M, Adzet T, Merlos M, Laguna JC.** Gemfibrozil modifies acyl composition of liver microsomal phospholipids from guinea-pigs without promoting peroxisomal proliferation. Biochem Pharmacol. 1993;46:1515-8.
15. **Kleinveld HA, Hak-Lemmers HL, Stalenhoef AF, Demacker PN.**

(65) [1447]

Cardiovascular Effects of Sildenafil During Exercise in Men With Known or Probable Coronary Artery Disease

A Randomized Crossover Trial

Adelaide M. Arruda-Olson, MD, PhD
Douglas W. Mahoney, MS
Ajay Nehra, MD
Marilyn Leckel, RN
Patricia A. Pellikka, MD

Context The relationship between sildenafil citrate use and reported adverse cardiovascular events in men with coronary artery disease (CAD) is unclear.

Objective To evaluate the cardiovascular effects of sildenafil during exercise in men with CAD.

Design, Setting, and Subjects Randomized, double-blind, placebo-controlled crossover trial conducted March to October 2000 at a US ambulatory-care referral center among 105 men with a mean (SD) age of 66 (9) years who had erectile dysfunction and known or highly suspected CAD.

Interventions All patients underwent 2 symptom-limited supine bicycle echocardiograms separated by an interval of 1 to 3 days after receiving a single dose of sildenafil (50 or 100 mg) or placebo 1 hour before each exercise test.

Main Outcome Measures Hemodynamic effects of sildenafil during exercise (onset, extent, and severity of ischemia) assessed by exercise echocardiography.

Results Mean (SD) resting ejection fraction was 56% (7%) (range, 39%-68%). After sildenafil use, resting systolic blood pressure was reduced from 135 (19) mm Hg to 128 (17) mm Hg, for a mean change of −7 mm Hg (95% confidence interval [CI], −9 to −4 mm Hg; $P<.001$). After placebo use, the mean (SD) change was from 135 (20) mm Hg to 133 (19) mm Hg, a difference of −2 mm Hg (95% CI, −6 to 0.3 mm Hg; $P=.08$). The difference between mean change after sildenafil and placebo use was 4.3 (95% CI, 0.9-7.7; $P=.01$). Resting heart rate, diastolic blood pressure, and wall motion score index (a measure of the extent and severity of wall motion abnormalities) did not change significantly in either group. Exercise capacity was similar with sildenafil use (mean [SD], 4.5 [1.0] metabolic equivalents) and placebo use (mean [SD], 4.6 [1.0] metabolic equivalents; mean difference, 0.07; 95% CI, −0.06 to 0.19; $P=.29$). Exercise blood pressure and heart rate increments were similar. Dyspnea or angina developed in 69 patients who took sildenafil and 70 patients who took placebo ($P=.89$); exercise electrocardiography was positive in 12 patients (11%) who took sildenafil and 17 patients (16%) who took placebo ($P=.09$). Exercise-induced wall motion abnormalities developed in similar numbers of patients after sildenafil and placebo use (84 and 86 patients, respectively; $P=.53$). Wall motion score index at peak exercise was similar after sildenafil and placebo use (mean [SD], 1.4 [0.4] vs 1.4 [0.4]; mean difference, 0.01; 95% CI, −0.01 to 0.03; $P=.40$).

Conclusion In men with stable CAD, sildenafil had no effect on symptoms, exercise duration, or presence or extent of exercise-induced ischemia, as assessed by exercise echocardiography.

JAMA. 2002;287:719-725 www.jama.com

Author Affiliations: Division of Cardiovascular Diseases and Internal Medicine (Drs Arruda-Olson and Pellikka and Ms Leckel) and Departments of Biostatistics (Mr Mahoney) and Urology (Dr Nehra), Mayo Clinic and Mayo Foundation, Rochester, Minn.

Corresponding Author and Reprints: Patricia A. Pellikka, MD, Division of Cardiovascular Diseases and Internal Medicine, Mayo Clinic, 200 First St SW, Rochester, MN 55905 (e-mail: pellikka.patricia@mayo.edu).

For editorial comment see p 766.

ERECTILE DYSFUNCTION AFFECTS 30 million men in the United States[1] and frequently coexists with coronary artery disease. Since the Food and Drug Administration approved the use of sildenafil citrate for the treatment of erectile dysfunction, millions of prescriptions have been issued.[2] Reported adverse cardiac events associated with sildenafil use include acute myocardial infarction, ventricular tachycardia, hypotension, and death,[3-7] raising concerns about the safety of this agent in patients with coronary artery disease.

Sildenafil is a cyclic guanosine monophosphate–specific type 5 phosphodiesterase inhibitor. Phosphodiesterase 5 is located not only in the corpus cavernosum, but also in other vascular tissue, including arteries and veins.[8] Adverse cardiovascular events associated with sildenafil may be due to myocardial ischemia during sexual activity, with aggravation of ischemia by a vasodilator effect. Published guidelines regarding the management of cardiac patients with erectile dysfunction suggest that sildenafil may be hazardous in patients with ischemic heart disease and that clinicians should use caution in prescribing this medication.[7]

However, it is also possible that the concern regarding adverse events associated with sildenafil may instead stem from the cardiovascular demands of sexual activity, the health of the population for whom sildenafil is prescribed, an adverse interaction with nitrates, or reporting bias. Hence, the purpose of this study was to assess the hemodynamic effects of sildenafil during exercise, including the effect on the onset, extent, and severity of electrocardiographic and echocardiographic evidence of ischemia in men with known coronary artery disease or high pretest probability of coronary artery disease.

METHODS

All study subjects were men older than 40 years with erectile dysfunction and either known coronary artery disease (≥50% diameter stenosis of a major epicardial vessel or one of its major branches, history of myocardial infarction, prior positive stress imaging test result, or prior coronary artery bypass surgery or angioplasty) or a high (>70%) pretest probability of coronary artery disease, according to the presence of typical angina pectoris.[8] Subjects were recruited via posted advertisements, Mayo Clinic newsletter announcements, and physician referrals; in all cases, primary care physicians agreed to the subjects' participation. All subjects had adequate resting echocardiographic images, were able to exercise, and agreed to participate in the study. No subject had asthma, severe aortic stenosis, hypertrophic obstructive cardiomyopathy, unstable angina, recent myocardial infarction (within 1 month), significant arrhythmia or atrial fibrillation, congestive heart failure, hepatic insufficiency, renal insufficiency, or a systolic blood pressure less than 90 mm Hg. No subjects were receiving therapy with dipyridamole, theophylline, erythromycin, or cimetidine, nor had they used sildenafil within the previous 24 hours. The use of long-acting nitrates was discontinued at least 72 hours before testing; the use of other cardioactive medications was continued. The sildenafil citrate dose was 50 mg, unless another dose was recommended by the man's physician. The study was approved by the institutional review board of the Mayo Clinic, Rochester, Minn. Written informed consent was obtained from all subjects.

Exercise Echocardiography

All patients underwent 2 symptom-limited exercise echocardiograms separated by an interval of 1 to 3 days. Cardioactive medications were not changed between the 2 tests. Subjects were randomized in a double-blind crossover design to receive a single dose of sildenafil or placebo 1 hour before the exercise test. The order of administration was determined by a randomization schedule in blocks of 10, generated within the Section of Biostatistics, Mayo Clinic, so that half of the study population underwent the initial test after receiving sildenafil and half after receiving placebo. Sildenafil and placebo preparations, identical in appearance, were prepared in the institution's pharmacy and labeled "first test" and "second test" for each study, according to the randomization schedule. Unblinding was performed only after the database was closed.

Baseline echocardiographic images (parasternal long-axis and short-axis views and apical 4-chamber and 2-chamber views) were obtained and repeated 1 hour after the administration of sildenafil or placebo. The exercise echocardiogram was performed with a supine bicycle (Medical Positioning, Kansas City, Mo) attached to a table tilted 30° to 45° to the left. Subjects began exercising at 25 W, with a 25-W increase in the resistance at 2-minute stages. Workload in metabolic equivalent tasks (METs) was calculated with a standard equation for ergometer exercise.[9] Echocardiographic imaging was performed continuously during each stage of the exercise protocol by using an ultrasound system (Acuson Sequoia, Mountain View, Calif) with a 3-MHz transducer and harmonic-imaging mode. The study was recorded on videotape, and representative cardiac cycles were acquired, digitized, and stored for each standard view at rest, 25 W, peak exercise, and immediate recovery. The criteria for test termination were development of symptoms, including fatigue, a systolic blood pressure decrease greater than 10 mm Hg, ventricular dilation or global reduction of systolic function, or significant arrhythmia.

Baseline blood pressure and heart rate were recorded before sildenafil or placebo administration and immediately before the exercise test. During exercise, blood pressure and pulse were recorded at the end of each stage, a 12-channel electrocardiogram was obtained each minute, and 3-channel monitoring of cardiac rhythm was performed continuously.[10,11] After termination of exercise testing, subjects were monitored for 15 minutes. Oxygen via nasal cannula and intravenous esmolol hydrochloride were available for treatment of persistent symptoms and evidence of marked ischemia.

Stress Echocardiogram Interpretation

Interpretation of the echocardiographic studies was performed by a single experienced reviewer (P.A.P.) blinded to clinical information, subject identity, and results of the other stress echocardiogram. Each study was scored semiquantitatively with a 16-segment model.[12] Each segment was analyzed individually and scored by motion and systolic thickening at rest, 1 hour after medication, and with exercise. Wall motion was scored according to a 5-point grading system (1 = normal or hyperdynamic, 2 = hypokinetic, 3 = akinetic, 4 = dyskinetic, and 5 = aneurysm).[10,11] The wall motion score index (WMSI) was calculated at rest and with exercise as the sum of the scores divided by the number of segments visualized. The normal response to exercise was an increase in contractility. Myocardial ischemia was diagnosed when the exercise echocardiographic images documented a new regional wall motion abnormality or worsening of preexisting wall motion.[10,11] Resting wall motion abnormalities were classified as infarction. For assessment of the number of vascular regions with echocardiographic abnormalities, the anterior, an-

teroseptal, midinferoseptal, and apical segments were attributed to the left anterior descending coronary artery, the anterolateral and inferolateral segments to the circumflex, and the inferior and basal inferoseptal segments to the right coronary artery. The percentage of ischemic segments was determined at peak exercise. The heart rate at the onset of new or worsening wall motion abnormalities was recorded. Left ventricular ejection fraction and end systolic volume were assessed at baseline, 1 hour after administration of sildenafil or placebo, and at peak exercise by using the biplane Simpson method.[13]

Exercise electrocardiography results were considered positive for ischemia if there was horizontal or downsloping ST-segment depression of 1 mm or more at 80 milliseconds after the J point, nondiagnostic if the baseline ST segment was abnormal, or negative for ischemia in the absence of these criteria. The heart rate at which the electrocardiogram result became positive was recorded.

Statistical Analysis

Continuous variables were summarized as the mean (SD). Categorical variables were summarized as percentages. In the first step of the analysis, a test for residual carryover effects of sildenafil was conducted for each variable and none were significant (all $P>.05$).[14] For continuous variables, treatment effects were assessed by calculating the difference between first and second study data. These differences were then compared between the subjects randomized to receive placebo and then sildenafil vs sildenafil and then placebo by using the 2-sample test with corresponding 95% confidence intervals (CIs).[15] The matched-pairs paired *t* test was used to evaluate differences between premedication and postmedication data (ie, before vs after sildenafil administration and before vs after placebo administration). To evaluate differences in categorical data, variables were categorized as follows: −1, positive result in the first study and negative result in the second study; 0, positive result in both studies or negative result in both studies; and 1, negative result in the first study and positive result in the second study. The linear trends test for percentages was then used to assess significant treatment differences by comparing these categorizations between the subjects randomized to placebo and then sildenafil vs sildenafil and then placebo.[16] Odds ratios and corresponding 95% CIs for these comparisons were estimated with the method of Gart[17] and converted to relative risk estimates by using methods described by Zhang and Yu.[18]

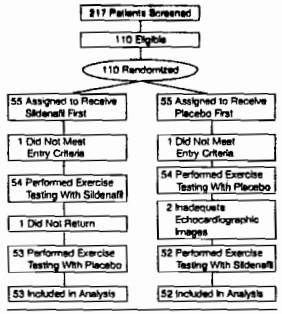
The sample size for this study was based on detectable differences of exercise WMSI between the 2 treatments. Our initial estimate of the SD for exercise WMSI for this study was 0.6 and was based on treadmill stress echocardiographic studies conducted in men who were older than 40 years and had known or high pretest probability of coronary artery disease between January 1990 and September 1998. A conservative estimate of the SD of the paired difference in exercise WMSI between the 2 examinations would be approximately 0.85 (ie, $\sqrt{2}\times 0.6$), which is considered conservative because observations on the same subject under the different treatments will be positively correlated, thus reducing the SD of paired differences. Given the estimates, this study had approximately 80% power to detect a mean increase in exercise WMSI of 0.24 between the 2 treatments, assuming a 2-sided significance level of .05 with 100 subjects (analyses performed with SAS software version 6.12 [SAS Institute Inc, Cary, NC]). In at least 3 of the 16 segments, this increase corresponds to 1 or more levels on average per subject (eg, 3 or more segments that change from hypokinesis to akinesis).

RESULTS

Study Group

From March 4, 1999, through October 4, 2000, 110 men were randomized into the study (FIGURE). Of the 105 subjects with evaluable data, the mean (SD) age was 66 (9) years (range, 43-87 years). Ninety-three (89%) had known coronary artery disease and 29 (28%) had typical angina pectoris. Ninety-seven men (92%) received 50 mg of sildenafil citrate; 8 (7%) received 100 mg. The median interval between tests was 24 hours (range, 22-77 hours). The use of long-acting nitrates was discontinued in 21 subjects (20%) 72 hours before exercise testing. Subjects' clinical characteristics are summarized in TABLE 1.

Figure. A Randomization of Study Patients



The baseline electrocardiogram result was abnormal in 59 patients (56%): 15 (14%) had previous Q-wave myocardial infarction, 42 (40%) had ST-segment abnormalities, 1 (1%) had left ventricular hypertrophy, 7 (7%) had right bundle-branch block, 1 (1%) had left bundle-branch block, and 13 (12%) had other conduction abnormalities.

Resting wall motion abnormalities were present in 60 patients (57%), whereas the mean (SD) resting WMSI was 1.2 (0.3). The mean (SD) resting ejection fraction was 56% (7%) (range, 39%-68%). A resting ejection fraction of less than 50% was present in 16 men (15%) and less than 40% in only 2 men (2%).

Blood Pressure and Heart Rate

Resting heart rate did not change significantly after sildenafil administration (mean difference, 1/min; 95% CI, −0.1 to 2.1; $P=.08$). Systolic blood pressure decreased significantly after sildenafil administration (mean [SD], 135 [19] mm Hg to 128 [17] mm Hg; mean difference, −7 mm Hg; 95% CI, −9 to −4 mm Hg; $P<.001$); this significant decrease was not observed with placebo (mean [SD], 135 [20] mm Hg to 133 [19] mm Hg; mean difference, −3 mm Hg; 95% CI, −6 to 0.3 mm Hg; $P=.08$). Diastolic blood pressure did not change significantly after sildenafil administration (mean difference, −2 mm Hg; 95% CI, −4 to 0.1 mm Hg; $P=.16$) or placebo.

Exercise hemodynamic data are summarized in TABLE 2. After exercise, the rate of recovery for heart rate and systolic and diastolic blood pressure was similar for sildenafil and placebo. The average rate of decrease in heart rate from peak exercise was 3% (1%) per minute for both sildenafil and placebo (mean difference, 0%/min; 95% CI, −0.1% to 0.1%; $P=.88$). For systolic blood pressure, the rate of decrease was 3.6 (1.5) mm Hg/min with sildenafil and 3.3 (1.5) mm Hg/min with placebo (mean difference, −0.3 mm Hg/min; 95% CI, −0.5 to 0.02; $P=.07$). Diastolic blood pressure decreased 1.0 (0.8) mm Hg/min with sildenafil and 0.9 (0.8) mm Hg/min with placebo (mean difference, −0.1 mm Hg/min; 95% CI, −0.3 to 0.1; $P=.30$).

Table 1. Clinical Profile of Study Group*

Characteristic	Patients, No. (%) Sildenafil First (n = 53)	Placebo First (n = 52)	P Value
Clinical			
Hypercholesterolemia†	35 (66)	45 (87)	.01
Hypertension†	35 (66)	25 (48)	.06
Family history of CAD	21 (40)	20 (38)	.90
Diabetes mellitus†	9 (17)	12 (23)	.43
Angina	13 (25)	16 (31)	.47
Previous CABG	18 (34)	15 (29)	.57
Previous PTCA	22 (42)	18 (35)	.47
Previous MI	20 (38)	17 (33)	.59
Medication use			
β-Adrenergic blocking agent	27 (51)	30 (58)	.49
Calcium channel blocking agent	16 (30)	15 (29)	.88
Diuretic	13 (25)	8 (12)	.08
Angiotensin-converting enzyme inhibitor or angiotensin II receptor blocker	16 (30)	17 (33)	.78
Long-acting nitrates (discontinued)	13 (25)	8 (15)	.24
α,-Adrenergic blocking agent	8 (15)	6 (12)	.59
HMG-CoA reductase inhibitor	30 (57)	33 (63)	.47
Aspirin or other antiplatelet agent	42 (79)	44 (85)	.47

*CAD indicates coronary artery disease; CABG, coronary artery bypass grafting surgery; PTCA, percutaneous transluminal coronary angiography; MI, myocardial infarction; and HMG-CoA, 3-hydroxy-3-methylglutaryl coenzyme A.
†Patients were considered to have hypercholesterolemia if their total cholesterol value was >200 mg/dL (5.17 mmol/L) or if they were receiving a cholesterol-lowering medication. Hypertension was defined as a systolic blood pressure >140 mm Hg, a diastolic blood pressure >90 mm Hg, or the requirement for an antihypertensive medication. Diabetes mellitus was defined by a fasting blood glucose level ≥126 mg/dL (7.0 mmol/L) on at least 2 occasions or the requirement for insulin or oral hypoglycemic agents.

Table 2. Exercise Test Hemodynamics (N = 105)*

Variable	Mean (SD) Sildenafil	Placebo	Mean Difference (95% CI)	P Value
Heart rate, beats/min				
Rest†	64 (11)	64 (11)	−0.19 (−1.6 to 1.2)	.78
Exercise	110 (18)	110 (18)	−0.42 (−2.5 to 1.7)	.69
Difference (exercise − rest)	46 (14)	45 (14)	−0.22 (−2.4 to 1.9)	.84
Systolic blood pressure, mm Hg				
Rest†	128 (17)	133 (19)	4.3 (0.9 to 7.7)	.01
Exercise	174 (29)	176 (30)	−2.4 (−2.5 to 7.4)	.34
Difference (exercise − rest)	46 (24)	44 (26)	−1.9 (−6.8 to 3.0)	.45
Diastolic blood pressure, mm Hg				
Rest†	76 (11)	79 (10)	3.7 (1.9 to 5.5)	<.001
Exercise	93 (15)	95 (16)	2.2 (−0.52 to 5.0)	.11
Difference (exercise − rest)	17 (11)	16 (12)	−1.5 (−4.4 to 1.4)	.32
Double product‡	19 294 (5317)	19 503 (5423)	209 (−589 to 1007)	.60
Exercise capacity, METs	4.5 (1.0)	4.6 (1.0)	0.07 (−0.06 to 0.19)	.29

*CI indicates confidence interval; METs, metabolic equivalent tasks.
†Rest was measured after medication receipt but before exercise testing.
‡Double product is the product of the heart rate and systolic blood pressure at peak exercise.

Clinical, Electrocardiographic, and Echocardiographic Response

The resting WMSI did not change significantly after administration of sildenafil (mean difference, 0; 95% CI, −0.005 to 0.003; $P=.53$) or placebo (mean difference, 0; 95% CI, −0.005 to 0.005; $P>.99$). The resting ejection fraction did not change significantly (56% [7%] before and 57% [7%] after sildenafil; mean difference, 0.3%; 95% CI, −0.5% to 1.0%; $P=.43$) or placebo administration.

Symptoms of dyspnea or angina developed in 69 men taking sildenafil and in 70 men taking placebo ($P=.89$). Reasons for termination of exercise and electrocardiographic interpretations were similar with sildenafil and placebo use (TABLE 3). One subject developed hypotension with exercise after taking 100 mg of sildenafil citrate. This subject's blood pressure decreased from 114/70 to 90/60 mm Hg at peak exercise. Blood pressure in recovery was 70/50 mm Hg. The subject was asymptomatic and was treated with a 500-mL intravenous bolus of isotonic sodium chloride solution. Hypotension persisted for 22 minutes. This subject was taking a calcium channel blocking agent, an α_1-adrenergic receptor blocking agent, and aspirin.

There were no deaths, acute myocardial infarction, or ventricular fibrillation associated with exercise studies. Nasal oxygen was administered

during the recovery period in 6 men (6%) taking sildenafil and 2 (2%) taking placebo ($P=.16$). No subjects required treatment with an intravenous β-blocker during recovery.

The interpretation of exercise echocardiograms was similar for sildenafil and placebo ($P=.49$). The overall summary of interpretations of exercise echocardiograms was as follows (TABLE 4): results were normal in 16 subjects (15%) taking sildenafil and 14 (13%) taking placebo, ischemia was present in 25 subjects (24%) taking sildenafil and 27 (26%) taking placebo, infarction was present in only 5 subjects (5%) taking each, and infarction with ischemia was present in 59 men (56%) taking sildenafil and placebo. In sildenafil and placebo groups, there was no difference in the numbers of subjects with any ischemia (84 and 86 subjects, respectively; $p=.53$) or multivessel ischemia (59 and 57 subjects, respectively; $P=.62$) or in WMSI during exercise, ejection fraction, or heart rate at onset of ischemia.

COMMENT

In this prospective, randomized, double-blind crossover study in men with erectile dysfunction and known or probable coronary artery disease, sildenafil administered 1 hour before maximal, symptom-limited exercise testing was well tolerated and did not change the onset, extent, or severity of ischemia, as assessed by exercise electrocardiography or echocardiography.

Risk factors for erectile dysfunction and coronary artery disease are similar and include age, diabetes mellitus, hypertension, and smoking.[19] The coexistence of coronary artery disease and sexual dysfunction in middle-aged and older men is common. Phase 2/3 studies of sildenafil include predominantly patients without coronary artery disease and patients at low risk for coronary artery disease. In those studies, sildenafil improved erectile function and was well tolerated, and the incidence of severe adverse effects was low.[20-22] However, in patients who have used sildenafil, 130 deaths have been reported to the Food and Drug Administration.[23] Seventy-seven had cardiovascular events, including 41 with definite or suspected myocardial infarction, 27 with cardiac arrest, 6 with cardiac symptoms, and 3 with coronary artery disease. Accordingly, there has been concern regarding the safety of sildenafil in patients with ischemic heart disease.[7] The men in our study are likely representative of many seeking treatment for erectile dysfunction.

This is the first report, to our knowledge, to describe exercise testing with sildenafil monitored by both electrocardiography and an imaging technique. Exercise echocardiography is a well-validated, noninvasive technique to evaluate patients with known or suspected coronary artery disease.[24-26] It is safe, sensitive, and specific, with an overall accuracy similar to that observed with other imaging techniques and higher than that of exercise electrocardiography.[25] In our study, bicycle exercise echocardiography allowed continuous echocardiographic imaging throughout exercise, which enhanced the safety of the study. It also permitted assessment of the heart rate at which new wall motion abnormalities, indicative of ischemia, first developed.

Table 3. Exercise Test Characteristics*

Variable	Patients, No. (%)		Relative Risk (95% CI)†	P Value
	Sildenafil	Placebo		
Reason for termination of exercise				
Fatigue	62 (59)	65 (62)	0.82 (0.28 to 1.35)	.57
Dyspnea	26 (25)	26 (25)	1.00 (0.18 to 2.81)	>.99
Leg distress	16 (15)	11 (10)	3.42 (0.32 to 8.91)	.16
Angina	0	1 (1)	NE	NE
Hypertension	0	2 (2)	NE	NE
Hypotension	1 (1)	0	NE	NE
Arrhythmias				
Ventricular ectopy	51 (49)	55 (52)	0.74 (0.27 to 1.39)	.48
Supraventricular ectopy	36 (34)	46 (44)	0.35 (0.09 to 1.09)	.07
Exercise ECG Interpretation				
Negative	85 (81)	80 (76)	0.05 (0.001 to 1.15)	.09‡
Positive	12 (11)	17 (16)		
Nondiagnostic	8 (8)	8 (8)		
Mean (SD) heart rate at onset of ECG positivity, beats/min	111 (21)	114 (22)	−2.87 (−8.88 to 1.55)§	.52

*NE indicates not estimable because of the small number of events; ECG, electrocardiogram.
†The number of discordant results was low, resulting in large confidence intervals (CIs) for some of the relative risk estimates.
‡Comparison of positive ECG result vs negative or nondiagnostic ECG result.
§Refers to mean difference rather than relative risk.

Table 4. Exercise Echocardiography*

Variable	Sildenafil	Placebo	Mean Difference (95% CI)	P Value
Rest echocardiography				
WMSI	1.2 (0.3)	1.2 (0.3)	−0.005 (−0.02 to 0.005)	.30
Ejection fraction, %	56 (7)	55 (7)	−1.08 (−1.90 to −0.26)	.01
Exercise echocardiography				
WMSI	1.4 (0.4)	1.4 (0.4)	0.01 (−0.01 to 0.03)	.40
Ejection fraction, %	60 (10)	60 (9)	0.38 (−0.74 to 1.49)	.50
Ejection fraction difference, %†	4.0 (6.1)	4.4 (6.4)	0.33 (−0.79 to 1.45)	.56
LV end systolic volume, mL†	−7 (14)	−8 (12)	0.61 (−2.33 to 3.55)	.68
Percentage of ischemic segments	19 (17)	20 (17)	0.01 (−0.01 to 0.02)	.51
Heart rate at onset of new wall motion abnormalities, beats/min	96 (14)	98 (16)	−0.32 (−2.36 to 1.73)	.76

*Data are given as mean (SD) except where noted. CI indicates confidence interval; WMSI, wall motion score index; and LV, left ventricle.
†Difference between exercise and rest measurements.

In most of our subjects, exercise echocardiography demonstrated exercise-induced myocardial ischemia. However, exercise was no more likely to induce ischemia after sildenafil use than after placebo use. The extent and severity of ischemia and the heart rate at which it developed were similar. In this study group of 105 men, despite frequent ischemia after either placebo or sildenafil use, there were no clinically significant events. However, the study was powered only to assess the impact of sildenafil on extent and severity of ischemia. Larger numbers would be needed to assess any potential impact on clinical events.

Despite the randomized, double-blind design of the study, subjects may have been able to determine when they received sildenafil. However, the paramedical staff members who administered the test were not told which drug the subject had received. The physician who interpreted the tests was not present during the performance of the tests. Therefore, it is unlikely that nonblinding could have affected the results of this study.

In our study subjects, there was a slight decrease in blood pressure at rest after sildenafil administration, without changes in heart rate. A discrete blood pressure reduction at rest after sildenafil administration has been attributed to vasodilation and described both in healthy subjects and patients with stable angina.[27,28] The peak plasma concentration following an oral dose occurs approximately 1 hour after sildenafil administration.[27] The maximum decrease in blood pressure occurs at this time.[27,29] Therefore, to maximize any potential adverse effects of the drug, stress testing was performed 1 hour after administration of the drug.

Exercise, including sexual activity, may trigger acute coronary events in patients with coronary artery disease.[30-32] In our subjects, the exercise-induced increments in blood pressure and heart rate with exercise were similar with and without sildenafil use. During sexual intercourse, heart rate and blood pressure increase as with other forms of exertion.[33,34] Hellerstein and Friedman[35] monitored middle-aged men with known or suspected coronary artery disease during sexual intercourse with their spouses at home and observed a mean peak heart rate of 117/min. The mean (SD) peak exercise heart rate of patients in our study was slightly less (110 [18]/min). Our observations suggest that myocardial ischemia during sexual activity may be common in men with stable coronary artery disease.

The typical maximum workload during coitus is approximately 3.3 to 3.4 METs for less than 30 seconds.[34,35] Guidelines from the American College of Cardiology and the American Heart Association suggest that if a patient can exercise more than 5 to 6 METs without demonstrating ischemia on exercise electrocardiography testing, the risk of ischemia during sexual intercourse is probably low.[7] In our study, which included only men who could exercise, the mean (SD) exercise capacity was not affected by sildenafil use (4.5 [1.0] METs with sildenafil and 4.6 [1.0] METs with placebo, $P=.29$).

In this study, men were not taking nitrates or had discontinued the use of nitrates 72 hours before exercise testing. Both nitrates and sildenafil promote increased cyclic guanosine monophosphate levels. An interaction during concomitant administration of sildenafil and nitrates promotes marked reductions in blood pressure because of vasodilation in both animal models and humans.[29,36] Of the cardiovascular deaths reported to the Food and Drug Administration, some involved a possible interaction between sildenafil and nitrates.[7,23] Therefore, the coadministration of nitrates and sildenafil was assiduously avoided in our study.

Oral sildenafil increases coronary flow reserve in severely stenotic coronary arteries to an extent comparable to the increase in normal coronary arteries, thus preserving the ratio of flow reserve in stenotic and normal vessels.[18] In another study[37] of patients with chronic heart failure, oral sildenafil increased epithelium-dependent, flow-mediated vasodilation when compared with placebo. These studies, albeit in small numbers of patients, support our conclusions that oral sildenafil does not have an adverse effect on stress-induced myocardial ischemia in patients with ischemic heart disease.

The number of men with left ventricular dysfunction included in our study was limited. Although a variety of medications were taken by the subjects in this study, numbers were too small to permit subgroup analysis. Significant hypotension developed in a single subject who had taken 100 mg of sildenafil citrate. Numbers of men receiving the 100-mg dose were too small for conclusions to be drawn about the safety of this dose in this population.

Patients with known or suspected coronary artery disease and erectile dysfunction should have an individualized assessment before sildenafil prescriptions are issued. Exercise testing can be performed after sildenafil administration and may be indicated for risk stratification of some patients. Patients with stable coronary artery disease who are able to exercise to 4.5 METs without angina or hypotension and with a negative or mildly positive stress test result can probably safely take sildenafil. Further research will be needed, though, to clarify what levels of functional capacity and severity of ischemia can be considered truly safe for men with coronary disease who wish to use sildenafil.

CONCLUSIONS

In men who had known or probable coronary artery disease and were able to exercise, sildenafil had no effect on the presence or extent of exercise-induced regional wall motion abnormalities, symptoms, exercise duration, or arrhythmias. In patients who have stable coronary artery disease and are not taking nitrates, sildenafil did not potentiate myocardial ischemia.

Author Contributions: *Study concept and design:* Arruda-Olson, Mahoney, Nehra, Pellikka
Acquisition of data: Leckel, Pellikka
Analysis and interpretation of data: Arruda-Olson, Mahoney, Nehra, Leckel, Pellikka
Drafting of the manuscript: Arruda-Olson, Mahoney, Nehra, Pellikka

Critical revision of the manuscript for important intellectual content: Mahoney, Leckal, Pellikka.
Statistical expertise: Mahoney
Obtained funding: Arruda-Olson, Pellikka.
Administrative, technical, or material support. Pellikka.
Study supervision: Pellikka.
Funding/Support: The study was supported by grants from the Mayo Foundation and from the American Heart Association, Northland Affiliate. Dr Arruda-Olson was supported by grants from the CAPES Foundation (Fundação Coordenação de Aperfeiçoamento de Pessoal de Nível Superior), Brasília, Brazil, and from the Mayo Foundation.

REFERENCES

1. National Institutes of Health Consensus Development Panel on Impotence. Impotence. *JAMA*. 1993; 270:83-90.
2. Harrold LR, Gurwitz JH, Field TS, et al. The diffusion of a novel therapy into clinical practice: the case of sildenafil. *Arch Intern Med*. 2000;160:3401-3405.
3. Feenstra J, Van Drie-Pierik RJ, Lade CF, Strickec BH. Acute myocardial infarction associated with sildenafil. *Lancet*. 1998;352:957-958.
4. Porter A, Mager A, Birnbaum Y, Sclarovsky S, Rechavia E. Acute myocardial infarction following sildenafil citrate (Viagra) intake in a nitrate free patient. *Clin Cardiol*. 1999;22:762-763.
5. Shah PK. Sildenafil in the treatment of erectile dysfunction (letter). *N Engl J Med*. 1998;339:699.
6. Hayashi K, Minekaki KK, Narukawa M, Ookubo M, Mitsuhashi T, Shimada K. Atrial fibrillation and continuous hypotension induced by sildenafil in an intermittent WPW syndrome patient. *Jpn Heart J*. 1999; 40:827-830.
7. Cheitlin MD, Hutter AM Jr, Brindis RG, et al. Use of sildenafil (Viagra) in patients with cardiovascular disease. *Circulation*. 1999;99:168-177.
8. Diamond GA, Forrester JS. Analysis of probability test as an aid in the clinical diagnosis of coronary artery disease. *N Engl J Med*. 1979;300:1350-1358.
9. Froelicher VF, Myers JN. Interpretation of hemodynamic responses to exercise testing. In: Froelicher VF, Myers JN, eds. *Exercise and the Heart*. 4th ed. Philadelphia, Pa: WB Saunders; 2000:93-120.
10. Roger VL, Pellikka PA, Oh JK, Miller FA, Seward JB, Tajik AJ. Stress echocardiography, part I: exercise echocardiography: techniques, implementation, clinical applications, and correlations. *Mayo Clin Proc*. 1995;70:5-15.
11. Armstrong WF, Pellikka PA, Ryan T, Crouse L, Zoghbi W. Stress echocardiography: recommendations for performance and interpretation of stress echocardiography. *J Am Soc Echocardiogr*. 1998;11:97-104.
12. Schiller NB, Shah PM, Crawford M, et al. Recommendations for quantification of the left ventricle by two-dimensional echocardiography: American Society of Echocardiography Committee on Standards, Subcommittee on Quantification of Two-dimensional Echocardiograms. *J Am Soc Echocardiogr*. 1989;2:358-367.
13. Schiller NB, Aquatella H, Ports TA, et al. Left ventricular volume from paired bi-plane two-dimensional echocardiograms. *Circulation*. 1979;60:547-555.
14. Grizzle JE. The two-period change over design and its use in clinical trials. *Biometrics*. 1965;21:467-480.
15. Hills M, Armitage P. The two-period cross-over clinical trial. *Br J Clin Pharmacol*. 1979;8:7-20.
16. Prescott RJ. The comparison of success rates in cross-over trials in the presence of an order effect. *Appl Stat*. 1979;30:9-15.
17. Gart JJ. An exact test for comparing matched proportions in crossover designs. *Biometrika*. 1969;56: 75-80.
18. Zhang J, Yu KF. What's the relative risk? a method of correcting the odds ratio in cohort studies of common outcomes. *JAMA*. 1998;280:1690-1691.
19. Feldman HA, Goldstein I, Hatzichristou DG, Krane RJ, McKinlay JB. Impotence and its medical and psychological correlates: results of the Massachusetts Male Aging Study. *J Urol*. 1994;151:54-61.
20. Goldstein I, Lue TF, Padma-Nathan H, Rosen RC, Steers WD, Wicker PA. Oral sildenafil in the treatment of erectile dysfunction. *N Engl J Med*. 1998; 338:1397-1404.
21. Morales A, Gingell C, Collins M, Osterloh IH. Clinical safety of oral sildenafil citrate (Viagra) in the treatment of erectile dysfunction. *Int J Impot Res*. 1998;10:69-74.
22. Padma-Nathan H, Steers WD, Wicker PA. Efficacy and safety of oral sildenafil in the treatment of erectile dysfunction: a double blind, placebo-controlled study of 329 patients. *Int J Clin Pract*. 1998; 52:375-380.
23. Postmarketing Safety of Sildenafil Citrate (Viagra). US Food and Drug Administration Center for Drug Evaluation and Research. Available at: http://www.verity.fda.gov/cder/consumerinfo/viagra/safety3.htm. Accessed November 20, 2001.
24. Pellikka PA. Stress echocardiography in the evaluation of chest pain and accuracy in the diagnosis of coronary artery disease. *Prog Cardiovasc Dis*. 1997; 39:523-532.
25. Fleischmann KE, Hunink MG, Kuntz KM, Douglas PS. Exercise echocardiography or exercise SPECT imaging? a meta-analysis of diagnostic test performance. *JAMA*. 1998;280:913-920.
26. Badruddin SM, Ahmad A, Mickelson J, et al. Supine bicycle versus post-treadmill exercise echocardiography in the detection of myocardial ischemia: a randomized single-blind crossover trial. *J Am Coll Cardiol*. 1999;33:1485-1490.
27. Jackson G, Benjamin N, Jackson N, Allen MJ. Effects of sildenafil citrate on human hemodynamics. *Am J Cardiol*. 1999;83(suppl 5A):13C-20C.
28. Herrmann HC, Chang G, Klugherz BD, Mahoney PD. Hemodynamic effects of sildenafil in men with severe coronary artery disease. *N Engl J Med*. 2000;342:1622-1626.
29. Webb DJ, Muirhead GJ, Wulf M, Sutton JA, Levi R, Dinsmore WW. Sildenafil citrate potentiates the hypotensive effects of nitric oxide donor drugs in male patients with stable angina. *J Am Coll Cardiol*. 2000; 36:25-31.
30. Master AM, Jaffee HL. Factors in the onset of coronary occlusion and coronary insufficiency: effort, occupation, trauma, and emotion. *JAMA*. 1952;148: 794.
31. Falk E. Why do plaques rupture? *Circulation*. 1992; 86:III30-III42.
32. Muller JE, Mettleman A, Maclure M, Sheerwood JB, Tofler GH. Triggering myocardial infarction by sexual activity. *JAMA*. 1996;275:1405-1409.
33. Jackson G. Sexual intercourse and stable angina pectoris. *Am J Cardiol*. 2000;86:35F-37F.
34. Bohlen JG, Held JP, Sanderson O, Patterson RP. Heart rate, rate-pressure product, and oxygen uptake during four sexual activities. *Arch Intern Med*. 1984;144:1745-1748.
35. Hellerstein HK, Friedman EH. Sexual activity and the post coronary patient. *Arch Intern Med*. 1970; 125:987-999.
36. Ishikura F, Beppu S, Hamada T, Khandheria BK, Seward JB, Nehra A. Effects of sildenafil citrate (Viagra) combined with nitrate on the heart. *Circulation*. 2000;102:2516-2521.
37. Katz SD, Balidemaj K, Homma S, Wu H, Wang J, Maybaum S. Acute type 5 phosphodiesterase inhibition with sildenafil enhances flow-mediated vasodilation in patients with chronic heart failure. *J Am Coll Cardiol*. 2000;36:845-851.

Sputum eosinophilia and short-term response to prednisolone in chronic obstructive pulmonary disease: a randomised controlled trial (178) [363]

Christopher E Brightling, William Monteiro, Richard Ward, Debbie Parker, Michael D L Morgan, Andrew J Wardlaw, Ian D Pavord

Summary

Background Some patients with chronic obstructive pulmonary disease (COPD) respond to corticosteroid therapy. Whether these patients have different airway pathology from other COPD patients is unclear. We tested the hypothesis that response to prednisolone is related to the presence of eosinophilic airway inflammation.

Methods We did a randomised, double-blind, crossover trial. Patients who had COPD treated with bronchodilators only were assigned placebo and 30 mg prednisolone daily for 2 weeks each, in a random order, separated by a 4-week washout period. Before and after each treatment period, we assessed patients with spirometry, symptom scores, the chronic respiratory disease questionnaire (CRQ), incremental shuttle walk test, and induced sputum. Analysis was done by intention to treat.

Findings 83 patients were recruited, of whom 67 were randomised. The geometric mean sputum eosinophil count fell significantly after prednisolone (from 2·4% to 0·4%; mean difference six-fold [95% CI 3·1–11·4]) but not after placebo. Other sputum cell counts did not change. After stratification into tertiles by baseline eosinophil count, postbronchodilator forced expiratory volume in 1 s (FEV_1) and total scores on the CRQ improved progressively after prednisolone from the lowest to the highest eosinophilic tertile, compared with placebo. The mean change in postbronchodilator FEV_1, total CRQ score, and shuttle walk distance with prednisolone compared with placebo in the highest tertile was 0·19 L (0·06–0·32), 0·62 (0·31–0·93), and 20 m (5–35), respectively.

Interpretation Our findings suggest that eosinophilic airway inflammation contributes to airflow obstruction and symptoms in some patients with COPD and that the short-term effects of prednisolone are due to modification of this feature of the inflammatory response. The possibility that sputum eosinophilia identifies a subgroup of patients who particularly respond to long-term treatment with inhaled corticosteroids should be investigated.

Lancet 2000; **356**: 1480–85

Institute for Lung Health, Department of Respiratory Medicine, Glenfield Hospital, Leicester LE3 9QP, UK (C E Brightling MRCP, W Monteiro BSc, R Ward BSc, D Parker BSc, M D L Morgan FRCP, A J Wardlaw FRCP, I D Pavord MRCP)

Correspondence to: Dr Christopher E Brightling (e-mail: chris.brightling@glenfield.tr-trent.nhs)

Introduction

The role of corticosteroids in stable chronic obstructive pulmonary disease (COPD) is uncertain.[1] Guidelines reflect the perception that a subgroup of patients respond to this treatment,[2] although identification of characteristics associated with a positive response to short-term or long-term treatment with corticosteroids has been difficult.[3] Corticosteroids effectively modify eosinophilic airway inflammation in asthma,[4,5] but there is less evidence that they effect the neutrophilic inflammation that predominates in COPD.[3] One possibility is that corticosteroids are effective in COPD patients who have eosinophilic airway inflammation.

Early reports suggested that sputum eosinophilia could predict clinical benefit from corticosteroids,[6] but the methods of sputum assessment were crude and the findings were not consistent.[3] Over the past 10 years, important advances have been made in the technique of sputum induction and analysis and methods are now reliable, valid,[7] responsive, and safe.[8] Pizzichini and colleagues[10] used these methods to show, in a small, single-blind study, that patients with sputum eosinophilia had a greater improvement in forced expiratory volume in 1 s (FEV_1) and health status after a short course of prednisone than those without. We did a double-blind placebo-controlled crossover study to define the sputum characteristics of patients with COPD and to investigate the relation between the sputum eosinophil count and the response to 2 weeks of treatment with prednisolone.

Methods

Patients

We recruited, from respiratory clinics, patients who had symptoms of chronic airflow obstruction, post-bronchodilator FEV_1 of less than 70% predicted, and an FEV_1/forced vital capacity (FVC) ratio of less than 70%. We enrolled those who had no substantial improvement in FEV, after taking 2·5 mg nebulised salbutamol (<15% or, if FEV_1 <1·2 L, <200 mL improvement). We excluded patients if they had a clinical diagnosis of asthma, a history of childhood respiratory disorders, variability in symptoms not associated with infections, a history of acute wheeze, breathlessness, or deterioration associated with allergens, or an exacerbation within 6 weeks of trial entry. Patients taking regular oral corticosteroids were excluded. In patients taking inhaled corticosteroids, these drugs were discontinued for at least 1 month before randomisation. We withdrew patients from the study if they had a moderate exacerbation requiring inhaled corticosteroids or antibiotics, a severe exacerbation needing oral corticosteroids, or a severe intercurrent illness. The study was approved by the local research ethics committee and all patients gave written informed consent.

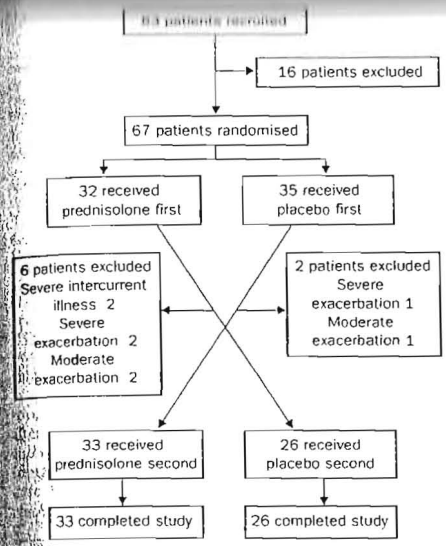


Figure 1: Trial profile

Study design

Patients were randomly assigned, by random numbers, prednisolone 30 mg daily and placebo for 2 weeks each in a random order (figure 1). The two agents were administered as single white capsules. These were prepared by the Royal Hallamshire Hospital Pharmacy Department, Sheffield, UK. Glenfield Hospital Pharmacy Department did the randomisation, distributed the study agents, and held the trial codes, which was disclosed after the study. The two treatment periods were preceded by a 4-week run-in period and separated by a 4-week washout period, since shorter washout times have been associated with a carryover effect.[9] Patients attended on four occasions before, and 4–6 h after the last dose of study medication, at the same time of day on each occasion more than 6 h after their last dose of bronchodilator, and 24 h after the last dose of long-acting β2-agonists. At each visit we assessed patients by spirometry before and 15 min after 2·5 mg nebulised salbutamol, the chronic respiratory disease questionnaire (CRQ) for health status, a 100 mm visual analogue scale for symptom scores (no symptom to the worst symptom ever), an incremental shuttle walk test, and sputum induction for differential cell count. The timing of events was kept constant in relation to the nebulised salbutamol.

We recorded details of smoking status, normal treatment, and atopic and childhood respiratory history. We recorded symptom scores for dyspnoea, cough, sputum production, and wheeze. Spirometry was done with a Compact Vitalograph spirometer (Vitalograph, Buckinghamshire, UK). Salbutamol was administered via a Flaem Nuova Type II nebuliser (Deva Medical, Runcorn, Cheshire) with a median particle size of 2 μm and the patient breathing tidally. We recorded FEV_1 as the better of two successive readings within 100 mL. Lung-function tests were done with a benchmark (P K Morgan, Chatham, UK) and lung volumes assessed by the helium dilution method. We took venous blood samples to measure peripheral blood cosinophil count, total IgE and radioallergosorbent tests to *Dermatophagoides pteronyssinus*, cat fur, and grass pollen.

We assessed health status with the CRQ, which consists of 20 questions in four domains: dyspnoea, fatigue, emotions, and mastery.[11] A seven-point Likert scale was used for each question and the total score and each domain score was recorded out of seven, with a minimum important difference of 0·5. The incremental shuttle walk test was done according to a standard protocol,[11] which has been shown to be repeatable after one practice walk. All patients had a practice walk before entry into the study. We recorded the total distance of the completed shuttles.

Sputum was induced and processed as previously described.[8,14] Briefly, sputum was induced by use of 3%, 4%, and 5% saline, inhaled in sequence for 5 min via an ultrasonic nebuliser (Medix, Harlow, UK) 30 min after 2·5 mg nebulised salbutamol. Once expectorated, we stored sputum on ice for analysis within 2 h. A differential cell count was obtained by counting more than 400 non-squamous cells on Romanovski-stained cytospins.

We measured sputum elastase in the cell-free supernatant by spectrofluorimetric assay, in which methoxy-succinyl-L-alanyl-L-alanyl-L-prolyl-L-valyl-amino-methylcoumarin (Sigma, Poole, UK) was used as the substrate.[15] We added 50 μL of the substrate solution (0·1 mg/mL) with appropriate negative controls to 50 μL of elastase standards (Elastin Products, Owensville, MI, USA), samples, and quality-control samples. The plate was read every 5 min with a fluorimeter (Victor Wallac, Perkin Life Sciences, Cambridge, UK) that used an

	Before prednisolone	After prednisolone	Before placebo	After placebo
FEV_1 (L)	1·03 (0·05)	1·09 (0·05)	1·09 (0·05)	1·08 (0·05)
Postbronchodilator FEV_1 (L)	1·10 (0·05)	1·15 (0·05)	1·15 (0·06)	1·14 (0·06)
CRQ total score	3·98 (0·13)	4·34 (0·14)	4·03 (0·14)	4·08 (0·14)
Walk (m)	217 (13)	230 (14)	213 (12)	214 (12)
Total cell count (10^6 cells/g sputum)*	2·79 (0·05)	2·37 (0·04)	2·66 (0·05)	2·63 (0·05)
Eosinophil count (%)*	2·35 (0·09)	0·39 (0·14)†	2·18 (0·09)	2·08 (0·13)
Neutrophils (%)	67·9 (3·2)	70·4 (2·7)	73·7 (2·5)	69·0 (2·7)‡
Macrophages (%)	21·5 (2·3)	25·4 (2·7)	22·1 (2·5)	22·7 (2·4)
Squamous epithelial cells (%)	1·17 (0·2)	1·37 (0·23)	0·99 (0·16)	1·47 (0·35)‡
Lymphocytes (%)	0·40 (0·07)	0·44 (0·15)	0·35 (0·05)	0·71 (0·25)‡
Mucus plug cells (%)‡	1·85 (0–51)	3·15 (0–49)	2·6 (0–40)	2·2 (0–51)
Viability (%)‡	76 (28–95)	67 (20–86)	73 (23–91)	71 (18–89)
Elastase (ng/mL sputum*	1454 (0·08)	801 (0·08)	1149 (0·09)	1236 (0·09)
Total IgE (kU/L)*	53·2 (0·12)	31 (0·12)	42·5 (0·11)	36·5 (0·11)
Interleukin 8 (ng/mL)*	117 (0·1)	105 (0·1)	112 (0·08)	97 (0·1)

ECP=eosinophilic cationic protein; CRQ=chronic respiratory disease questionnaire. *Geometric mean (log SE). †p<0·01. ‡Median (range)

Table 1: Mean (SE) outcome measures before and after treatment

Patients' characteristics			
Male/female	11 (55%)/ 9 (45%)	13 (57%)/ 10 (43%)	15 (68%)/ 7 (32%)
Mean (range) age (years)	68 (42–82)	66 (49–83)	64 (47–78)
Current smoker	6	6	5
Ex smokers	14	17	15
Mean (SE) number of pack years	33 (4·1)	35 (2·5)	37 (4·3)
Test outcomes (mean [SE])			
Atopic by positive RAST	7	7	6
IgE (KU/L)	110 (40)	145 (52)	108 (31)
FBC Eo (×10⁹/L)	0·12 (0·02)	0·23 (0·04)	0·17 (0·02)
Shuttle walk distance (m)	176 (21·1)*	247 (22·5)	246 (23·3)
Total CRQ score	3·92 (0·25)	4·04 (0·19)	4·08 (0·25)
FEV_1 (L)	1·15 (0·09)	1·11 (0·09)	0·96 (0·08)
Postbronchodilator FEV_1 (L)	1·19 (0·14)	1·15 (0·09)	1·01 (0·08)
FEV_1, % predicted	44·3 (2·6)	43·6 (3·5)	37·8 (3·0)
FVC (L)	2·30 (0·14)	2·25 (0·11)	2·32 (0·17)
Total lung capacity % predicted	93 (4·7)	96 (3·6)	97 (3·6)
Corrected carbon monoxide transfer coefficient (KCO) % predicted	89 (7·5)	91 (7·7)	84 (8·5)
Sputum characteristics†			
Eosinophils (%)	0·5 (0·08)	2·5 (0·02)	12·3 (0·07)
Neutrophils (%)	79·4 (4·0)	77·4 (2·5)	59·1 (5·3)*
Total cell count (10^6 cells/g sputum)	2·69 (0·1)	2·74 (0·08)	2·88 (0·08)
ECP (ng/mL)	1183 (0·13)	869 (0·13)	3162 (0·12)*
Elastase (µg/mL)	78 (0·19)	27 (0·25)	48 (0·17)
Interleukin 8 (ng/mL)	97 (0·18)	98 (0·14)	106 (0·17)
CRQ domain scores (mean [SE])			
Dyspnoea	3·54 (0·24)	3·42 (0·24)	3·36 (0·26)
Fatigue	3·36 (0·30)	3·82 (0·26)	3·75 (0·26)
Emotion	4·12 (0·26)	4·53 (0·25)	4·66 (0·30)
Mastery	4·64 (0·28)	4·4 (0·25)	4·56 (0·32)
Symptom scores (mean [SE])			
Cough VAS (mm)	47·3 (6·1)	26·1 (3·5)*	36·3 (5·6)
Sputum VAS (mm)	36·4 (5·6)	19·9 (2·5)*	30·8 (5·7)
Dyspnoea VAS (mm)	50·1 (5·0)	43·8 (4·7)	54·2 (5·3)
Wheeze VAS (mm)	43·0 (6·3)	26·7 (5·6)	31·3 (5·9)
Sputum characteristics†			
Lymphocytes (%)	0·3 (0·09)	0·3 (0·07)	0·6 (0·13)*
Macrophages (%)	19·7 (4·2)	22·6 (3·9)	21·3 (3·5)
Bronchial epithelial (%)	0·9 (0·4)	1·0 (0·2)	1·3 (0·3)
Squamous cells (%)	3·4 (0–42)	1·1 (0–51)	2·6 (0–51)
Viability (%)	77 (35–89)	79 (27–95)	68 (23–83)

RAST=radioallergosorbent test. *p<0·05. †Values are geometric mean (SE), except neutrophils, which are mean (SE). ‡Values are geometric mean (log SE), except squamous cells and viability, which are median (IQR).

Table 2: **Baseline characteristics after stratification into tertiles by baseline eosinophil counts**

excitation wavelength of 340 nm and emission of 455 nm over 2 h at 37°C. We calculated the standard curve and unknown values from the maximum slope of fluorescence compared with time curves. We measured interleukin 8 with a commercial ELISA (OptEIA Set, Becton Dickinson, Cowley, Oxford, UK) and eosinophilic cationic protein with a commercial fluoroimmunoassay (Unicap, Pharmacia, Milton Keynes, UK). The limits of detection for elastase, interleukin 8 and eosinophilic cationic protein were 1·5 µg/mL, 30 pg/mL, and 18 ng/mL sputum, respectively, and the between-assay and within-assay variabilities were between 5% and 10%. Spiking experiments confirmed complete recovery of mediators added to the cell-free supernatant.

The primary outcomes were change in postbronchodilator FEV_1, total CRQ score, and shuttle walk distance after prednisolone compared with placebo. The secondary outcome measures were changes in the individual domains of chronic respiratory disease questionnaire, symptom scores, and sputum characteristics.

Statistical analysis

To assess the association between sputum eosinophil count and primary outcomes, we stratified patients into tertiles by baseline sputum eosinophil count. Baseline data were derived from measurements taken before the first treatment phase. The study had more than 80% power at the 5% level to detect a 150 mL difference in the change in FEV_1 within tertiles, assuming a within-patient SD of 100 mL.[16] A sample size of eight patients in each tertile would have been sufficient to achieve this power. If, however, we allowed for the possibility of a carry-over effect and a 10% dropout rate before completion of the first treatment phase, a sample size of 20 patients in each tertile was required.

We analysed all data with the Minitab statistical package for Windows (version 11). Spirometry, shuttle walk distance, total and domain CRQ scores, and symptom scores are expressed as mean (SE). Eosinophil count, total cell count, and mediator concentrations were log normally distributed and are described as geometric mean (log SE) with fold change (95% CI). We report other baseline

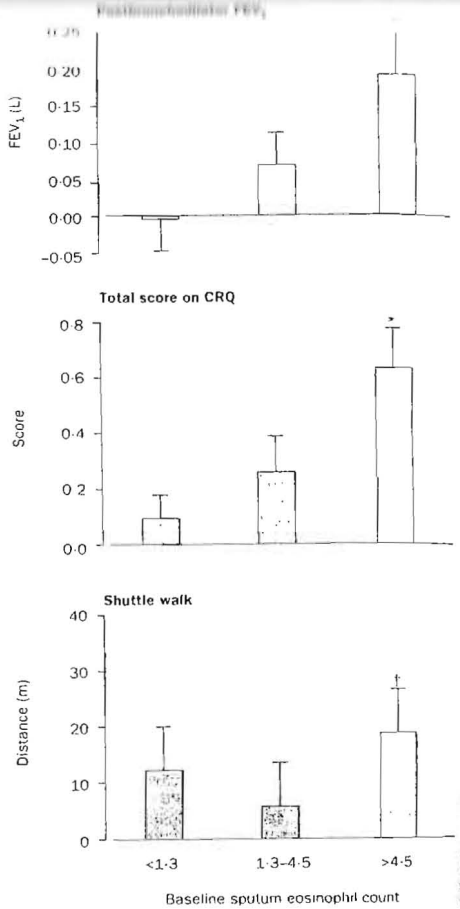


Figure 2: **Mean (SE) absolute increase in primary outcomes for each tertile after prednisolone compared with placebo**
*p<0·01. †p<0·05

	<1·3 (n=22)	1·3–4·5 (n=21)	>4·5 (n=22)
CRQ domain*			
Dyspnoea CRQ	0·15 (−0·04 to 0·35)	0·27 (−0·11 to 0·65)	0·93 (0·50 to 1·35)†
Fatigue CRQ	0·09 (−0·28 to 0·46)	0·42 (0·06 to 0·79)‡	0·56 (0·20 to 0·91)†
Emotion CRQ	0·08 (−0·18 to 0·35)	0·12 (−0·15 to 0·39)	0·19 (−0·11 to 0·49)
Mastery CRQ	0·03 (−0·23 to 0·30)	0·20 (−0·202 to 0·59)	0·82 (0·37 to 1·26)†
Symptom scores§			
Cough VAS‡	2 (−27 to 31)	41 (9 to 74)‡	53 (−1 to 106)
Sputum VAS‡	6 (−15 to 26)	34 (2 to 66)‡	45 (18 to 72)†
Dyspnoea VAS‡	5 (−3 to 14)	14 (3 to 27)‡	32 (9 to 55)†
Wheeze VAS‡	−1 (−15 to 13)	21 (−1 to 14)	42 (−0 to 85)
Sputum indices‖			
Total cell count	1·03 (0·62 to 1·72)	1·38 (0·93 to 2·03)	1·11 (0·74 to 1·65)
Eosinophil count	3·1 (0·9 to 10·6)	5·4 (1·2 to 24·2)†	6·9 (1·3 to 36·4)†
Neutrophil count	−3·0 (−12·5 to 6·5)	−4·5 (−15·2 to 6·2)	−19·3 (−31·8 to −6·8)‡
Oesinophilic cationic protein	1·2 (0·6 to 2·7)	1·5 (0·7 to 3·1)	3·4 (1·4 to 8·1)†
Elastase	1·9 (0·5 to 7·5)	1·2 (0·2 to 6·5)	2·8 (0·5 to 13·8)
Interleukin 8	1·3 (0·3 to 5·6)	0·7 (0·1 to 2·9)	2·1 (0·7 to 6·2)

*Absolute decrease. †p<0·01 ‡p<0·05. §Percentage decrease. ‖Fold decrease except neutrophil count, which is absolute decrease.

Table 3: **Mean (95% CI) change in secondary outcomes after prednisolone treatment compared with placebo**

sputum differential cell counts as mean (SE) and change after treatment for these as mean (95% CI). Cell viability and squamous-cell contamination are expressed as median (range). Improvement in the primary and secondary outcome variables after prednisolone compared with placebo are reported as paired mean differences (95% CI). We compared differences in patients' characteristics between tertiles by ANOVA, with the Student-Newman-Keuls procedure to correct for multiple comparisons. The significance of the trend between tertiles in the change in primary and secondary outcome variables was analysed by linear regression. All p values are two-tailed. Analyses were done by intention to treat. Patients who withdrew in the washout phase were assigned a net change of zero for the second treatment phase.

Results

83 patients were recruited, of whom 67 were randomised (figure 1). Seven of 45 patients taking inhaled corticosteroids at recruitment were excluded before randomisation because they developed moderate exacerbations in the run-in period. The baseline characteristics of the remaining 38 patients did not differ significantly from the corticosteroid-naïve patients. Eight patients withdrew during the washout period (figure 1). Treatment period or order did not influence values before treatment or the changes in the primary outcome variables.

The geometric mean sputum eosinophil count significantly decreased after treatment with prednisolone from 2·4% to 0·4% (mean difference six-fold [95% CI 3·1–11·4], p<0·0001) but not after placebo (table 1). Treatment-associated change did not differ between treatments for other sputum cell counts. 29 (43%) patients had a baseline eosinophil differential count higher than the normal range in our laboratory (>3%).

The mean paired difference between prednisolone and placebo treatment for the change in primary outcomes for the whole group were: postbronchodilator FEV_1 0·07 L (95% CI 0·01–0·14, p=0·02), total score on CRQ 0·32 (0·17–0·47, p=0·0001), and shuttle walk distance 12 m (3–21, p=0·01). After stratification into tertiles by the baseline sputum eosinophil count, patients were well matched for age, sex, smoking history, and atopy, and did not differ in baseline spirometry, lung function, or total score on CRQ (table 2).

Mean difference between prednisolone and placebo increased progressively from the lowest to the highest eosinophilic tertile for change in postbronchodilator FEV_1 (p=0·003) and total score for CRQ (p=0·02), but not for shuttle walk distance (p=0·56, figure 2). The mean change in postbronchodilator FEV_1, CRQ scores, and shuttle walk distance with prednisolone compared with placebo in the highest tertile were 0·19 L (0·06–0·32, p=0·005), 0·62 (0·31–0·93, p=0·0005), and 20 m (5–35, p=0·013), respectively. The change in secondary outcome measures showed a similar progressive improvement from least to most eosinophilic tertile. The trend was significant for the chronic respiratory disease scores for dyspnoea (p=0·005), fatigue (p=0·04), and mastery domains (p=0·01), and the symptoms of wheeze (p=0·02), sputum production (p=0·04), and breathlessness (p=0·01, table 3).

The decrease in eosinophilic cationic protein in response to prednisolone was significant for the highest eosinophilic tertile (table 3). The concentration of elastase and interleukin 8 did not differ between the tertiles or in response to prednisolone (tables 2 and 3).

Discussion

Eosinophilic airway inflammation was common among patients who had stable moderate and severe COPD, and postbronchodilator FEV_1, health status, and exercise capacity were improved by prednisolone. Greater improvements were seen in patients with higher baseline sputum eosinophil counts than for those with lower counts, and were associated with striking reductions in sputum eosinophil count and sputum concentration of the activated eosinophil product eosinophilic cationic protein. Cell and molecular markers of neutrophilic inflammation were not affected. These findings strongly suggest that eosinophilic airway inflammation contributes to airflow obstruction and symptoms in some patients with COPD, and the effects of corticosteroids are due to inhibition of this feature of the inflammatory response in COPD.

The mean effect of prednisolone on FEV_1 after bronchodilator use was slight, even in the highest tertile of sputum eosinophil count, which suggests that other factors, such as a permanent structural defect and corticosteroid-unresponsive non-eosinophilic airway inflammation are the main causes of airflow obstruction in COPD. Changes in health status and symptom scores associated with treatment were more obvious, and in most domains of the CRQ surpassed the minimum difference for clinical importance.[12] As with FEV_1, improvement in these variables was related to the sputum eosinophil count, which suggests that it

represents an effect on the airways rather than a non-specific corticosteroid effect.

Our findings for FEV_1, health status, sputum inflammatory cell counts, and sputum mediators are in accord with those of Pizzichini and colleagues.[16] The findings are also consistent with the those of Chanez and colleagues[17] who showed that patients with COPD who improved after open-label treatment with prednisolone generally had pathological features of asthma on bronchial biopsy and bronchoalveolar lavage. By contrast, Keatings and colleauges[5] saw no effect on FEV_1 or sputum eosinophil count after 2 weeks of treatment with oral prednisolone in a small single-blind study, although there was no strong sputum evidence of eosinophilic airway inflammation. The effects of inhaled corticosteroids on airway inflammation in COPD are less consistent. In one open study, 2 months of treatment with inhaled beclomethasone dipropionate lowered the sputum neutrophil count,[18] whereas Keatings and co-workers[5] found no change in sputum differential cell counts after 2 weeks of inhaled budesonide compared with placebo. Whether the effects of inhaled and oral corticosteroids differ is unclear and needs further study.

COPD is associated with sputum and bronchial biopsy evidence of neutrophilic airway inflammation and the extent of this inflammation, as reflected by the induced sputum neutrophil differential count, inversely correlates with FEV_1 and the decline in FEV_1, which suggests that it is functionally important.[19–21] The patients in our study, who had moderate and severe COPD, had sputum neutrophilia, in keeping with predominant neutrophilic airway inflammation. Little attention has been paid to the presence of eosinophilic airway inflammation in stable COPD, although reports of smaller numbers of patients with COPD of comparable severity have noted similar group mean induced sputum eosinophil counts.[16,22]

The origin of eosinophilic airway inflammation in COPD is unclear, although the presence of an asthmatic component to the fixed-airways obstruction is assumed.[23] An asthmatic component was unlikely in our population since we rigorously excluded patients who had variable airflow obstruction and clinical features suggesting asthma. Furthermore features such as a blood eosinophilia, atopy, smoking history, and physiological evidence of emphysema were no more or less common in patients in the highest eosinophilic tertile. It is more likely that smoking[24] and other mechanisms that recruit neutrophils into the airway mucosa in COPD cause a degree of eosinophil influx. In some patients this influx might amount to functionally important eosinophilic airway inflammation. Explanation of the high sputum eosinophil counts observed in some of our patients is, however, difficult. One possibility is that eosinophilic COPD starts as eosinophilic bronchitis. This disorder is a common cause of chronic cough in middle age, characterised by a sputum eosinophilia but no symptoms and functional evidence of variable airflow obstruction or airway hyper-responsiveness.[25] Although characterised by normal spirometric values at the time of diagnosis, eosinophilic bronchitis has been associated with an accelerated decline in FEV_1 and the development of COPD.[26]

There is currently much interest in the role of long-term inhaled corticosteroid treatment in COPD.[1,4] Investigators, in two large, placebo-controlled trials, have shown a small improvement in FEV_1 in the first 3 months of treatment;[27,28] such improvement was not seen in a third study in which patients with a positive oral corticosteroid response were excluded.[29] Treatment did not modify the subsequent rate of decline in FEV_1 over 3 years in any of those studies, but in one, 500 µg inhaled fluticasone twice daily lowered the exacerbation frequency, especially in patients who had more severe disease.[27] The important remaining question is whether these minor clinical benefits are confined to a definable subgroup of patients. Although we cannot discount that corticosteroids might have additional or different effects when given for long periods or via the inhaled route, our findings raise the possibility that a simple sputum test might allow treatment to be targeted to a population who would benefit especially. Further studies to investigate the role of induced sputum as a predictor of long-term response to inhaled corticosteroids are a priority.

Contributors

Chris Brightling assisted with the protocol development, undertook study recruitment and assessments, did the analysis, and prepared the paper. Will Monteiro, Richard Ward, and Debbie Parker analysed the induced sputum. Andrew Wardlaw and Michael Morgan assisted with protocol development and recruitment of patients. Ian Pavord conceived, designed, and obtained funding for the study, and supervised the clinical and laboratory assesments, study analysis, and preparation of the paper.

Acknowledgments

This study was funded by a grant from Trent Regional Research Scheme. We thank the Pharmacy Department, Glenfield Hospital for randomisation of patients and distribution of medicines, and Astra-Zeneca, Charnwood, for the measurement of eosinophilic cationic protein.

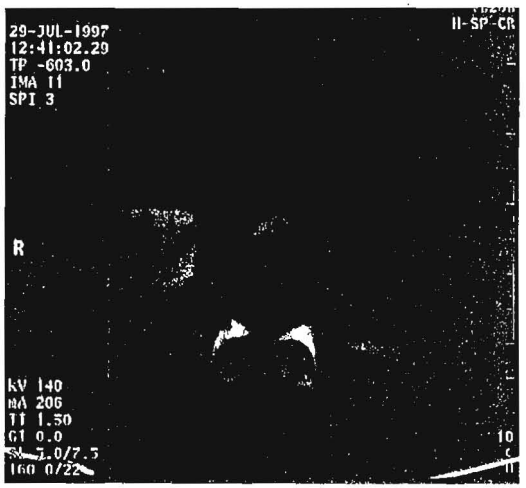
References

1 Boushey HA. Glucocorticoid therapy for chronic obstructive pulmonary disease. *N Engl J Med* 1999; 340: 1090–91.
2 British Thoracic Society. Guidelines for the management of chronic obstructive pulmonary disease. *Thorax* 1997; 52 (suppl 5): 9.
3 Calverley PMA Re-assessing the evidence about inhaled corticosteroids in chronic obstructive pulmonary disease *Thorax* 1999; 54: 3–4
4 Pavord ID, Brightling CE, Woltmann G, Wardlaw AJ Non-eosinophilic corticosteroid unresponsive asthma. *Lancet* 1999; 353: 2213–14.
5 Keatings VM, Jatakanon A, Worsdell YM, Barnes PJ Effects of inhaled and oral glucocorticoids on inflammatory indices in asthma and COPD. *Am J Respir Crit Care Med* 1997; 155: 542–48.
6 Shim C, Stover GE, Williams MP. Response to corticosteroids in chronic bronchitis. *J Allergy Clin Immunol* 1978; 62: 363–67.
7 Clifton M. Steroid therapy in chronic bronchitis. *Lancet* 1962, i 1311–13
8 Pizzichini E, Pizzichini MMM, Efthimiadis A, et al. Indices of airway inflammation in induced sputum: reproducibility and validity of cell and fluid phase measurements. *Am J Respir Crit Care Med* 1996, 154: 808–17.
9 Hunter CJ, Ward R, Woltmann G, Wardlaw AJ, Pavord ID. The safety and success rate of sputum induction using a low output ultrasonic nebuliser. *Respir Med* 1999; 93: 345–48.
10 Pizzichini E, Pizzichini MMM, Gibson P, et al. Sputum eosinophilia predicts benefit from prednisone in smokers with chronic obstructive bronchitis. *Am J Respir Crit Care Med* 1998; 158: 1511–17.
11 Weir DC, Gove RJ, Robertson AS, Burge PS. Corticosteroid trials in non-asthmatic chronic airflow obstruction: a comparison of oral prednisolone and inhaled beclomethasone dipropionate. *Thorax* 1990; 45: 112–17.
12 Guyatt GH, Berman LB, Townsend M, Pugsley S, Chambers L. A measure of quality of life for clinical trials in chronic lung disease *Thorax* 1987; 136: 1285–98.
13 Singh SJ, Morgan MDL, Scott S, Walters A, Hardman AE. The development of the shuttle walking test of disability in patients with chronic airways obstruction. *Thorax* 1992; 47: 1019–24.
14 Pavord ID, Pizzichini MMM, Pizzichini E, Hargreave FE The use of induced sputum to investigate airway inflammation. *Thorax* 1997; 52: 498–501.
15 Janoff A, Raju L, Dearing R. Levels of elastase activity in bronchoalveolar lavage fluids of healthy smokers and nonsmokers *Am Rev Respir Dis* 1983; 127: 540–44.
16 Tweeddale PM, Alexander F, McHardy GJR. Short term variability in FEV, and bronchodilator responsiveness in patients with obstructive ventilatory defects. *Thorax* 1987, 42: 87–90.
17 Chanez P, Vignola AM, O'Shaugnessy T, et al. Corticosteroid reversibility in COPD is related to features of asthma. *Am J Respir Crit Care Med* 1997; 155: 1529–34.
18 Confalonieri M, Mainardi E, Della Porta R, et al. Inhaled corticosteroids reduce neutrophilic bronchial inflammation in patients with chronic obstructive pulmonary disease. *Thorax* 1998; 53: 583–85.
19 Keatings VM, Collins PD, Scott DM, Barnes PJ Differences in interleukin-8 and tumor necrosis factor-α in induced sputum from patients with chronic obstructive pulmonary disease or asthma. *Am J Respir Crit Care Med* 1996; 153: 530–34.
20 Di Stefano A, Capelli A, Lusuardi M, et al. Severity of airflow limitation is associated with severity of airway inflammation in smokers. *Am J Respir Crit Care Med* 1998; 158: 1277–85.
21 Stanescu D, Sanna A, Veriter C, et al. Airways obstruction, chronic expectoration, and rapid decline of FEV_1 in smokers are associated with increased levels of sputum neutrophils. *Thorax* 1996; 51: 267–71
22 Saetta M, Stefano AD, Maestrelli P, et al. Airway eosinophilia in chronic bronchitis during exacerbations. *Am J Respir Crit Care Med* 1994; 150: 1646–52.
23 Barnes NC. Inhaled steroids in COPD *Lancet* 1998; 351: 706–07.
24 Matsumoto K, Aizawa H, Inoue H, et al. Eosinophilic airway inflammation induced by repeated exposure to cigarette smoke. *Eur Respir J* 1998; 12: 387–94.
25 Brightling CE, Ward R, Goh KL, Wardlaw AJ, Pavord ID. Eosinophilic bronchitis is an important cause of cough. *Am J Respir Crit Care Med* 1999; 160: 406–10.
26 Brightling CE, Woltmann G, Wardlaw AJ, Pavord ID The development of irreversible airflow obstruction in a patient with eosinophilic bronchitis without asthma. *Eur Respir J* 1999; 14: 1228–30
27 Burge PS, Calverley PM, Jones PW, Spencer S, Anderson JA, Maslen TK. Randomised, double blind, placebo controlled study of fluticasone propionate in patients with moderate to severe chronic obstructive pulmonary disease: the ISOLDE trial *BMJ* 2000; 320: 1297–303.
28 Pauwels RA, Lofdahl CG, Laitinen LA, et al. Long-term treatment with inhaled budesonide in persons with mild chronic obstructive pulmonary disease who continue smoking. *N Engl J Med* 1999; 340: 1948–53
29 Vestbo J, Sorensen T, Lange P, Brix A, Torre P, Viskum K. Long-term effect of inhaled budesonide in mild and moderate chronic obstructive pulmonary disease; a randomised controlled trial. *Lancet* 1999; 353: 1819–23.

Clinical picture

Migrating lumbar vertebra

Dheeraj Gandhi, Raju Sharma, P P Kotwal, Pradeep Hatimota



A 22-year-old woman was admitted to hospital with low backache and intermittent low-grade fever. Laboratory investigations revealed mild anaemia, an increased erythrocyte sedimentation rate, and a positive Mantoux test. Radiographs of the lumbo-sacral spine showed the body of fourth lumbar (L4) vertebra floating in the right paravertebral gutter. Contrast-enhanced computed tomography scan revealed a right-sided psoas abscess containing the migrated L4 vertebral body (figure). The posterior elements of L4 vertebra were identified in their normal location but there was destruction of bilateral pedicles. The patient underwent surgery and debridement of granulation tissue, sequestrectomy, and vertebral cage fixation. Histopathology of the removed vertebra revealed tuberculosis. Pathological fracture through the diseased pedicles allowed the vertebral body to migrate laterally into the psoas abscess.

Department of Radiodiagnosis, All India Institute of Medical Sciences, New Delhi, India (D Gandhi MD, R Sharma MD, P Hatimota MD); **and Department of Orthopaedics, All India Institute of Medical Sciences** (P Kotwal MS)

(212) [240]

The Effect of Two Different Dosages of Intravenous Immunoglobulin on the Incidence of Recurrent Infections in Patients with Primary Hypogammaglobulinemia

A Randomized, Double-Blind, Multicenter Crossover Trial

Heleen W. Eijkhout, MD; Jos W.M. van der Meer, MD, PhD; Cees G.M. Kallenberg, MD, PhD; Ron S. Weening, MD, PhD; Jaap T. van Dissel, MD, PhD; Lieke A.M. Sanders, MD, PhD; Paul F.W. Strengers, MD; Henriët Nienhuis, MSc; and Peter Th.A. Schellekens, MD, PhD, for the Inter-University Working Party for the Study of Immune Deficiencies

Background: In patients with hypogammaglobulinemia, substitution with immunoglobulin is the treatment of choice to reduce both frequency and severity of bacterial infections. Even with treatment, however, infections still occur in these patients.

Objective: To determine whether doubling the standard dose of intravenous immunoglobulin would affect the incidence and duration of infections.

Design: Multicenter, double-blind, randomized, crossover study.

Setting: 15 outpatient clinics in the Netherlands.

Patients: 43 patients with primary hypogammaglobulinemia, 41 of whom completed the protocol.

Intervention: Patients received standard-dose immunoglobulin therapy for 9 months, followed by a 3-month washout period, and high-dose intravenous immunoglobulin therapy for 9 months, or vice versa.

Measurements: The primary outcome measures were total number and duration of infections. Other measures were periods of fever, hospital admissions, use of antibiotics, absence from school or work, and trough levels of serum immunoglobulin. Side effects from the study medication were also recorded.

Results: Compared with the standard dose of intravenous immunoglobulin (adults, 300 mg/kg of body weight every 4 weeks; children, 400 mg/kg every 4 weeks), high-dose therapy (adults, 600 mg/kg every 4 weeks; children, 800 mg/kg every 4 weeks) significantly reduced the number (3.5 vs. 2.5 per patient; $P = 0.004$) and duration (median, 33 days vs. 21 days; $P = 0.015$) of infections. Trough levels of IgG increased significantly during high-dose therapy. The incidence and type of side effects did not differ significantly for the two dosages.

Conclusion: In patients with hypogammaglobulinemia, doubling the standard dose of intravenous immunoglobulin significantly reduced the number and duration of infections.

Ann Intern Med. 2001;135:165-174. www.annals.org

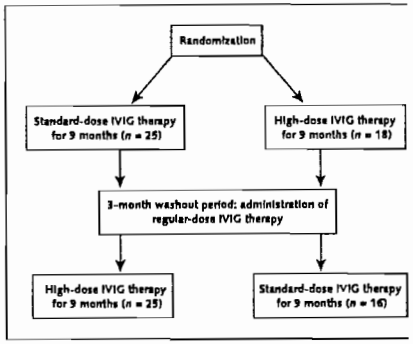
For author affiliations, current addresses, and contributions, see end of text.

Patients with primary hypogammaglobulinemia, such as X-linked agammaglobulinemia and common variable immunodeficiency, have recurrent infections, predominantly of the respiratory and intestinal tract (1, 2). Most respiratory infections in these patients are caused by *Haemophilus influenzae* and *Streptococcus pneumoniae* and, without proper treatment, may lead to severe pneumonia, bronchiectasis, decreased pulmonary function, and death (1, 3–6). Recurrent *Giardia lamblia* infections may result in chronic diarrhea, whereas chronic *Campylobacter jejuni* infections may cause recurrent bacteremia and cellulitis (1, 3, 4, 7, 8). In patients with X-linked agammaglobulinemia, persistent enterovirus infections, notably those caused by echovirus, are associated with chronic meningoencephalitis (3, 6).

Since immunoglobulin therapy was introduced for the treatment of immunodeficiency diseases, the frequency and severity of infections have decreased (1, 9, 10). However, repeated long-term use of intramuscular immunoglobulin has serious limitations. Injections cause pain at the injection site, and their volume is limited, in turn limiting the prospective increase in IgG level (11). With subcutaneous infusion of immunoglobulin, normal IgG concentrations can be achieved. In most patients, however, local tissue reactions, such as swelling, induration, and soreness, are observed; long-term subcutaneous infusion can result in fibrosis at the injection site.

These disadvantages were overcome when intravenous immunoglobulin preparations became available in the 1980s (12). Several studies have compared the clinical efficacy of various immunoglobulin products and dosage regimens in patients with primary hypogammaglobulinemia. Results have shown that intravenous immunoglobulin, when compared with subcutaneous or intramuscular administration, reduces the incidence of infections because it offers increased bioavailability and allows administration of higher doses (5, 12, 13). Furthermore, it has been suggested that when the serum IgG trough level exceeds 5 g/L, protection against bacterial infections is improved (12, 14). However, even in patients who have trough IgG levels greater than 5 g/L during replacement therapy, infections continue to be present. We performed this crossover study to determine whether doubling the commonly advised (standard) dose of intravenous immunoglobulin (300 mg/kg of body weight every 4 weeks in adults, 400 mg/kg every 4 weeks in children) decreases the incidence and duration of infections.

Figure 1. Flow of patients through the study.



IVIG = intravenous immunoglobulin.

METHODS

Patients

Between September 1995 and February 1998, 46 patients with established humoral primary immunodeficiency who had X-linked agammaglobulinemia and common variable immunodeficiency (as defined by the World Health Organization) and an IgG trough level of 4 g/L or less at the time of diagnosis were studied in 15 hospitals in the Netherlands (15). They represented 40% to 50% of the total number of patients with X-linked agammaglobulinemia and common variable immunodeficiency in the country. Exclusion criteria were age younger than 1 year; anti-IgA antibodies; chronic active diseases, such as hepatitis, AIDS, and malignant conditions; history of anaphylactic reactions to intravenous immunoglobulin; and participation in a clinical trial 3 months before the start of the study. Thirteen of the patients in the study group (28%) had clinical and radiographic evidence of preexisting chronic bronchopulmonary disease. Before the start of the study, all patients received regular replacement therapy with intravenous immunoglobulin, with the exception of four patients who were treated with subcutaneously administered 16% immunoglobulin.

The ethics committee of each participating hospital approved the study protocol. Before enrollment in the study, all patients or their legal representatives provided written informed consent.

Study Design and Treatment Protocol

The study was a multicenter, double-blind, randomized, crossover trial. After providing written informed consent, patients were randomly divided into two groups according to a computer-generated randomization list. During the first 9 months of the study, one group was treated with standard-dose intravenous immunoglobulin followed by a 3-month washout period, during which patients received the dose of intravenous immunoglobulin used before the study began. Patients were then treated with high-dose immunoglobulin for 9 months. In the second group, the treatment sequence was reversed (Figure 1). The time schedule was chosen to prevent seasonal variations from influencing the infection rate.

To ensure that patients, nurses, and physicians remained unaware of the dosages of intravenous immunoglobulin, each hospital pharmacy provided the study product after dissolving the freeze-dried immunoglobulin in the required volume of sterile water. An equal volume of sodium chloride (NaCl 0.9%) was added to the solution of the standard dose immunoglobulin to mimic the high-dose volume. The two preparations did not differ in appearance.

Intravenous immunoglobulin (Immunoglobuline I.V., Sanquin Blood Supply Foundation, Amsterdam, the Netherlands) was manufactured from the pooled plasma of at least 1000 voluntary, nonremunerated donors by using cold ethanol fractionation according to the Cohn method. At the end of the purification process, immunoglobulin was treated by using mild prote-

olysis at a pH of 4. After the freeze-dried product is reconstituted with water, Immunoglobuline I.V. is composed of ar least 95% IgG and 0.24 mol of glucose per L as stabilizer. The distribution of IgG subclasses in the product (by mean percentage [±SD]) is as follows: IgG1, 61.9% ± 4.8%; IgG2, 32.6% ± 4.8%; IgG3, 2.1% ± 0.4%; and IgG4, 3.3% ± 0.4%. The product contains only small amounts of IgA and IgM.

The standard dose of intravenous immunoglobulin was 300 mg/kg every 4 weeks for adults and 400 mg/kg every 4 weeks for children (those ≤20 years of age). High-dose therapy was 600 mg/kg every 4 weeks for adults and 800 mg/kg every 4 weeks for children. For patients who were receiving regular treatment (one infusion every 2 to 3 weeks) before the start of the study, the dose and frequency were adjusted to ensure that a similar total quantity of intravenous immunoglobulin was received.

Follow-up and Outcome Measures

Before each infusion of intravenous immunoglobulin, we reviewed each patient's health status using his or her diary, medical record, and an interview. We recorded previous hospital admissions (diagnosis, frequency, and duration); number, type, and duration of infections (as defined by the Infectious Diseases Society of America [16]); prophylactic and therapeutic use of antibiotics; febrile periods (temperature ≥ 38.5 °C); and absence from school or work. Duration of infection was determined from the number of days with symptoms. The attending physician recorded whether infections were considered to be associated with immunodeficiency.

We measured the peak expiratory flow rate at the onset of each study period, as well as 6 and 9 months thereafter, because previous studies have shown that lung function can improve in patients with hypogammaglobulinemia who receive higher doses of intravenous immunoglobulin (14, 17). During and after each infusion of intravenous immunoglobulin, vital signs (blood pressure, temperature, pulse rate) were measured and side effects were monitored.

Laboratory Analysis

Trough levels of IgG: levels of subclasses IgG, IgA, and IgM; levels of antibodies against a variety of microorganisms; and safety measures such as blood cell count and chemistry values (for example, kidney and liver function) were determined immediately before the intravenous immunoglobulin infusion at the onset of each study period and at 6 and 9 months. We quantified levels of immunoglobulins and IgG subclasses by nephelometry and levels of antibodies to microorganisms by specific enzyme-linked immunosorbent assays. If possible, sputum and stools for culture were obtained at the same time periods. The trough levels of both immunoglobulins and levels of antibodies to microorganisms were withheld from each patient's physician during the study period.

Statistical Analysis

The primary end point, occurrence of infection, was used to calculate the required sample size. From the results of the study by Bernatowska and colleagues (17), we assumed that infections would develop at a rate of approximately 200 days per 1000 observation-days during standard-dose treatment, compared with 180 days per 1000 observation-days during high-dose treatment. This implied that a minimum of 30 patients was required to allow us to detect a 10% difference between the two groups (17).

The final analysis was conducted on an intention-to-treat basis. All patients were included in the study, and those who had at least one efficacy result while receiving high-dose therapy as well as standard-dose therapy were analyzed. Before the analysis of efficacy, the data were investigated for a carryover effect by using the sequence of treatment (high-dose vs. standard-dose and standard-dose vs. high-dose) as one of the factors and the primary variable of efficacy (number of infections related to immunodeficiency) as the dependent variable. Thus, we created a multivariate model in which the number of infections was the dependent variable and dosage, sequence of treatment, and patient were the independent variables. We then determined whether the sequence of treatment was statistically significant.

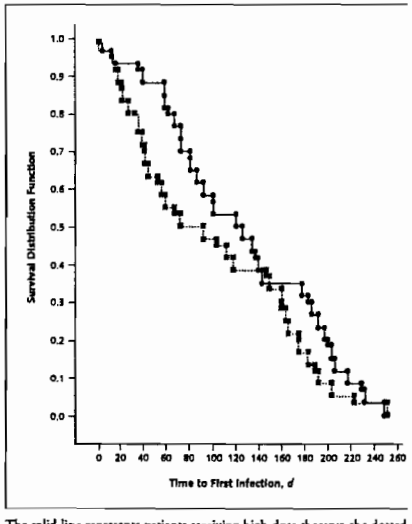
The total number of acute infections was compared by using the paired *t*-test. In addition, the total duration of all infections per patient and per dosing for all patients was computed. These results were compared by using the paired *t*-test, as were differences in time period from the start of treatment until occurrence of first infection. For the primary efficacy variables, we conducted subanalyses for both diagnoses (X-linked agammaglobulinemia vs. common variable immunodeficiency) and for age (adults vs. children).

Secondary end points were the total number of hospital admissions, days missed from school or work, frequency and duration of antibiotic courses, and total number of days with fever. In addition, trough serum levels of immunoglobulins and the results of the cultures were compared. Most results are expressed as the mean (±SD), but some results are expressed as the median because the data were relatively skewed. All reported *P* values are two tailed. A *P* value less than 0.05 was considered statistically significant. We used SAS/STAT software (SAS Institute, Inc., Cary, North Carolina) for all calculations.

Role of the Funding Source

The funding source had no role in the collection, analysis, or interpretation of the data or in the decision to submit the manuscript for publication.

Figure 2. Kaplan–Meier estimate of survivor function.



The solid line represents patients receiving high-dose therapy; the dotted line represents patients receiving standard-dose therapy.

RESULTS

Of the 46 patients originally enrolled in the study, 3 withdrew informed consent before the first infusion. Of the remaining 43, 24 had common variable immunodeficiency (16 women, 8 men) and 19 had X-linked agammaglobulinemia. Ages varied from 1.6 to 70.3 years (mean age, 29.9 years). Twenty-five patients received infusions every 3 weeks, 10 patients were treated every 2 weeks, and 8 patients were treated every 4 weeks. All but 2 patients completed both study periods. One patient withdrew from the study because of side effects (headache and nausea) after nine infusions of high-dose intravenous immunoglobulin during the first study period. The other patient dropped out after the first study period, during which he received high-dose intravenous immunoglobulin, because of severe illness and poor life expectancy after aspiration pneumonia.

The mean of the total study period (including the 3-month washout period) for each patient was 562 days: 230 days with standard-dose intravenous immunoglobulin and 233 days with high-dose immunoglobulin. A total of 1057 infusions were given (518 standard-dose, 539 high-dose). Statistical analysis showed that the 3-month washout period was long enough to prevent a carryover effect from high-dose intravenous immunoglobulin to standard-dose treatment, and vice versa ($P > 0.2$).

Primary Variables

The total number of infections was 135 during standard-dose treatment and 103 during high-dose treatment. Four infections (1 during standard-dose treatment and 3 during high-dose treatment) involved contributing conditions other than immunodeficiency, such as trauma, and were not included in the analysis.

Infections were predominantly upper respiratory tract infections and bronchitis. Other infections included pneumonia, urinary infections, otitis media, conjunctivitis, diarrhea, and cellulitis. During treatment with standard-dose intravenous immunoglobulin, one or more infections were observed in 37 of 41 (90.2%) patients, compared with 36 of 43 (83.7%) patients treated with high-dose intravenous immunoglobulin (Table 1). The mean number of infections associated with immunodeficiency was 3.5 ± 2.6 per patient during standard-dose therapy and 2.5 ± 2.4 per patient

Table 1. Infections in Patients with Primary Hypogammaglobulinemia Treated with Two Different Dosages of Intravenous Immunoglobulin

Variable	Standard-Dose Therapy (n = 41)	High-Dose Therapy (n = 43)	Difference (95% CI)	P Value
Patients with infections, n	37	36		
Total infections related to immunodeficiency, n*	134	100		
Mild	54	38		
Moderate	17	11		
Severe	63	51		
Mean total immunodeficiency-related infections per patient ± SD (95% CI), n†	3.5 ± 2.6 (2.7–4.3)	2.5 ± 2.4 (1.8–3.2)	1.1 (0.4 to 1.8)	0.004
Median duration of immunodeficiency-associated infections (range), d†	33 (1–185)	21 (1–125)		0.015
Total respiratory infections, n	61	50		
Mean respiratory infections per patient ± SD (95% CI), n†	1.5 ± 1.6 (1.0–2.0)	1.2 ± 1.7 (0.7–1.7)	0.46 (−0.18 to 0.78)	0.18
Median duration of respiratory infections (range), d†	29 (5–178)	22 (2–125)		0.16

* Type of infection was categorized according to the criteria described in reference 18.
† Results are based on data from 41 patients. Two patients did not receive standard-dose intravenous immunoglobulin; therefore, no results were obtained for them during their respective standard-dose periods. Results were compared by using the paired *t*-test.

during high-dose therapy ($P = 0.004$). Sixty-three of the 134 infections (47.0%) that occurred during standard-dose treatment were reported as severe (that is, respiratory infections, cellulitis, and sepsis or an infection that resulted in hospital admission) compared with 51 of 100 infections (51%) that occurred during high-dose treatment (Table 1) (18). Although the number of respiratory infections tended to decrease during high-dose therapy, the difference from standard-dose therapy was not statistically significant ($P = 0.18$). It can therefore be concluded that the number of the infections decreased proportionally regardless of infection severity. The 13 patients with preexisting bronchopulmonary disease did not differ from other patients.

Median duration of infection was significantly shorter during high-dose therapy than during standard-dose therapy (21 days [range, 1 to 125 days] vs. 33 days [range, 1 to 185 days]; $P = 0.015$). The estimated median time from the start of the study period until occurrence of the first infection was 81.5 days during standard-dose therapy and 123 days during infusion of high-dose intravenous immunoglobulin (Figure 2). The mean time until occurrence of a second, third, and fourth infection was also longer with high-dose therapy than with standard-dose therapy. The number and duration of infections were similar for adults and children, as well as for patients with X-linked agammaglobulinemia and those with common variable immunodeficiency (Table 2). In addition, the frequency of infusions (two, three, or four per week) did not influence the number of infections per patient (analysis of variance; $P = 0.12$).

Table 2. Subgroup Analysis of Number and Duration of Infections during Standard-Dose and High-Dose Intravenous Immunoglobulin Therapy*

Variable	Standard-Dose Therapy (n = 41)	High-Dose Therapy (n = 43)
Patients with recurrent infections, n		
Adults	23	21
Children	14	15
Patients with X-linked agammaglobulinemia	16	14
Patients with common variable immunodeficiency	21	22
Infections related to immunodeficiency, n		
Adults	72	50
Children	62	50
Patients with X-linked agammaglobulinemia	65	44
Patients with common variable immunodeficiency	69	56
Mean immunodeficiency-related infections per patient ± SD, n		
Adults	3.5 ± 2.0	2.7 ± 2.0
Children	4.6 ± 2.9	3.4 ± 2.6
Patients with X-linked agammaglobulinemia	4.3 ± 3.0	3.4 ± 2.8
Patients with common variable immunodeficiency	3.6 ± 1.8	2.7 ± 1.9
Median duration of immunodeficiency-related infections (range), d		
Adults	35 (2–185)	27 (2–125)
Children	31 (1–79)	12 (1–45)
Patients with X-linked agammaglobulinemia	28.5 (1–185)	19 (2–51)
Patients with common variable immunodeficiency	35 (2–155)	24 (1–125)

* Based on data from 25 adults and 18 children. Nineteen patients had X-linked agammaglobulinemia, and 24 patients had common variable immunodeficiency.

Secondary Variables

During standard-dose treatment, five patients were admitted to the hospital eight times. During high-dose treatment, four patients were admitted seven times (Table 3). All of these patients were admitted to the hospital because of an infection. Among all patients, we observed no differences in the median duration of hospital admission (19.0 days for standard-dose therapy vs. 20.5 days for high-dose therapy), frequency of absence from school or work (0.8 ± 1.2 periods per patient for both treatments), or the number of days absent from work or school (median, 17.5 days for standard-dose therapy vs. 12.5 days for high-dose therapy) (Table 3).

Seventy-one periods of fever were reported, 39 in 23 patients during standard-dose therapy and 32 in 18 patients during high-dose therapy. The mean number of febrile periods for all included patients was lower during high-dose therapy than during standard-dose therapy (0.7 vs. 1.0) (Table 3).

Peak expiratory flow rate was measured during both study periods in all but four patients (children ≤ 6 years of age). Compared with the results at the start of the study period, the peak expiratory flow rate after 9 months of therapy increased by 11.4 L/min during standard-dose therapy and 37.3 L/min during high-dose therapy. However, these increases were not significant. No significant difference was seen between groups ($P > 0.2$). Eleven patients used prophylactic antibiotics,

Table 3. Results of Secondary Variables during Standard-Dose and High-Dose Intravenous Immunoglobulin Therapy

Variable	Standard-Dose Therapy (n = 41)	High-Dose Therapy (n = 43)
Patients who were admitted to the hospital, n	5	4
Hospital admissions, n	8	7
Median duration of hospital admission (range), d	19.0 (5–24)	20.5 (9–50)
Mean periods absent from work or school per patient, n	0.8	0.8
Median days absent from work or school (range), n	17.5 (1.0–38.0)	12.5 (2.0–67.0)
Patients who had a febrile period, n	23	18
Febrile periods, n	39	32
Mean febrile periods for all included patients, n	1.0	0.7
Patients using prophylactic antibiotics, n	11	11
Patients using therapeutic antibiotics, n	32	26

Table 4. Serum Immunoglobulin Trough Levels during Standard-Dose and High-Dose Intravenous Immunoglobulin Therapy*

Type of Immunoglobulin	Level at Onset		Level at Month 6†		Level at Month 9†	
	Standard-Dose Therapy	High-Dose Therapy	Standard-Dose Therapy	High-Dose Therapy	Standard-Dose Therapy	High-Dose Therapy
	g/L					
Total IgG	6.5 ± 1.7	6.3 ± 1.6	6.4 ± 1.6	9.0 ± 2.1	6.6 ± 1.6	9.4 ± 2.7
IgG1	3.9 ± 1.3	3.8 ± 1.0	3.9 ± 1.4	5.3 ± 1.7	4.0 ± 1.4	5.4 ± 1.9
IgG2	2.4 ± 0.8	2.4 ± 0.6	2.3 ± 0.6	3.5 ± 0.9	2.5 ± 0.5	3.7 ± 1.0
IgG3	0.14 ± 0.12	0.12 ± 0.08	0.13 ± 0.12	0.14 ± 0.07	0.14 ± 0.11	0.14 ± 0.07
IgG4	0.19 ± 0.08	0.19 ± 0.08	0.18 ± 0.08	0.25 ± 0.08	0.19 ± 0.08	0.25 ± 0.08
IgA	0.07 ± 0.21	0.08 ± 0.26	0.09 ± 0.24	0.09 ± 0.24	0.09 ± 0.26	0.08 ± 0.25
IgM	0.3 ± 0.6	0.3 ± 0.7	0.3 ± 0.7	0.3 ± 0.6	0.3 ± 0.6	0.2 ± 0.4

* Values are the mean ± SD.
† The difference in IgG trough levels was statistically significant ($P = 0.001$)

Table 5. Median Trough Levels of Antibodies to Several Microorganisms

Microorganism	Level at Onset		Level at Month 6		Level at Month 9	
	Standard-Dose Therapy	High-Dose Therapy	Standard-Dose Therapy	High-Dose Therapy	Standard-Dose Therapy	High-Dose Therapy
Haemophilus influenzae type B (range), μg/mL	2.2 (1–9)	1.9 (1–9)	2.2 (1–8)	2.9 (1–10)	1.8 (1–7)	2.9 (1–17)
Nonencapsulated *H. influenzae* (range), μg/mL	6.0 (3–23)	5.5 (3–24)	6.0 (2–20)	7.3 (4–30)	5.3 (2–15)	8.4 (4–39)
Streptococcus pneumoniae type 3 (range), IU/mL	55.4 (20–393)	56.8 (23–445)	69.7 (24–453)	92.0 (40–660)	67.8 (27–403)	110.5 (41–667)
S. pneumoniae type 9 (range), IU/mL	37.2 (17–346)	37.8 (18–474)	41.2 (16–334)	65.0 (33–660)	41.1 (23–379)	68.4 (30–659)
S. pneumoniae type 14 (range), IU/mL	45.8 (18–744)	43.6 (23–815)	49.8 (21–734)	87.5 (44–1158)	48.7 (30–785)	86.8 (38–1474)

mostly cotrimoxazole, during both study periods. In addition, 32 patients received antibiotics for treatment of acute infections during standard-dose therapy and 26 patients did so during high-dose therapy; however, this is not a staristically significant difference. The mean number of antibiotic courses per patient was 2.5 ± 2.2 during standard-dose therapy and 1.8 ± 2.5 during high-dose therapy.

At the start of each study period, the IgG trough level was similar in both groups. The IgG trough level increased from 6.3 g/L to 9.4 g/L ($P < 0.001$) during high-dose therapy but did not change during standard-dose therapy. The difference in IgG trough levels between the two dosages at month 6 and month 9 was statistically significant ($P = 0.001$) (Table 4). During high-dose therapy, the mean level of subclass IgG1 increased from 3.8 g/L to 5.4 g/L and the mean level of IgG2 increased from 2.4 g/L to 3.7 g/L (Table 4). No changes were observed in the trough levels of IgG3, IgG4, IgA, or IgM during either treatment. During high-dose treatment, we found a significant positive correlation between the trough level of IgG and trough levels of antibodies to *H. influenzae* ($r = 0.42$; $P < 0.001$), *H. influenzae* type B ($r = 0.37$; $P < 0.001$), and *S. pneumoniae* type 3 ($r = 0.19$; $P = 0.003$) and type 9 ($r = 0.14$; $P = 0.028$) (Tables 4 and 5).

The results of routine sputum and stool cultures did not differ with respect to the number of positive cultures and the types of microorganisms found (in sputum, mostly *H. influenzae*, and in stools, mostly *C. jejuni* and *C. lari*) during standard-dose or high-dose therapy. Cultures were positive in 14 of 19 patients receiving standard-dose therapy and in 8 of 17 patients receiving high-dose therapy. The microorganisms most frequently cultured during infectious periods were *H. influenzae* and *S. pneumoniae*.

Side effects of intravenous immunoglobulin were documented in 10 of 41 patients (24%) during or directly after 23 of 518 standard-dose infusions (4.4%) and in 13 of 43 patients (30%) during or directly after 35 of 539 high-dose infusions (6.5%). Side effects did not differ significantly between the two groups. Headache and fever were most common, especially during high-dose treatment in one specific patient. In this patienr, treatment had ro be discontinued because of the severity of these side effects after infusion of high-dose therapy. When we adjusted for the fact that more patients received more infusions of intravenous immunoglobulin during the high-dose periods, no relation between the

Table 6. Adverse Events Related to Standard-Dose or High-Dose Immunoglobulin Therapy

Variable	Standard-Dose Therapy	High-Dose Therapy
Patients, *n*	41	43
Infusions, *n*	518	539
Infusions with adverse events, n	23	35
Total adverse events, n	36	51
Patients with adverse events	10	13
Type of adverse event		
Headache	7	14*
Fever	8	8
Malaise	2	2
Backache	4	3
Nausea	0	5
Perspiration	3	1
Chills	1	3
Tachycardia	0	1
Dyspnea	0	1
Other	11	13

* One patient reported headache after receiving eight of nine infusions of high-dose intravenous immunoglobulin. This patient dropped out of the second part of the study because of this side effect.

type of the reactions and the dose could be demonstrated (Table 6). No signs of meningeal irritation, hemolysis, neutropenia, or renal and liver abnormalities were found to be related to intravenous immunoglobulin therapy in either study period.

Discussion

In this double-blind, randomized, crossover study, we found statistically significant reductions in the number ($P = 0.004$) and duration ($P = 0.015$) of infections when high-dose intravenous immunoglobulin (adults, 600 mg/kg every 4 weeks; children, 800 mg/kg every 4 weeks) was compared with standard-dose intravenous immunoglobulin (adults, 300 mg/kg every 4 weeks; children, 400 mg/kg every 4 weeks). These differences were found in patients with common variable immunodeficiency, in those with X-linked agammaglobulinemia, and in adults as well as children. Our study sample represented 40% to 50% of the total number of patients with these disorders in the Netherlands. The dosage frequency (two, three, or four times weekly) did not seem to influence the results. In addition, the time from the start of immunoglobulin treatment until occurrence of the first, second, third, and fourth infection was longer when high-dose intravenous immunoglobulin was infused. Although slightly more nonsevere adverse effects were observed during high-dose treatment, none of the side effects were dose related.

Previous studies have shown a relationship between the amount of immunoglobulin administered and the infection rate. None, however, fulfilled the stringent criteria applied in our study (12–14, 18). In addition, some studies used different immunoglobulin preparations (intramuscular vs. intravenous immunoglobulin) to compare efficacy or included only a relatively small number of patients (12, 13, 18, 19). In the parallel study by Ochs and colleagues (20), which compared intravenous immunoglobulin dosages of 100 mg/kg every 4 weeks and 400 mg/kg every 4 weeks, patients with humoral immunodeficiency who received high doses showed no significant improvements in the number of infections, use of antibiotics, or the number of days absent from work or school. Comparison of this study with our trial is hampered by difference in study design and dosage regimen. We chose a crossover design, in which each patient serves as his own control, to eliminate between-subject variability.

Pruzanski and associates (18) performed a randomized, double-blind, crossover, dose-assessing study in 21 patients with common variable immunodeficiency. They compared three dosages of intravenous immunoglobulin (200 vs. 400 vs. 600 mg/kg every 4 weeks) over a period of 6 months each and found that 400 mg/kg every 4 weeks tended to be more effective than 200 mg/kg every 4 weeks, although the differences were not statistically significant. Increasing the dose to 600 mg/kg every 4 weeks did not provide better protection. The study by Pruzanski and associates (18) has several methodologic limitations, including the relatively short observation period (6 months), which makes it possible that that seasonal differences may have played a role. Since a washout period between the dosages was not included, a carryover effect from high-dose to low-dose immunoglobulin therapy is also likely to have occurred (18).

In our study, we observed no differences between the two dosage regimens in the number of hospital admissions and in days absent from school or work. Most infections were treated in outpatient clinics, which explains the low number of hospital admissions. Our results concerning the duration of hospital admission and absence from work during high-dose intravenous immunoglobulin therapy were negatively influenced by one patient who had aspiration pneumonia for 67 days. This patient remained in the hospital for 50 of these 67 days and dropped out of the study after completing the first study period. Because our study had an intention-to-treat design, this patient was included in the analysis.

Studies by Roifman and coworkers (14) and Bernatowska and colleagues (17) have shown improvement in the pulmonary functioning of patients with hypogammaglobulinemia and chronic lung disease who were treated with high-dose intravenous immunoglobulin. The results of our study confirm that peak expiratory flow rate tended to increase more during high-dose treatment than during standard-dose treatment, although this difference was not statistically different. We hypothesize that when both the number and the duration of infections are reduced, morbidity, including the risk for bronchiectasis and respiratory failure, will improve. This hypothesis must be confirmed by future long-term prospective studies.

Patients with hypogammaglobulinemia are often

chronic carriers of *H. influenzae* and *S. pneumoniae* (21). In our study, positive sputum cultures were often found in patients with chronic bronchopulmonary disease. High-dose intravenous immunoglobulin did not decrease the number of positive sputum cultures, possibly because intravenous immunoglobulin does not contain sufficient amounts of IgA, the predominant immunoglobulin on mucosal surfaces that protects against bacterial colonization (11, 21). This may also explain why we observed no difference between the two dosages with respect to the presence of *Campylobacter* and *Giardia* species in the gastrointestinal tract.

In patients with primary hypogammaglobulinemia, it has been recommended that the IgG trough level should be maintained above 5 g/L to prevent recurrent infections (12, 14). Our results show that a mean IgG trough level of 9.4 g/L results in a significant difference in the number and duration of infections compared with a mean IgG trough level of 6.5 g/L.

Because intravenous immunoglobulin is expensive, a cost–benefit analysis should be done to determine the advantages of increasing immunoglobulin doses for all patients with hypogammaglobulinemia. A 70-kg patient receiving high-dose therapy would require 252 g of immunoglobulin instead of 126 g during a 6-month period, an additional cost of $5000. These extra costs for prevention of infections and, in the long term, bronchiectasis and respiratory failure should be balanced against the costs of repeated antibiotic treatment, hospital admissions, and absence from work.

Although our study demonstrates that high-dose intravenous immunoglobulin significantly reduces the number as well the duration of infections, we do not believe that this treatment regimen is routinely required, given the costs involved. In our study, an increased dose decreased incidence of infections in many but not all patients. In contrast, at a lower dose, some patients experienced fewer than two infections per year. Because IgG levels differed considerably among patients in the same treatment group, we advocate the use of IgG levels instead of dosage schemes. Therefore, we propose that patients with newly diagnosed primary hypogammaglobulinemia should first be treated with a standard dose of intravenous immunoglobulin (adults, 300 mg/kg every 4 weeks; children, 400 mg/kg every 4 weeks). When patients develop more than two severe infections per year despite this therapy, we recommend adjusting the therapy to increase the IgG level by 1 to 1.5 g/L. This cycle may be repeated until a level of approximately 9.5 g/L is reached. It remains to be seen whether an IgG level exceeding 9.5 g/L will contribute to reducing periods of infection. When increasing the dose does not decrease the incidence of infections, maintenance combination therapy with a lower dose of intravenous immunoglobulin and antibiotics may be required.

From CLB, Academic Medical Center, and Parexel Mirai, Amsterdam; University Hospital Nijmegen, Nijmegen; University Hospital Groningen, Groningen; Leiden University Medical Center, Leiden; and University Hospital for Children, Utrecht, the Netherlands.

Acknowledgments: The authors thank the internists, pulmonologists, and pediatricians who participated in the study; the nurses who provided medical care to the patients, the pharmacists who provided study medication, and the patients who participated. They also thank Professor W.G. van Aken for commenting on the manuscript; Machteld Tissing and Angelique van den Broek for data entry; and T.H. The, K. Brinkman, J.C.C. Borleffs, C.A.J.J. Jaspers, A.J.P. Veerman, P.J. van den Broek, M.C. Wallis-Spit, A.M. Hemmes, E.J.F.M. ten Berge, J.J.M. Peters, F.A.E. Nabben, L. Aurema, J.W. van 't Wout, and H.H.M. Hassing for provision of study materials or patients.

Grant Support: By Sanquin Blood Supply Foundation, Amsterdam, the Netherlands.

Requests for Single Reprints: Heleen W. Eijkhout, MD, CLB, Medical Department, Plesmanlaan 125, 1066 CX Amsterdam, the Netherlands; e-mail, H_Eijkhout@CLB.nl.

Current Author Addresses: Drs. Eijkhout and Strengers: CLB, Medical Department, Plesmanlaan 125, 1066 CX Amsterdam, the Netherlands.
Dr. van der Meer: Department of Medicine, University Hospital Nijmegen, Geert Grooteplein Zuid 8, 6525 GA Nijmegen, the Netherlands.
Dr. Kallenberg: Department of Clinical Immunology, University Hospital Groningen, Hanzeplein 1, 9713 GZ Groningen, the Netherlands.
Dr. Weening: Department of Paediatrics, Academic Medical Center, Meibergdreef 9, 1105 AZ Amsterdam, the Netherlands.
Dr. van Dissel: Department of Infectious Diseases, Leiden University Medical Center, Albinusdreef 2, 2333 ZA Leiden, the Netherlands.
Dr. Sanders: Department of Immunology, University Hospital for Children, Lundlaan 6, 3584 EA Utrecht, the Netherlands.
Ms. Nienhuis: Parexel Mirai, Herman Heijermansweg 20, Box 75079, 1070 AB Amsterdam, the Netherlands.
Dr. Schellekens: Department of Medicine, Academic Medical Center, Meibergdreef 9, 1105 AZ Amsterdam, the Netherlands.

Author Contributions: Conception and design: H.W. Eijkhout, J.W.M. van der Meer, C.G.M. Kallenberg, R.S. Weening, J.T. van Dissel, P.Th.A. Schellekens.
Analysis and interpretation of the data: H.W. Eijkhout, C.G.M. Kallenberg, J.T. van Dissel, E.A.M. Sanders, H. Nienhuis, P.Th.A. Schellekens.
Drafting of the article: H.W. Eijkhout, H. Nienhuis, P.Th.A. Schellekens.
Critical revision of the article for important intellectual content: J.W.M. van der Meer, C.G.M. Kallenberg, E.A.M. Sanders, P.F.W. Strengers, P.Th.A. Schellekens.
Final approval of the article: H.W. Eijkhout, J.W.M. van der Meer, C.G.M. Kallenberg, R.S. Weening, J.T. van Dissel, E.A.M. Sanders, P.F.W. Strengers, H. Nienhuis, P.Th.A. Schellekens.
Provision of study material or patients: J.W.M. van der Meer, C.G.M. Kallenberg, R.S. Weening, J.T. van Dissel, E.A.M. Sanders, P.Th.A. Schellekens.
Statistical expertise: H. Nienhuis.
Obtaining of funding: H.W. Eijkhout, J.W.M. van der Meer.
Collection and assembly of data: H.W. Eijkhout, P.F.W. Strengers.

References

1. Stiehm ER, Chin TW, Haas A, Peerless AG. Infectious complications of the primary immunodeficiencies. Clin Immunol Immunopathol. 1986;40:69-86. [PMID: 3521971]
2. Washington K, Stenzel TT, Buckley RH, Gottfried MR. Gastrointestinal pathology in patients with common variable immunodeficiency and X-linked agammaglobulinemia. Am J Surg Pathol. 1996;20:1240-52. [PMID: 8827031]
3. Rosen FS, Cooper MD, Wedgwood RJ. The primary immunodeficiencies. N Engl J Med. 1995;333:431-40. [PMID: 7616993]
4. Hurton DP, Kavanaugh AF, Rohane PW, Hurton MM. Immunoglobulin deficiency syndromes and therapy. J Allergy Clin Immunol. 1991;87:1-17. [PMID: 1704022]
5. Kainulainen L, Nikoskelainen J, Vuorinen T, Tevola K, Liippo K, Ruuskanen O. Viruses and bacteria in bronchial samples from patients with primary hypogammaglobulinemia. Am J Respir Crit Care Med. 1999;159:1199-204. [PMID: 10194166]
6. Ochs HD, Smith CI. X-linked agammaglobulinemia. A clinical and molecular analysis. Medicine (Baltimore). 1996;75:287-99. [PMID: 8982147]
7. Hermaszewski RA, Webster AD. Primary hypogammaglobulinaemia: a survey of clinical manifestations and complications. Q J Med. 1993;86:31-42. [PMID: 8438047]
8. Kerstens PJ, Endtz HP, Meis JF, Oyen WJ, Knopman RJ, van den Broek PJ, et al. Erysipelas-like skin lesions associated with *Campylobacter jejuni* septicemia in patients with hypogammaglobulinemia. Eur J Clin Microbiol Infect Dis. 1992;11:842-7. [PMID: 1468426]
9. Bruton OC. Agammaglobulinemia. Pediatrics. 1952;9:722-8.
10. Janeway CA, Rosen FS. The gamma globulins. IV. Therapeutic uses of gamma globulin. N Engl J Med. 1966;275:826-31. [PMID: 4161939]
11. Pirofsky B. Intravenous immune globulin therapy in hypogammaglobulinemia. A review. Am J Med. 1984;76:53-60. [PMID: 6424457]
12. Liese JG, Wintergerst U, Tympner KD, Belohradsky BH. High- vs low-dose immunoglobulin therapy in the long-term treatment of X-linked agammaglobulinemia. Am J Dis Child. 1992;146:335-9. [PMID: 1543181]
13. Cunningham-Rundles C, Siegal FP, Smithwick EM, Lion-Boulé A, Cunningham-Rundles S, O'Malley J, et al. Efficacy of intravenous immunoglobulin in primary humoral immunodeficiency disease. Ann Intern Med. 1984; 101:435-9. [PMID: 6206756]
14. Roifman CM, Levison H, Gelfand EW. High-dose versus low-dose intravenous immunoglobulin in hypogammaglobulinaemia and chronic lung disease. Lancet. 1987;1:1075-7. [PMID: 2883406]
15. Primary immunodeficiency diseases. Report of a WHO scientific group. In: Rosen FS, Seligmann M, eds. Immunodeficiencies. Chur, Switzerland: Harwood Academic Publishers; 1993:1-29.
16. Beam TR Jr, Gilbert DN, Kunin CM. General guidelines for the clinical evaluation of anti-infective drug products. Infectious Diseases Society of America and the Food and Drug Administration. Clin Infect Dis. 1992;15(Suppl 1):S5-32. [PMID: 1477251]
17. Bernatowska E, Madaliński K, Janowicz W, Wiertemowicz R, Gutkowski P, Wolf HM, et al. Results of a prospective controlled two-dose crossover study with intravenous immunoglobulin and comparison (retrospective) with plasma treatment. Clin Immunol Immunopathol. 1987;43:153-62. [PMID: 3568457]
18. Pruzanski W, Sussman G, Dorian W, Van T, Ibanez D, Reddelmeier D. Relationship of the dose of intravenous gammaglobulin to the prevention of infections in adults with common variable immunodeficiency. Inflammation. 1996;20:353-9. [PMID: 8872499]
19. Remvig L, Andersen V, Hansen NE, Karle H. Prophylactic effect of self-administered pump-driven subcutaneous IgG infusion in patients with antibody deficiency: a triple-blind cross-over study comparing P-IgG levels of 3 g L − 1 versus 6 g L − 1. J Intern Med. 1991;229:73-7. [PMID: 1995766]
20. Ochs HD, Fischer SH, Wedgwood RJ, Ward DW, Cowan MJ, Ammann AJ, et al. Comparison of high-dose and low-dose intravenous immunoglobulin therapy in patients with primary immunodeficiency diseases. Am J Med. 1984; 76:78-82. [PMID: 6424461]
21. Lindberg K, Samuelson A, Rynnel-Dagöö B, Smith E, Hammarström L. Nasal administration of IgA to individuals with hypogammaglobulinemia. Scand J Infect Dis. 1993;25:395-7. [PMID: 8362237]

(452) 5117

Effects of Fexofenadine, Diphenhydramine, and Alcohol on Driving Performance

A Randomized, Placebo-Controlled Trial in the Iowa Driving Simulator

John M. Weiler, MD; John R. Bloomfield, PhD; George G. Woodworth, PhD; Angela R. Grant, BS; Teresa A. Layton, BSN; Timothy L. Brown, MS; David R. McKenzie, MS; Thomas W. Baker, MS; and Ginger S. Watson, PhD

Background: Sedating antihistamines may impair driving performance as seriously as alcohol.

Objective: To compare the effects of fexofenadine, diphenhydramine, alcohol, and placebo on driving performance.

Design: Randomized, double-blind, double-dummy, four-treatment, four-period crossover trial.

Setting: The Iowa Driving Simulator.

Participants: 40 licensed drivers with seasonal allergic rhinitis who were 25 to 44 years of age.

Intervention: One dose of fexofenadine (60 mg), diphenhydramine (50 mg), alcohol (approximately 0.1% blood alcohol concentration), or placebo, given at weekly intervals before participants drove for 1 hour in the Iowa Driving Simulator.

Measurements: The primary end point was coherence, a continuous measure of participants' ability to match the varying speed of a vehicle that they were following. Secondary end points were drowsiness and other driving measures, including lane keeping and response to a vehicle that unexpectedly blocked the lane ahead.

Results: Participants had significantly better coherence after taking alcohol or fexofenadine than after taking diphenhydramine. Lane keeping (steering instability and crossing the center line) was impaired after alcohol and diphenhydramine use compared with fexofenadine use. Mean response time to the blocking vehicle was slowest after alcohol use (2.21 seconds) compared with fexofenadine use (1.95 seconds). Self-reported drowsiness did not predict lack of coherence and was weakly associated with minimum following distance, steering instability, and left-lane excursion.

Conclusions: Participants had similar performance when treated with fexofenadine or placebo. After alcohol use, participants performed the primary task well but not the secondary tasks; as a result, overall driving performance was poorer. After participants took diphenhydramine, driving performance was poorest, indicating that diphenhydramine had a greater impact on driving than alcohol did. Drowsiness ratings were not a good predictor of impairment, suggesting that drivers cannot use drowsiness to indicate when they should not drive.

Ann Intern Med. 2000;132:354-363.

For author affiliations, current addresses, and contributions, see end of text.

Allergic rhinitis afflicts more than 39 million persons in the United States (1). Only about 4.8 million persons (12%) take prescription drugs for this condition; most go without treatment or self-treat with over-the-counter medications, which generally contain a first-generation antihistamine. These medications may be effective but carry potential risks, including drowsiness and impairment in performing everyday tasks (2–6). These adverse events may be sufficient to dissuade some persons from treating their symptoms. Other patients take these sedating drugs, become impaired, and try nonetheless to perform complex tasks; as a result, they are more likely to be involved in collisions (2, 7, 8).

Our goal was to examine automobile driving performance, a complex multiaspect task requiring mental alertness; visual, auditory, and kinesthetic information processing; eye–hand coordination; and manual dexterity. By using the Iowa Driving Simulator, a unique state-of-the-art facility, we evaluated driving performance measures and self-ratings of drowsiness to determine the effects of alcohol and first- and second-generation antihistamines on driving performance. No previous study has compared the effects of these drugs in the highly controlled environment of a driving simulator.

Methods

Study Design

During ragweed season, we compared the effects of fexofenadine (60 mg), a second-generation antihistamine; diphenhydramine (50 mg) (Benadryl, Warner-Lambert Co., Morris Plains, New Jersey), a first-generation antihistamine; alcohol; and placebo on driving performance and self-reported drowsiness of persons who were allergic to ragweed. A randomized, double-blind, double-dummy, crossover design was used (9). The University of Iowa Institutional Review Board approved the study, and all participants signed a consent form before participation in the study.

See editorial comment on pp 405-407.

Inclusion and Exclusion Criteria

Key inclusion criteria were ability to remain for 5 hours after the drives, history of alcohol use and willingness to consume alcohol, age 25 to 45 years, seasonal allergic rhinitis caused by ragweed pollen, previous successful use of antihistamine to treat seasonal allergic rhinitis, status as a currently licensed experienced driver who drove an average of at least three times a week for at least 3 years, and 20/20 corrected vision. Key exclusion criteria were medical conditions that might interfere with ability to perform the study, pregnancy or lactation, unusual sleep patterns (including those of third-shift workers), excessive alcohol consumption, use of tobacco in the past year or excessive caffeine consumption, previous experience in the Iowa Driving Simulator, and a positive result on a drug screening test.

Procedures

At visit 1, participants were selected on the basis of inclusion and exclusion criteria. Qualified participants drove in the Iowa Driving Simulator for 8 minutes; those with a tendency to develop simulator sickness were excluded.

Visits 2, 3, 4, and 5 (treatment visits) occurred weekly on the same day at the same time. Participants avoided consuming food or beverages, except water, for 2 hours before these visits. Participants completed the baseline drowsiness visual analogue scale immediately before taking a capsule of fexofenadine, diphenhydramine, or placebo; the drive was scheduled to start 2.5 hours later to coincide with peak levels of antihistamine. Both researchers and participants were blinded to the treatment given. After treatment, participants were permitted to consume only fluids; caffeine, stimulants, and depressants were excluded. Vital signs were determined and participants completed the second drowsiness scale 1 hour after taking the capsule. The study beverage was dispensed 60 minutes before the scheduled drive and was consumed over 20 to 30 minutes with a light snack. The dose of alcohol (or placebo alcohol) was derived by using an algorithm that included the participant's sex and weight to reach an estimated blood alcohol concentration of 0.1% (21.7 mmol/L) (10). Male participants received the equivalent of 800 mg of absolute alcohol per kg of body weight, and female participants received 640 mg/kg. Ninety-five percent alcohol (or placebo alcohol) was added to a glass, which was filled with the participant's choice of noncaffeinated carbonated soda. Alcohol was swabbed on the rim of each glass to maintain blinding. Immediately before and after the drive, participants again completed drowsiness scales. After the drive, study staff determined vital signs. Participants were observed until they were sober. To maintain the double-blinding of the alcohol treatment, participants remained for 5 hours or until the blood alcohol level was less than 0.03% after alcohol and after one of the other treatments (selected randomly). An unblinded Clinical Research Center nurse with no other study role determined alcohol levels by using a breath analyzer (Alco-Sensor, Intoximeters, Inc., St. Louis, Missouri).

Treatment Preparation and Randomization

Capsules (fexofenadine, diphenhydramine, and placebo) were blinded and packaged by Hoechst Marion Roussel, Inc. (Kansas City, Missouri). The Division of Pharmaceutical Service, College of Pharmacy, University of Iowa, Iowa City, Iowa, prepared alcohol and placebo beverages.

Driving Simulation

The Iowa Driving Simulator allowed collection of data on driving performance measures in a manner not available with on-street driving (11, 12). Briefly, the simulator consists of a domed enclosure mounted on a hexapod motion platform. The inner walls of the dome act as a screen on which correlated images are projected.

The experimental drive was conducted in dry weather conditions, with good visibility, on a two-lane rural highway that was 72.4 km (45 miles) long. The lane widths were standard (3.66 m [12 ft]) and the road surface was standard blacktop. The posted speed limit was 88.6 km/h (55 miles/h) for most of the course. Vehicles in the oncoming lane simulated low-density traffic. Participants practiced driving in the simulator for 8 to 10 minutes before each experimental drive. The experimental drive consisted of two phases driven consecutively without interruption. In phase 1 (30% of the total driving distance), the driver followed a Volkswagen Golf. Phase 2 began when this lead vehicle turned off the main road and participants continued to drive "as you normally would" along the designated route. In the first three sessions, the experimental drive ended uneventfully. At the end of the fourth and final session, participants encountered a vehicle that unexpectedly pulled out from a driveway into the lane of the experimental vehicle. A truck with trailer simultaneously occupied the oncoming lane.

Outcome Measures

During the first phase, participants were instructed to maintain a constant distance behind a lead car, which had realistic random velocity fluctuations. The primary end point was coherence—the correlation between the velocity of the participant's

vehicle and the velocity of the lead vehicle. Participants with high coherence were able to maintain a relatively uniform distance from the lead vehicle, whereas those with low coherence had more variability in distance between their cars and the lead vehicle.

In both phases of the drive, we evaluated steering instability, the root-mean-square deviation (in meters), of the participant's car around the participant's preferred position in the lane. Participants with high instability wandered left and right within (and sometimes out of) the lane. We measured deviations from the preferred position rather than the geometric center of the lane to avoid penalizing otherwise steady drivers who simply preferred to be closer to the center line or to the right shoulder line. We also evaluated left-lane excursions—the total number of times the participant partially or totally crossed the center line during the second phase of the driving session.

We measured participants' responses to the blocking vehicle (the last event on the final drive). Videotapes and a computer-generated aerial view of the driving course and vehicles (generated by using Scenario Authoring Tool software [National Advanced Driving Simulator, Iowa City, Iowa]) were reviewed by three blinded investigators who evaluated two aspects of the participants' responses to the blocking vehicle. *Response time* was the time from the moment the blocking vehicle began to move until the instant the participant started to respond. The *blocking vehicle response category* was based on whether the participant's car came into contact with the incoming car or approaching truck (collision), stopped completely in the lane before passing the plane of the incoming car (clear avoidance), or either passed the plane of the incoming car or was more than a tire's width out of lane before stopping (potentially unsafe avoidance). Finally, we evaluated drowsiness by using a visual analogue scale (3, 4, 13–15) that asked participants to rate drowsiness from "I feel wide awake" to "I feel extremely sleepy."

Data Capture, Reduction, and Management

Simulator data were collected in real time at 30 Hz and were then reduced. During the data reduction stage, checks were performed to ensure that output was correct and meaningful. Data were visually inspected, sorted to identify extreme values, and plotted to ensure that all points were within naturally occurring boundaries. When extreme values were found, operator and experimenter source documents were consulted to determine an explanation. Videotape records were inspected to establish the origin of any anomalies in the data, and, if necessary, the raw data were fed into the Scenario Authoring Tool, which replayed the drive using animation.

Statistical Analysis

The experiment was run as a crossover design with four periods and four treatments so that each participant received all four treatments (alcohol, diphenhydramine, fexofenadine, and placebo) on four successive sessions in the driving simulator. With few exceptions, the sessions were 1 week apart at the same time of day. Treatments were presented in 24 different sequences (such as AFDP and FDAP). To ensure that each treatment occurred equally often in each period, the sequences were arranged in six Latin squares (for example, ADFP/DFPA/FPDA/PDAF). Four of the squares were replicated twice, for a total of 40 participants. The design was balanced so that each treatment preceded and followed the others equally often. Each treatment effect was estimated with equal precision in a model with treatment, period, and first-order carryover effects. In the design phase, we did an extensive Monte Carlo investigation of the robustness of this design to random loss of participants (rows of data) and found that the selected design (along with several similar designs) was robust, much more so than a design consisting of 10 replications of one Latin square.

Crossover designs have advantages and drawbacks. With four treatments, a crossover design requires one fourth the number of participants required by a completely randomized design in which each participant receives only one treatment. Furthermore, because each participant acts as his or her own control, it is, in theory, possible to compare treatments with much greater precision. One drawback of a crossover design is that early dropout of participants complicates the analysis and may have a comparatively greater impact on the precision of the results than the loss of a participant from a completely randomized trial. The most problematic aspect of crossover designs may be the effect of previous experiences on a participant's reaction to the current treatment. Such effects can be broadly classified as *period effects*, such as learning or habituation, which are unrelated to previous treatments, and *carryover effects*, which are related to previous treatments. Although it is unlikely that any residual study drug remained after an interval of 1 week, drug effects can carry over in other ways. For example, if a drug promoted simulator motion sickness, the participant may have driven more cautiously the week after receiving that drug. Without statistical adjustment, this drug-induced caution is attributed to whichever drug was administered the week after that drug. Statistical adjustment to remove period and carryover effects from the treatment means was accomplished by including variables for these effects in the analysis of variance model.

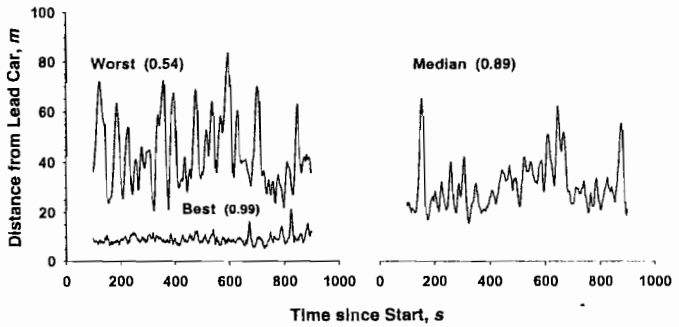


Figure 1. Maintenance of following distance for individual participants with near-best, near-worst, and near-median coherence scores. Initial and final transients are removed. The lower the score, the more erratic the following distance. The best driver (coherence, 0.99) varied about ±2.5 m, the worst driver (coherence, 0.54) varied about ±35 m, and the median driver (coherence, 0.89) varied about ±10 m.

Another complication of crossover designs is the statistical relation among repeated measures in the same participant. Participants' performance in the simulator is expected to be similar from week to week (that is, positively correlated), and variability may increase or decrease over time. Specifying the form of the "covariance structure" of the data deals with such issues (16). For simplicity, we chose the most general possible covariance structure.

Finally, the statistical method we used (the mixed general linear model) requires that the data be approximately normally distributed. Most of the outcomes we measured were significantly non-normally distributed, and it was necessary to re-express (transform) them to achieve normality. We used Box-Cox analysis (17) to select an appropriate power transformation of each variable. Specifically, left-lane excursion counts were re-expressed as log(count + 1), coherence (c) was re-expressed as $(1 - \sqrt{1 - c^2})^{1.25}$, steering instability ($s$) was re-expressed as s^{-1}, and minimum following distance (d) was re-expressed as $d^{-1/4}$. Statistics for re-expressed data are difficult to interpret; what does it mean, for example, that "log crossing count plus one" is 2.1 points higher for 1 drug than for placebo? To make our statistics interpretable, we converted all statistical results—means, differences, and CIs—back to the original, more interpretable measurement scales. Thus, crossing counts are reported as counts, steering instability and minimum following distance are expressed in meters, and coherence is expressed in its original form as the correlation between the velocity of the lead car and that of the participant's car. We used a Markov chain Monte Carlo procedure (18, 19) to compute these statistics and CIs.

All data were analyzed by using SAS software, versions 6.12 and 7.0 for Windows (SAS Institute, Inc., Cary, North Carolina). The contrast tests were two-sided, and an α level of 0.05 was required. Markov chain Monte Carlo computations were made by using WinBUGS, version 1.2 (19). For the primary and secondary outcome measures, we report treatment means and differences; CIs are given for differences between treatment means.

Response time to the blocking vehicle, which was measured in the fourth driving session only (so that only 25% of each treatment group was confronted with this situation), was analyzed by using a general linear model with treatments as the only effect. Response to the blocking vehicle (clear avoidance, potentially unsafe avoidance, or collision) was analyzed by using an exact permutation test (20).

Missing Data

One participant fell asleep after receiving alcohol and could not be roused for a driving session. Data for four other participants were missing for the second half of phase 2 in one session because these participants had simulator sickness. Mechanical problems resulted in the loss of phase 1 data for one participant in one session and data from the second half of phase 2 for another participant in one session. Thus, 2 of 160 sessions lacked phase 1 data and 6 of 160 sessions lacked data from the second half of phase 2.

The theory of missing data distinguishes between random and informative missing values (21). *Randomly missing data* are those that are missing for reasons unrelated to the participant's response to the treatment; they are therefore distributed like the observed data and can be predicted from the observed values of this participant and other participants. *Informative missing data* are missing for reasons related to the participant or treatment and are likely to have been somewhat anomalous if observed. Consequently, the fact that these data are

missing gives some information about the unobserved value. For example, the participant who could not be woken up would probably have driven badly if she had been awakened, and the participants who developed simulator sickness would probably have driven badly if they had completed the session.

The statistical software that we used imputes randomly missing data with the predicted value but adjusts degrees of freedom and SEs to reflect the fact that these values are not real data. Regarding informative missing data, Chow and Liu (21) remark that "There is no satisfactory, well-developed methodology to account for missing values or intermittent missing values." We believed that we should probe the sensitivity of the results to a range of plausible imputed values of the missing data. Therefore, we did analyses to assess whether the results were sensitive to possible values for the informative missing data. In one analysis, we treated them as missing at random; in another (the worst-case analysis), we imputed the nonrandomly missing values of impairment measures (high = bad) as the predicted value plus 2.5 SEs of the predicted value. We chose 2.5 SEs because it is pessimistic but does not distort the analysis by adding outliers. For performance measures (high = good), we subtracted 2.5 SEs from the predicted value. The results of the two analyses did not differ substantively. In this article, we report the results of the second analysis.

Role of the Study Sponsor

The industry sponsor had a consulting role in the design, conduct, and reporting of the study. Decisions in all aspects of the study, including the decision to publish the results, were made by the authors.

Results

Study Participants

Seventy-one participants were screened; 41 were randomly assigned and received double-blind treatment. One participant elected to discontinue participation during the first portion of her first drive and was not included in the efficacy analysis. Fifteen men (37.5%) and 25 women (62.5%) were included in the analysis. The mean age was 31 years (range, 25 to 44 years); 37 were white. Participants had a mean duration of ragweed allergy of 20 years.

Table 1. Primary End Point: Coherence*

Treatment	Participants, *n*	Mean Coherence Value * (95% CI)
Alcohol	40	0.920 ± 0.014 (0.891 to 0.945)
Diphenhydramine	40	0.877 ± 0.019 (0.837 to 0.911)
Fexofenadine	40	0.915 ± 0.014 (0.884 to 0.940)
Placebo	40	0.906 ± 0.015 (0.875 to 0.933)
Alcohol vs. diphenhydramine	40	0.043 ± 0.012 (0.021 to 0.068)
Alcohol vs. fexofenadine	40	0.005 ± 0.009 (−0.012 to 0.024)
Alcohol vs. placebo	40	0.014 ± 0.009 (−0.004 to 0.033)
Diphenhydramine vs. fexofenadine	40	−0.038 ± 0.013 (−0.063 to −0.013)
Diphenhydramine vs. placebo	40	−0.029 ± 0.012 (−0.054 to −0.005)
Fexofenadine vs placebo	40	0.009 ± 0.010 (−0.010 to 0.028)

* Data are expressed as the mean ± SD.

Phase 1

Coherence

As explained above, coherence was a participant's ability to maintain a constant distance from a lead car that varied its speed randomly. **Figure 1** provides a plot of distance fluctuations for three representative participants, one each with near-best, near-median, and near-worst coherence. Differences in coherence (**Table 1**) were observed among the four treatments. Pairwise comparisons revealed that after taking diphenhydramine, participants performed car-following with significantly less coherence than after taking alcohol, fexofenadine, or placebo (the CI excludes zero).

Minimum Following Distance

Significant differences in minimum following distance (**Table 2**) were observed among the four treatments. Pairwise comparisons indicated that when participants performed car-following after consuming alcohol, they had significantly smaller minimum following distances (15.1 m) than they did after taking fexofenadine (17.1 m) or placebo (17.4 m).

Steering Instability

Significant differences in steering instability (**Table 2**) were observed among the four treatments. Pairwise comparisons showed that after participants took fexofenadine, they had significantly less steering instability than after taking diphenhydramine or alcohol (but not placebo). After participants took placebo, they had significantly less steering instability than after consuming alcohol or diphenhydramine.

Phase 2

After completing phase 1, participants drove the remaining 30 miles of the course "as you normally would drive." Road signs and markings were the only guidance that they received in this phase.

Steering Instability

Significant differences in steering instability (**Table 2**) were again observed among the four treatments. Pairwise comparisons demonstrated that after participants took fexofenadine, they had significantly less steering instability than after taking diphenhydramine or alcohol (but not placebo). After participants took placebo, they had significantly less steering instability than after consuming alcohol or diphenhydramine. After participants consumed alcohol, they had the same or less steering instability than after taking diphenhydramine.

Table 2. Secondary End Points*

Treatment	Phase 1		Phase 2	
	Minimum Following Distance (95% CI)	Steering Instability (95% CI)	Steering Instability (95% CI)	Left-Lane Excursions (95% CI)
	m			*n*
Alcohol	15.07 ± 1.11 (13.04 to 17.43)	0.376 ± 0.010 (0.359 to 0.397)	0.512 ± 0.0088 (0.498 to 0.531)	2.12 ± 0.56 (1.16 to 3.34)
Diphenhydramine	16.25 ± 1.22 (14.05 to 18.79)	0.380 ± 0.010 (0.363 to 0.402)	0.527 ± 0.0095 (0.508 to 0.546)	3.15 ± 0.75 (1.85 to 4.82)
Fexofenadine	17.05 ± 1.29 (14.72 to 19.77)	0.354 ± 0.009 (0.338 to 0.372)	0.492 ± 0.0080 (0.477 to 0.509)	1.17 ± 0.38 (0.52 to 2.01)
Placebo	17.43 ± 1.32 (15.06 to 20.20)	0.359 ± 0.009 (0.344 to 0.378)	0.495 ± 0.0083 (0.480 to 0.513)	1.32 ± 0.40 (0.63 to 2.21)
Alcohol vs. diphenhydramine	−1.18 ± 0.78 (−2.80 to 0.32)	−0.004 ± 0.007 (−0.017 to 0.009)	−0.014 ± 0.0073 (−0.029 to 0.000)	−1.03 ± 0.60 (−2.30 to 0.06)
Alcohol vs. fexofenadine	−1.98 ± 0.85 (−3.67 to −0.36)	0.022 ± 0.006 (0.011 to 0.034)	0.020 ± 0.0067 (0.007 to 0.033)	0.94 ± 0.45 (0.14 to 1.90)
Alcohol vs. placebo	−2.36 ± 0.85 (−4.10 to −0.76)	0.017 ± 0.006 (0.006 to 0.029)	0.017 ± 0.0068 (0.003 to 0.031)	0.79 ± 0.43 (0.01 to 1.70)
Diphenhydramine vs. fexofenadine	−0.80 ± 0.86 (−2.52 to 0.88)	0.026 ± 0.007 (0.014 to 0.040)	0.034 ± 0.0074 (0.020 to 0.049)	1.98 ± 0.61 (0.87 to 3.31)
Diphenhydramine vs. placebo	−1.18 ± 0.84 (−2.84 to 0.44)	0.021 ± 0.006 (0.009 to 0.034)	0.031 ± 0.0074 (0.017 to 0.046)	1.83 ± 0.61 (0.74 to 3.14)
Fexofenadine vs. placebo	−0.38 ± 0.87 (−2.09 to 1.29)	−0.005 ± 0.005 (−0.016 to 0.005)	−0.003 ± 0.0066 (−0.017 to 0.010)	−0.15 ± 0.34 (−0.85 to 0.52)

* Data are expressed as the mean ± SD. All data are based on 40 participants.

Lane Excursions

We determined the effect of treatment on the probability that the participant's vehicle moved to the right and partially or totally crossed the right-edge lane marker or moved to the left and partially or totally crossed the center line (**Table 2**). No significant differences for lane excursions to the right were noted among the four treatments. For excursions to the left, however, significant differences were noted the four treatments. Pairwise comparisons demonstrated that after participants took diphenhydramine, they crossed the center line significantly more often than after taking fexofenadine or placebo. After participants took alcohol, they crossed the center line significantly more often than after taking fexofenadine and placebo. Fexofenadine and placebo did not differ significantly; the 95% CIs indicate that the difference is small (**Table 2**).

Response to Blocking Vehicle

No significant main effect of treatment on the response time to the blocking vehicle was observed, although pairwise comparisons showed that after consuming alcohol, participants responded significantly more slowly (2.21 seconds) to the event than after they took fexofenadine (1.95 seconds) (difference, 0.26 seconds [CI, 0.02 to 0.66 seconds]).

Responses to the blocking vehicle were categorized as clear avoidance, potentially unsafe avoidance, or collision (**Table 3**). The overall differences were not significant ($P > 0.2$, Fisher exact test). Pairwise comparisons, expressed as odds ratios, were also insignificant. However, because this event occurred only during the fourth drive, there were only 9 to 11 participants in each group (rather than 40, as was the case for all of the other measures). As a result, this analysis had far less power than the analyses of the other secondary measures.

Crashes were evaluated for speed of the driver's vehicle at the instant of the crash. For the 5 collisions, the speed at impact was 46 and 14 miles per hour after alcohol, 37 and 8 miles per hour after diphenhydramine, and 6 miles per hour after fexofenadine.

Subjective Drowsiness Ratings

Drowsiness ratings were expressed as differences between the drowsiness scales completed after treatments and the baseline scale (**Figure 2**). Scores on the second visual analogue scale, given 1 hour after capsule administration, had small average differences from baseline (<10 points), and no significant

Table 3. Clear Avoidance, Potentially Unsafe Avoidance, and Collision in the Final Driving Session

Treatment	Clear Avoidance	Potentially Unsafe Avoidance	Collision	Odds Ratio for Collision vs. Clear Avoidance or Potentially Unsafe Avoidance (95% CI)	Odds Ratio for Potentially Unsafe Avoidance or Collision vs. Clear Avoidance (95% CI)
	← n (%) →				
Fexofenadine (n = 11)	8 (72.7)	2 (18.2)	1 (9.1)		
Diphenhydramine (n = 10)	5 (50.0)	3 (30.0)	2 (20.0)		
Alcohol (n = 9)	6 (66.7)	1 (11.1)	2 (22.2)		
Placebo (n = 9)	8 (88.9)	1 (11.1)	0 (0.0)		
Diphenhydramine vs. alcohol				0.88 (0.051–15.3)	2.00 (0.226–19.4)
Fexofenadine vs. alcohol				0.35 (0.005–8.41)	0.75 (0.073–7.87)
Alcohol vs. placebo				– (2.14)*	4.00 (0.229–238.1)
Diphenhydramine vs. fexofenadine				2.22 (0.106–161.3)	2.67 (0.319–24.5)
Diphenhydramine vs. placebo				– (0.191)*	8.00 (0.558–416.7)
Fexofenadine vs. placebo				– (0.303)*	3.00 (0.181–175.4)

* The denominator of the odds ratio was zero. Only the lower limit of the CI is given; the upper limit was unbounded.

differences were seen among the treatment groups (the confidence limits for differences less than ±10 points). At the time of the third visual analogue scale, just before the drive, participants were most drowsy after taking diphenhydramine and least drowsy after taking fexofenadine or placebo. The differences between diphenhydramine and fexofenadine or placebo were significant (confidence limits ranged from 5 to 27 points on the 100-point visual analogue scale). The difference between fexofenadine and placebo was less than 1 point, with confidence limits of ±11 points. After the drive, participants were most drowsy with diphenhydramine and least drowsy with placebo. The difference between fexofenadine and placebo was insignificant (confidence limits were −7 to 19 points). With diphenhydramine, participants reported significantly higher levels of drowsiness than with fexofenadine (confidence limits of 9 to 35 points) and placebo (confidence limits of 15 to 41 points).

We examined whether self-reported drowsiness immediately before driving predicted impaired driving performance. Drowsiness was expressed as the difference between the third and first self-ratings. The correlation between drowsiness and the primary end point, coherence, was not statistically significant (**Table 4**). Statistically significant but small correlations were found between subjective drowsiness and minimum following distance, steering instability, and left-lane excursions; no correlation was greater than 0.21.

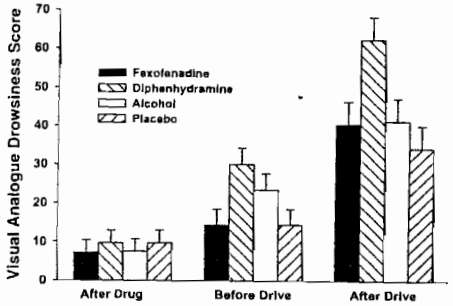


Figure 2. Change from baseline in visual analogue drowsiness scores. Participants rated drowsiness on a scale from "wide awake" to "extremely drowsy," which corresponded to a score of 1 to 100 on a 159-mm scale.

Although a significant correlation indicates some relation between two variables, the size of the correlation coefficient is not a good indicator of the strength of that relation. To give an idea of the practical meaning of the correlations we observed, **Table 4** shows mean driving performance values for participants who had increases in drowsiness scores in the upper quartile and lower three quartiles. Clearly, drowsiness was a weak predictor of poor driving. Indeed, only one of the five collisions occurred among the participants who were in the drowsiest quartile (as measured before or after the drive). Thus, "grounding" the drowsiest 25% of drivers would have prevented only 20% of the collisions. In contrast, three of five collisions occurred in participants who had the lowest quartile of following distances (following distance < 12.2 m), and four of five collisions occurred in participants who had the highest quartile of left-lane crossings (seven or more crossings).

Adverse Events

No unusual or serious adverse events were observed in this study. Adverse events occurred with similar frequency after all four treatments, with no significant differences between any two treatments in any adverse event category.

Discussion

First-generation antihistamines, such as diphenhydramine, cause sedation (2–6), which Gengo and Gabos (22) have distinguished as impairment and drowsiness. Cognitive impairment refers to some interference with the patient's ability to perform tasks and is measured by objective tests; drowsiness, which may or may not limit performance, is measured subjectively. The least desirable condition would be impairment without drowsiness because a patient might have no subjective clues suggesting impairment.

The second-generation antihistamines have difficulty crossing the blood–brain barrier and are believed to cause little or no central nervous system depression. In previous studies, fexofenadine and its parent compound, terfenadine, did not impair the performance of automobile drivers or airplane pilots (6, 23, 24).

In this study, participants in a driving simulator were first instructed to match the speed of the car they were following, then to drive "as you normally would." Coherence was chosen as the primary end point because the added complexity of trying to match the variable speed of the lead car might lead to greater sensitivity if impairment did occur. Coherence was significantly better after participants took alcohol or fexofenadine than after they took diphenhydramine. The minimum following distance was slightly shorter than the recommended distance after all four treatments (15.1 m [49.4 ft] to 17.4 m [57.2 ft]). The mean minimum following distance was about one-half car length longer (and safer) after participants had taken fexofenadine or placebo than after they had consumed alcohol. The shorter following distance might also have contributed to increased coherence. However, during the car-following phase, steering instability scores were highest after diphenhydramine or alcohol use, indicating poorer steering control.

Thus, although participants under the influence of alcohol did surprisingly well at matching the velocity of the lead car, they did so at the expense of driving closer to that vehicle and having less control over steering. These findings agree with the results obtained in other studies in which alcohol was administered to participants who were engaged in complex tasks that required divided attention. Horne and Baumber (25) reported that drivers who had consumed alcohol were able to maintain lateral position in wind gusts but did not perform well at following another vehicle. Landauer and Howat (10) used a nondriving task involving reaction time and tracking accuracy and found that after participants consumed alcohol, reaction time improved slightly but the number of tracking errors increased. Moskowitz (26) and Kerr and Hindmarch (27) reviewed studies of alcohol and divided attention and suggested that if one part of a divided attention task is perceived to be primary and the other part secondary, only the secondary task becomes impaired.

When we examined how participants performed when driving "normally," we found more steering instability after participants took diphenhydramine or alcohol than after they took fexofenadine or placebo. Participants with poorer steering were more likely to drive with part of the vehicle out of the lane. Lane excursions over the center line (causing potential exposure to oncoming traffic) may seriously affect safety. The numbers of lane excursions over the center line more than doubled after the participants had taken diphenhydramine compared with fexofenadine or placebo.

We also examined drowsiness and found that participants were significantly drowsier after taking diphenhydramine than after taking any of the other treatments. However, we found that subjective drowsiness either did not predict driving performance measures (coherence) or was a relatively weak predictor (for minimum following distance, steering instability, and left-lane excursion). This suggests that drivers who use alcohol or diphenhydramine are probably mistaken if they believe that lack of drowsiness means that they will be able to drive without impairment.

The potential crash scenario on the last drive provided some additional evidence of impairment. Participants had to react to a vehicle that unexpectedly pulled out of a driveway and blocked their lane.

Table 4. Performance Measures according to Degree of Sleepiness before Drive 4

Drowsiness Category*	Mean Coherence Value (95% CI)†	Mean Minimum Following Distance (95% CI)‡	Mean Steering Instability (95% CI)§	Left-Lane Excursions (95% CI)‖	Mean Collision Rate (95% CI)
		m		n	
Lower three fourths	0.893 (0.869 to 0.923)	16.1 (13.8 to 18.7)	0.530 (0.503 to 0.557)	2.06 (1.00 to 3.53)	0.138 (0.012 to 0.263)
Upper one fourth	0.917 (0.864 to 0.919)	12.9 (10.0 to 16.6)	0.539 (0.538 to 0.591)	4.88 (1.69 to 10.41)	0.100 (0.000 to 0.263)
Difference	0.024 (−0.037 to 0.077)	3.2 (−1.27 to 7.11)	0.009 (−0.045 to 0.067)	2.82 (−0.72 to 8.40)¶	
Odds ratio					0.694 (0.013 to 8.44)

* Increase in drowsiness between baseline (first drowsiness evaluation) and the third evaluation (immediately before a drive).
† r = 0.06; P > 0.2.
‡ r = 0.20; P = 0.01.
§ r = 0.20; P = 0.01.
‖ r = 0.21; P = 0.01.
¶ Statistically significant (P = 0.006).

Participants responded significantly more slowly to the event after consuming alcohol than after taking fexofenadine. At the posted speed, this slower reaction time resulted in a stopping distance that was approximately 8 m (26 ft) longer.

The observations reported here, combined with past reports, indicate that diphenhydramine clearly impairs driving performance, whereas the second-generation antihistamine fexofenadine was indistinguishable from placebo. Vermeeren and O'Hanlon (24) studied one driving variable, lateral position, and also reported that fexofenadine did not affect standard deviation of lateral position in an instrumented car used in an on-the-road study, nor did it affect various nondriving psychomotor tasks. In contrast, the first-generation antihistamine clemastine caused significant impairment.

In the United States, diphenhydramine is the top-selling over-the-counter medication sold for treatment of allergic rhinitis (28). It is estimated that 47% of persons with allergies treat themselves with over-the-counter products, most of which contain a sedating antihistamine (29). Consequently, millions of patients use first-generation antihistamines.

Several health programs have been developed that limit patient access to nonsedating antihistamines and emphasize the use of first-generation antihistamines (30, 31). The cost savings of these programs should be weighed against the potential increased risk to the driving public and against the laws of 27 states that prohibit driving under the influence of any drug or any substance (32, 33).

We conclude that participants performed similarly when treated with fexofenadine or placebo. Participants who consumed alcohol did well in performing the primary driving task but not the secondary tasks, resulting in poorer overall driving performance. This study demonstrates that the first-generation antihistamine diphenhydramine may have an even greater impact than does alcohol on the complex task of operating an automobile.

From the University of Iowa, Iowa City, Iowa; and Hoechst Marion Roussel, Inc., Kansas City, Missouri.

Disclosure: Dr. Weiler serves as a consultant and Dr. Woodworth has provided consulting services to Hoechst Marion Roussel, Inc.

Acknowledgments: The authors thank the following for their assistance in the conduct or analysis of the study: Susan Quinn, Sue Ellen Salisbury, Elizabeth Lawler, Cindy Mitchell, Emily Meis, Kathy Phipps, Suzanne Sack, Jagadish Boggavarapu, Dixie Ecklund and the Clinical Research Center staff, Twila Finkelstein, Julie Qidwai, Christopher Miller, Joss Nichols, Brent Caven, Mark Young, Dawn Kenyon, Kristen Rassbach, Brad Graham, Chris McMillan, Nick Taiber, Sneha Viratia, Srinivas Maddhi, Rohit Goal, Lucas Davisson, Brian Berentsen, Shaheen Bahauddin, Peter Grant, Katie Enstrom, Omar Ahmad, Imran Pirwani, Ludovic Moineau, Yiannis Papelis, Matthew Schikore, Tim Van Fossen, Dave Bronder, Shawn Allen, Rachel Nador, Steven Zellers, Ianos Schmidt, Paul Debbins, and Dave Muller.

Grant Support: By a grant from Hoechst Marion Roussel, Inc., and by grant M01-RR-59 from the National Center for Research Resources, General Clinical Research Centers Program, National Institutes of Health.

Requests for Single Reprints: John M. Weiler, MD, Department of Internal Medicine, University of Iowa, 200 Hawkins Drive, T307GH, Iowa City, IA 52242-1081; e-mail, john-weiler@uiowa.edu.

Requests To Purchase Bulk Reprints (minimum, 100 copies): Barbara Hudson, Reprints Coordinator; phone, 215-351-2657; e-mail, bhudson@mail.acponline.org.

Current Author Addresses: Dr. Weiler: University of Iowa Hospitals and Clinics, T307, 200 Hawkins Drive, Iowa City, IA 52242.
Dr. Woodworth: Department of Statistics, University of Iowa, 241 Schaeffer Hall, Iowa City, IA 52242.
Ms. Grant: 1660 Fulmer Street, Ann Arbor, MI 48103.
Ms. Layton: University of Iowa Hospitals and Clinics, 200 Hawkins Drive, Iowa City, IA 52242.
Mr. Brown: 4133 SC, University of Iowa, Iowa City, IA 52242.
Mr. McKenzie: Quintiles, Inc., Box 9708, Kansas City, MO 64134-0708.
Mr. Baker: Aventis Pharmaceuticals, Inc., Box 9627, Kansas City, MO 64134-0627.
Dr. Watson: National Advanced Driving Simulator, 2401 Oakdale Boulevard, University of Iowa, Iowa City, IA 52242.

Author Contributions: Conception and design: J.M. Weiler, J.R. Bloomfield, A.R. Grant, G.G. Woodworth, D.R. McKenzie, T.W. Baker.

Analysis and interpretation of the data: J.M. Weiler, J.R. Bloomfield, A.R. Grant, T.L. Brown, G.G. Woodworth, G.S. Watson.

Drafting of the article: J.M. Weiler, T.A. Layton, D.R. McKenzie, T.W. Baker, G.S. Watson.

Critical revision of the article for important intellectual content: J.M. Weiler, J.R. Bloomfield, A.R. Grant, G.G. Woodworth, D.R. McKenzie, T.W. Baker, G.S. Watson.

Final approval of the article: J.M. Weiler, G.G. Woodworth, T.A. Layton, T.L. Brown, T.W. Baker, G.S. Watson.

Provision of study materials or patients: J.M. Weiler, T.W. Baker.

Statistical expertise: G.G. Woodworth, D.R. McKenzie.

Obtaining of funding: J.M. Weiler, J.R. Bloomfield, D.R. McKenzie, T.W. Baker.

Administrative, technical, or logistic support: A.R. Grant, T.A. Layton.

Collection and assembly of data: J.M. Weiler, J.R. Bloomfield, A.R. Grant, T.A. Layton, T.L. Brown.

References

1. **Malone DC, Lawson KA, Smith DH, Arrighi HM, Battista C.** A cost of illness study of allergic rhinitis in the United States. J Allergy Clin Immunol. 1997;99:22-7.
2. **Starmer G.** Antihistamines and highway safety. Accid Anal Prev. 1985;17: 311-7.
3. **Gengo F, Gabos C, Miller JK.** The pharmacodynamics of diphenhydramine-induced drowsiness and changes in mental performance. Clin Pharmacol Ther. 1989;45:15-21.
4. **Gengo FM, Gabos C, Mechtler L.** Quantitative effects of cetirizine and diphenhydramine on mental performance measured using an automobile driving simulator. Ann Allergy. 1990;64:520-6.
5. **Brookhuis KA, De Vries G, De Waard D.** Acute and subchronic effects of the H1-antihistamine receptor antagonist ebastine in 10, 20 and 30 mg dose, and triprolidine 10 mg on car driving performance. Br J Clin Pharmacol. 1993;36:67-70.
6. **D'Hanlon JF, Ramaekers JG.** Antihistamine effects on actual driving performance in a standard test: a summary of Dutch experience, 1989-94. Allergy. 1995;50:234-42.
7. **Cimbura G, Lucas DM, Bennett RC, Warren RA, Simpson HM.** Incidence and toxicological aspects of drugs detected in 484 fatally injured drivers and pedestrians in Ontario. J Forensic Sci. 1982;27:855-67.
8. **Gengo FM, Manning C.** A review of the effects of antihistamines on mental processes related to automobile driving. J Allergy Clin Immunol. 1990;86: 1034-9.
9. **Chow SC, Liu JP.** Section 5.3. Crossover designs. In: Design and Analysis of Clinical Trials: Concept and Methodologies. New York: Wiley; 1998:183-90.
10. **Landauer AA, Howat P.** Low and moderate alcohol doses, psychomotor performance and perceived drowsiness. Ergonomics. 1983;26:647-57.
11. **Kuhl JG, Evans D, Papelis YE, Romano RA, Watson GS.** The Iowa Driving Simulator: an immersive research environment. IEEE Computer. 1995;28:35-41.
12. **Freeman JS, Watson GS, Papelis YE, Lin TC, Tayyab A, Romano RA, et al.** The Iowa Driving Simulator: An Implementation and Application Overview. Warrendale, PA: Soc of Automotive Engineers International; 1995. SAE Technical Paper Series 950174.
13. **Witek TJ, Canestrari DA, Miller RD, Yang JY, Riker DK.** Characterization of daytime sleepiness and psychomotor performance following H1 receptor antagonists. Ann Allergy Asthma Immunol. 1995;74:419-26.
14. **McCormack HM, Horne DJ, Sheather S.** Clinical applications of visual analogue scales: a critical review. Psychol Med. 1988;18:1007-19.
15. **Chemik DA, Gillilings D, Laine H, Hendler J, Silver JM, Davidson AB, et al.** Validity and reliability of the Observer's Assessment of Alertness/Sedation Scale: study with intravenous midazolam. J Clin Psychopharmacol. 1990;10: 244-51.
16. The MIXED procedure. In: SAS/STAT Software: Changes and Enhancements through Release 6.12. Cary, NC: SAS Institute; 1997:571-702.
17. **Welsberg S.** Applied Linear Regression. 2d ed. New York: Wiley; 1985.
18. **Gelman A, Rubin DB.** Markov chain Monte Carlo methods in biostatistics. Stat Methods Med Res. 1996;5:339-55.
19. **Spiegelhalter DJ, Thomas A, Best NG.** WinBUGS Version 1.2. User Manual. Cambridge, UK: MRC Biostatistics Unit; 1999.
20. The FREQ Procedure, Exact Statistics. In: SAS/STAT Software: Changes and Enhancements through Release 6.12. Cary, NC: SAS Institute; 1997:221-46.
21. **Chow SC, Liu JP, eds.** Issues in efficacy evaluation. In: Design and Analysis of Clinical Trials: Concept and Methodologies. New York: Wiley, 1998:531-2.
22. **Gengo FM, Gabos C.** Antihistamines, drowsiness, and psychomotor impairment: central nervous system effect of cetirizine. Ann Allergy. 1987;59:53-7.
23. **Philpot EE, Brooker AE, Blegalski CS.** Effects of sedating and nonsedating antihistamines on flying performance. Mil Med. 1993;158:654-60.
24. **Vermeeren A, O'Hanlon JF.** Fexofenadine's effects, alone and with alcohol, on actual driving and psychomotor performance. J Allergy Clin Immunol. 1998;101:306-11.
25. **Horne JA, Baumber CJ.** Time-of-day effects of alcohol intake on simulated driving performance in women. Ergonomics. 1991;34:1377-83.
26. **Moskowitz H.** Validity of driving simulator studies for predicting drug effects in real driving situations. In: Israelstam S, Lambert S, eds. Alcohol, Drugs and Traffic Safety: Proceedings of the Sixth International Conference on Alcohol, Drugs, and Traffic Safety. Toronto: Addiction Research Foundation of Ontario, 1974:295-303.
27. **Kerr JS, Hindmarch I.** The effects of alcohol alone or in combination with other drugs on information processing, task performance and subjective responses. Human Psychopharmacology. 1998;13:1-9.
28. Red Book. Montvale, NJ: Medical Economics; 1998:609.
29. **Smith EA.** Blame El Niño. Drug Topics. 1998;142:60.
30. **Meszaros L.** Allergy: old nemesis, new approaches. Managed Healthcare. 1996;6:532-4.
31. Use of OTC antihistamines, NSAIDs saves 12% over brand names. Managed Pharmaceutical Report. 1996;3:7-8.
32. Digest of State Alcohol-Highway Safety Related Legislation. 14th ed. Washington, DC: Office of Alcohol and State Programs, National Highway Traffic Safety Administration, U.S. Department of Transportation; 1996.
33. **Slavin RG.** Occupational and allergic rhinitis: impact on worker productivity and safety. Allergy Asthma Proc. 1998;19:277-84.

Vapreotide for Variceal Bleeding

"The combination of vapreotide and endoscopic treatment is more effective than endoscopic treatment alone."

Patients with bleeding from esophageal varices may be treated with somatostatin or endoscopy. In this study patients with cirrhosis and variceal bleeding were treated with the somatostatin analogue vapreotide or placebo for five days; all underwent endoscopic treatment within 12 hours. At endoscopy, 31 percent of the vapreotide group were actively bleeding, as compared with 46 percent of the placebo group. During the five-day treatment period the rates of survival with control of bleeding in the two groups were 66 percent and 50 percent, respectively.

Achieving hemostasis and preventing recurrent bleeding in patients with cirrhosis who have variceal bleeding is a difficult clinical problem. The combined approach used in this study seems to offer benefits, with respect to immediate cessation of bleeding and prevention of recurrence. However, the short-term mortality rates were similar in the two groups.

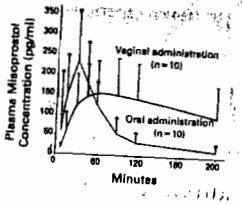
see page 23

Advances in Immunology: **Allergy and Allergic Diseases**

This article, which is part of the *Journal*'s Immunology series, is the first of a comprehensive two-part review of the pathogenesis and treatment of allergic diseases.

see page 30

Drug Therapy: **Misoprostol and Pregnancy**



Misoprostol has important uses in pregnant women. These uses include induction of abortion in conjunction with mifepristone or methotrexate, treatment of missed abortion, induction of cervical ripening and of labor, and prevention and treatment of postpartum hemorrhage. This article summarizes the results of studies of the efficacy and side effects of misoprostol in obstetrical practice.

see page 38 (editorial, page 59)

Clinical Problem-Solving: **High Time for Action**

A 58-year-old man receiving hemodialysis after failed renal transplantation presents with unstable angina. He also reports malaise, nausea and vomiting since repair of an intraabdominal aortic aneurysm and removal of the renal graft less than four weeks earlier. The prothrombin time and activated partial-thromboplastin time are checked before planned cardiac catheterization. Although previously normal, these values are prolonged at 88 seconds and 77 seconds, respectively.

see page 48

The New England Journal of Medicine

VOLUME 344 JANUARY 4, 2001 NUMBER 1

(292) [334]

EFFECTS ON BLOOD PRESSURE OF REDUCED DIETARY SODIUM AND THE DIETARY APPROACHES TO STOP HYPERTENSION (DASH) DIET

FRANK M. SACKS, M.D., LAURA P. SVETKEY, M.D., WILLIAM M. VOLLMER, PH.D., LAWRENCE J. APPEL, M.D., GEORGE A. BRAY, M.D., DAVID HARSHA, PH.D., EVA OBARZANEK, PH.D., PAUL R. CONLIN, M.D., EDGAR R. MILLER III, M.D., PH.D., DENISE G. SIMONS-MORTON, M.D., PH.D., NJERI KARANJA, PH.D., AND PAO-HWA LIN, PH.D., FOR THE DASH–SODIUM COLLABORATIVE RESEARCH GROUP

ABSTRACT

Background The effect of dietary composition on blood pressure is a subject of public health importance. We studied the effect of different levels of dietary sodium, in conjunction with the Dietary Approaches to Stop Hypertension (DASH) diet, which is rich in vegetables, fruits, and low-fat dairy products, in persons with and in those without hypertension.

Methods A total of 412 participants were randomly assigned to eat either a control diet typical of intake in the United States or the DASH diet. Within the assigned diet, participants ate foods with high, intermediate, and low levels of sodium for 30 consecutive days each, in random order.

Results Reducing the sodium intake from the high to the intermediate level reduced the systolic blood pressure by 2.1 mm Hg ($P<0.001$) during the control diet and by 1.3 mm Hg ($P=0.03$) during the DASH diet. Reducing the sodium intake from the intermediate to the low level caused additional reductions of 4.6 mm Hg during the control diet ($P<0.001$) and 1.7 mm Hg during the DASH diet ($P<0.01$). The effects of sodium were observed in participants with and in those without hypertension, blacks and those of other races, and women and men. The DASH diet wes associated with a significantly lower systolic blood pressure at each sodium level; and the difference was greater with high sodium levels than with low ones. As compared with the control diet with a high sodium level, the DASH diet with a low sodium level led to a mean systolic blood pressure that was 7.1 mm Hg lower in participants without hypertension, and 11.5 mm Hg lower in participants with hypertension.

Conclusions The reduction of sodium intake to levels below the current recommendation of 100 mmol per day and the DASH diet both lower blood pressure substantially, with greater effects in combination than singly. Long-term health benefits will depend on the ability of people to make long-lasting dietary changes and the increased availability of lower-sodium foods. (N Engl J Med 2001;344:3-10.)

HYPERTENSION affects almost 50 million people in the United States and places them at higher risk for cardiovascular diseases.[1,2] Furthermore, this risk increases with progressive elevations in blood pressure, beginning at even normal levels of blood pressure.[3] The Dietary Approaches to Stop Hypertension (DASH) trial demonstrated that a diet that emphasizes fruits, vegetables, and low-fat dairy products, that includes whole grains, poultry, fish, and nuts, that contains only small amounts of red meat, sweets, and sugar-containing beverages, and that contains decreased amounts of total and saturated fat and cholesterol lowers blood pressure substantially both in people with hypertension and those without hypertension, as compared with a typical diet in the United States.[4] The DASH diet is now recommended in national guidelines.[1,5] Clinical trials have shown that reducing the sodium chloride content of typical diets in the United States or northern Europe lowers blood pressure,[6-8] and guidelines recommend reducing the daily dietary sodium

From the Endocrine–Hypertension Division and the Channing Laboratory, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston (F.M.S., P.R.C.); the Duke Hypertension Center and the Sarah W. Stedman Center for Nutritional Studies, Duke University School of Medicine, Durham, N.C. (L.P.S., P.-H.L.); the Kaiser Permanente Center for Health Research, Portland, Oreg. (W.M.V., N.K.); the Welch Center for Prevention, Epidemiology, and Clinical Research, Johns Hopkins University, Baltimore (L.J.A., E.R.M.); the Pennington Biomedical Research Center, Baton Rouge, La. (G.A.B., D.H.); and the Division of Epidemiology and Clinical Applications, National Heart, Lung, and Blood Institute, Bethesda, Md. (E.O., D.G.S.-M.). Address reprint requests to Dr. Sacks at the Nutrition Department, Harvard School of Public Health, 665 Huntington Ave., Boston, MA 02115, or at fsacks@hsph.harvard.edu.

Other authors were Mikel Aickin, Ph.D., Kaiser Permanente Center for Health Research, Portland, Oreg.; Marlene M. Most-Windhauser, Ph.D., Pennington Biomedical Research Center, Baton Rouge, La.; Thomas J. Moore, M.D., Merck, West Point, Pa.; and Michael A. Proschan, Ph.D., and Jeffrey A. Cutler, M.D., National Heart, Lung, and Blood Institute, Bethesda, Md.

intake to 100 mmol (equivalent to 2.3 g of sodium or 5.8 g of sodium chloride) or less.[1]

We undertook this trial to address several questions relevant to the prevention and treatment of hypertension. Does reducing the level of sodium from the average intake in the United States (approximately 150 mmol per day, which is equivalent to 3.5 g of sodium, or 8.7 g of sodium chloride) to below the currently recommended upper limit of 100 mmol per day lower blood pressure more than reducing the sodium level only to the recommended limit? We hypothesized that it would, on the basis of both the blood-pressure levels in populations with an average consumption of less than 60 mmol of sodium per day[9] and data from incompletely controlled[10-14] or small-scale[15] clinical trials. Does the DASH diet lower the blood pressure beyond the level achievable by simply reducing sodium intake? What is the combined effect of the DASH diet and reduced sodium intake? The extent to which the reduction of the sodium level, in the context of a typical United States diet and in combination with the DASH diet, lowers blood pressure in people without hypertension is a much-debated[6-8] issue critical to the prevention of hypertension.

METHODS

Study Design

The study was a multicenter, randomized trial comparing the effects on blood pressure of three levels of sodium intake in two diets among adults whose blood pressure exceeded 120/80 mm Hg, including those with stage 1 hypertension (a systolic blood pressure of 140 to 159 mm Hg or a diastolic blood pressure of 90 to 95 mm Hg). The design of the trial, which was conducted from September 1997 through November 1999, has been described in detail elsewhere.[16] The three sodium levels were defined as high (a target of 150 mmol per day with an energy intake of 2100 kcal, reflecting typical consumption in the United States[1]), intermediate (a target of 100 mmol per day, reflecting the upper limit of the current national recommendations[1]), and low (a target of 50 mmol per day, reflecting a level that we hypothesized might produce an additional lowering of blood pressure). The daily sodium intake was proportionate to the total energy requirements of individual participants, so that larger or very active persons would receive more food and therefore more sodium than smaller or less active persons.

The two diets were a control diet[4] typical of what many people in the United States eat, and the DASH diet, which emphasizes fruits, vegetables, and low-fat dairy foods; includes whole grains, poultry, fish, and nuts; and contains smaller amounts of red meat, sweets, and sugar-containing beverages than the typical diet in the United States.[4,17] The DASH diet (originally termed the "combination diet"[4]) also contains smaller amounts of total and saturated fat and cholesterol and larger amounts of potassium, calcium, magnesium, dietary fiber, and protein than the typical diet.[4,17] The nutrient composition of the diets was calculated and monitored with the use of chemical analysis. Specific dietary patterns were composed to achieve the high, intermediate, and low levels of sodium in both the control and the DASH diets. Participants were provided with all of their food, including snacks and cooked meals. Taste tests were performed to ensure that the diets were palatable.

During a two-week run-in period, eligible persons ate the high-sodium control diet. Participants were then randomly assigned to follow one of the two diets according to a parallel-group design. They ate their assigned diet at each of the three sodium levels for 30 consecutive days in random order in a crossover design. Each participant's energy intake was adjusted to ensure that his or her weight remained constant throughout the study. Each of four clinical centers conducted the trial in four or five cohorts of participants. The primary outcome was systolic blood pressure at the end of each 30-day period of dietary intervention, and the secondary outcome was diastolic blood pressure. The study was approved by the human subjects committees of the centers, and written informed consent was given by each participant.

Criteria for Eligibility

To be eligible participants had to be at least 22 years old and to have an average systolic blood pressure on three screening visits of 120 to 159 mm Hg and an average diastolic blood pressure of 80 to 95 mm Hg. We targeted an enrollment that was 50 percent blacks and 50 percent women. The criteria for exclusion were heart disease, renal insufficiency, poorly controlled hyperlipidemia or diabetes mellitus, diabetes requiring insulin, special dietary requirements, intake of more than 14 alcoholic drinks per week, or the use of antihypertensive drugs or other medications that would affect blood pressure or nutrient metabolism.

Measurements

Blood pressure was measured with random-zero sphygmomanometers while participants were seated at three screening visits, twice during the run-in period, weekly during the first 3 weeks of each of the three 30-day intervention periods, and at five clinic visits during the last 9 days (at least two during the final 4 days) of each intervention period. During the screening period and during the last week of each intervention period, a 24-hour urine collection was obtained. The participants and the dietary staff were unaware of the outcome data; the personnel involved in the collection of the outcome data were unaware of participants' diet assignment. We assessed participants' adherence to the diet by reviewing their daily food diaries, having them eat their weekday lunches or dinners on site, and measuring 24-hour urinary excretion of sodium, potassium, phosphorus, and urea nitrogen. Side effects were monitored by means of questionnaires regarding symptoms and illnesses. According to the study protocol, a systolic blood pressure of more than 170 mm Hg or a diastolic blood pressure of more than 105 mm Hg at a single visit was considered to necessitate a second measurement; if the reading was sustained, the participant was referred to his or her physician for further evaluation and treatment.

Table 1. Base-Line Characteristics of the Participants.*

Characteristic	DASH Diet (N=208)	Control Diet (N=204)
Age (yr)	47±10	49±10
Female sex (%)	59	54
Race or ethnic group (%)		
Black	57	56
Non-Hispanic white	40	40
Asian or other	3	5
Hypertension (%)†	41	41
Blood pressure (mm Hg)‡		
Systolic	134±10	135±10
Diastolic	86±5	86±4
Body-mass index§	29±5	30±5
Waist circumference (cm)	96±12	100±14
Urinary sodium (mmol/day)¶	158±79	152±72
Educational level (%)		
High-school graduate	12	17
Attended college	41	32
College graduate	45	48
Annual income (%)		
<$30,000	32	34
$30,000–$60,000	33	41
>$60,000	35	25

*Plus–minus values are means ±SD. Because of rounding, not all percentages total 100.

†Hypertension was defined as an average systolic blood pressure of 140 to 159 mm Hg or an average diastolic blood pressure of 90 to 95 mm Hg during the three screening visits.

‡Base-line blood pressure was the average of three screening measurements and two measurements during the run-in period.

§The body-mass index is the weight in kilograms divided by the square of the height in meters.

¶Base-line urinary sodium was determined from a 24-hour urine collection during the screening period, when the participants were eating their customary, self-selected diets. Data were missing for four participants in the DASH-diet group.

Statistical Analysis

The analyses were structured according to a two-by-four design to compare the two diets (control and DASH) during the four periods (the run-in period and three intervention periods). The base-line blood pressure used for the analyses was the average of the measurements taken during the screening and run-in periods, and the blood pressure used for the end of each intervention period was the average of the last five measurements. A unified generalized-estimating-equation[18] model with an exchangeable covariance matrix was used for all primary analyses. Blood pressure was the outcome. The base-line blood pressure, the clinical center, and the cohort were represented in the model as fixed effects, whereas the intervention periods were included as random effects to allow for within-person correlation among blood-pressure measurements. The model included indicators of the cohort, the clinical center, and the carryover effect from the previous intervention. Results were similar with and without carryover in the model. Indicators for the subgroups specified in the study protocol (hypertensive status, race, and sex) and for the relevant interactions with the effects of the diet assignments and sodium levels were included in the subgroup analyses.

The linearity of the effects of sodium within the control diet or the DASH diet was assessed by comparing the decrease in blood pressure from the high to the intermediate level of sodium with the decrease from the intermediate to the low level of sodium. Multiple comparisons were accounted for by means of the method of Holm[19]; the resulting adjusted P values could be compared to 0.05 to determine significance.[20] The adjusted P values were used for the blood-pressure changes in the total cohort, but not in subgroups, as specified in the study protocol. All analyses were performed according to the intention-to-treat approach; in 22 instances, missing blood-pressure measurements during an intervention period, including those owing to a participant's withdrawal from the study, were replaced by base-line values. The planned sample size of 400 was calculated in order to provide the study with a power of 90 percent to detect a difference in systolic blood pressure of 2.1 mm Hg between sodium levels, and a difference of 3.0 mm Hg between the DASH and control diets at each sodium level.

RESULTS

The base-line characteristics of the participants are shown in Table 1. A total of 95 percent of the participants assigned to the DASH-diet group (198 of 208) and 94 percent of those assigned to the control-diet group (192 of 204) completed the study and provided blood-pressure measurements during each intervention period. Mean urinary sodium levels averaged 142 mmol per day during the high-sodium period, 107 mmol per day during the intermediate-sodium period, and 65 mmol per day during the low-sodium period (Table 2). The levels of urinary potassium, phosphorus, and urea nitrogen (reflective of the intake of fruit and vegetables, dairy products, and protein, respectively) were higher in the DASH-diet group than in the control-diet group, and were nearly identical for all three sodium levels. Weight remained stable, as intended.

The reduction of sodium intake significantly lowered systolic and diastolic blood pressure in a stepwise fashion, with both the control diet and the DASH diet (Fig. 1). The level of dietary sodium had approximately twice as great an effect on blood pressure with the control diet as it did with the DASH diet ($P<0.001$ for the interaction). There was a greater response of blood pressure to progressively lower levels of sodi-

Table 2. Urinary Excretion and Body Weight According to Dietary Sodium Level and Assigned Diet.

Variable	High Sodium Level		Intermediate Sodium Level		Low Sodium Level	
	DASH Diet	Control Diet	DASH Diet	Control Diet	DASH Diet	Control Diet
	mean ±SD					
Urinary excretion						
Sodium						
mmol/day	144±58	141±55	107±52	106±44	67±46	64±37
g/day	3.3±1.3	3.3±1.3	2.5±1.2	2.4±1.0	1.5±1.0	1.5±0.8
Potassium						
mmol/day	75±27	40±14	81±31	41±14	81±29	42±14
g/day	2.9±1.1	1.6±0.5	3.2±1.2	1.6±0.5	3.2±1.1	1.6±0.5
Phosphorus (mg/day)	778±285	666±248	825±350	646±264	783±286	672±243
Urea nitrogen (g/day)	11.5±4.0	9.6±3.2	12.4±4.5	9.7±3.4	11.8±4.1	10.0±3.3
Creatinine (g/day)	1.4±0.5	1.5±0.5	1.5±0.6	1.5±0.6	1.4±0.5	1.6±0.6
Weight (kg)	82.3±14.5	85.3±15.6	82.1±14.4	85.1±16.0	82.2±14.5	85.0±15.7

um intake. In the control diet, a reduction in the sodium intake of about 40 mmol per day from the intermediate sodium level lowered blood pressure more than a similar reduction in the sodium intake from the high level ($P=0.03$ for systolic blood pressure, $P=0.045$ for diastolic blood pressure).

The DASH diet, as compared with the control diet, resulted in a significantly lower systolic blood pressure at every sodium level and in a significantly lower diastolic blood pressure at the high and intermediate sodium levels (Fig. 1). It had a larger effect on both systolic and diastolic blood pressure at high sodium levels than it did at low ones ($P<0.001$ for the interaction).

As compared with the high-sodium control diet, the low-sodium DASH diet produced greater reductions in systolic and diastolic blood pressure than either the DASH diet alone or a reduction in sodium alone (Fig. 1). The reductions in blood pressure caused by the combination of dietary interventions were smaller than they would have been if the effects of each dietary intervention were strictly additive ($P<0.001$ for the interaction).

Reducing the sodium intake from the high to the low level, with either the control diet or the DASH diet, reduced systolic blood pressure in participants with and in those without hypertension (among blacks as well as among participants of other races or ethnic groups), and in men and women (Fig. 2). The effects of sodium were greater in participants with hypertension than in those without hypertension (interaction, $P=0.01$ on the control diet; $P=0.003$ on the DASH diet), in blacks on the control diet than in participants of other races or ethnic groups on that diet ($P=0.007$), and in women on the DASH diet than in men on that diet ($P=0.04$). As compared with the combination of the control diet and a high level of sodium, the combination of the DASH diet and a low level of sodium lowered systolic blood pressure by 11.5 mm Hg in participants with hypertension (12.6 mm Hg for blacks; 9.5 mm Hg for others), by 7.1 mm Hg in participants without hypertension (7.2 mm Hg for blacks; 6.9 mm Hg for others), and by 6.8 mm Hg in men and 10.5 mm Hg in women ($P<0.001$ in all subgroups). The combination of the two dietary interventions lowered systolic blood pressure more in participants with hypertension than in those without hypertension ($P=0.004$), and more in women than in men ($P=0.02$).

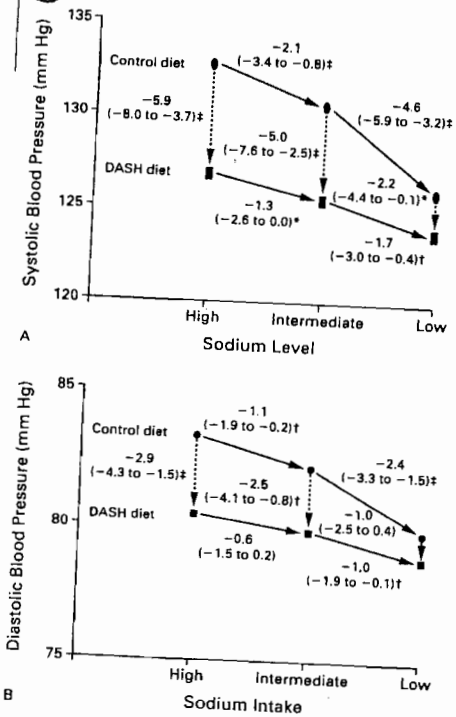


Figure 1. The Effect on Systolic Blood Pressure (Panel A) and Diastolic Blood Pressure (Panel B) of Reduced Sodium Intake and the DASH Diet.

The mean systolic and diastolic blood pressures are shown for the high-sodium control diet. The mean changes in blood pressure are shown for various sodium levels (solid lines), and the mean differences in blood pressure between the two diets at each level of sodium intake are shown. Unidirectional arrows are used for simplicity, although the order in which participants were given the sodium levels was random with a crossover design. The numbers next to the dotted lines connecting the data points are the mean changes in blood pressure. The 95 percent confidence intervals are given in parentheses. There was a significant difference in systolic blood pressure between the high-sodium and low-sodium phases of the control diet (mean, −6.7 mm Hg; 95 percent confidence interval, −5.4 to −8.0; $P<0.001$) and the DASH diat (mean, −3.0 mm Hg; 95 percent confidence interval, −1.7 to −4.3; $P<0.001$) and between the high-sodium phase of the control diet and the low-sodium phase of the DASH diet (mean, −8.9 mm Hg; 95 percent confidence interval, −6.7 to −11.1; $P<0.001$). There was also a significant difference in diastolic blood pressure between the high-sodium and low-sodium phases of the control diet (mean, −3.5 mm Hg; 95 percent confidence interval, −2.6 to −4.3; $P<0.001$) and of the DASH diet (mean, −1.6 mm Hg; 95 percent confidence interval, −0.8 to −2.5; $P<0.001$) and between the high-sodium phase of the control diet and the low-sodium phase of the DASH diet (mean, −4.5 mm Hg; 95 percent confidence interval, −3.1 to −5.9; $P<0.001$). Asterisks ($P<0.05$), daggers ($P<0.01$), and double daggers ($P<0.001$) indicate significant differences in blood pressure between groups or between dietary sodium categories.

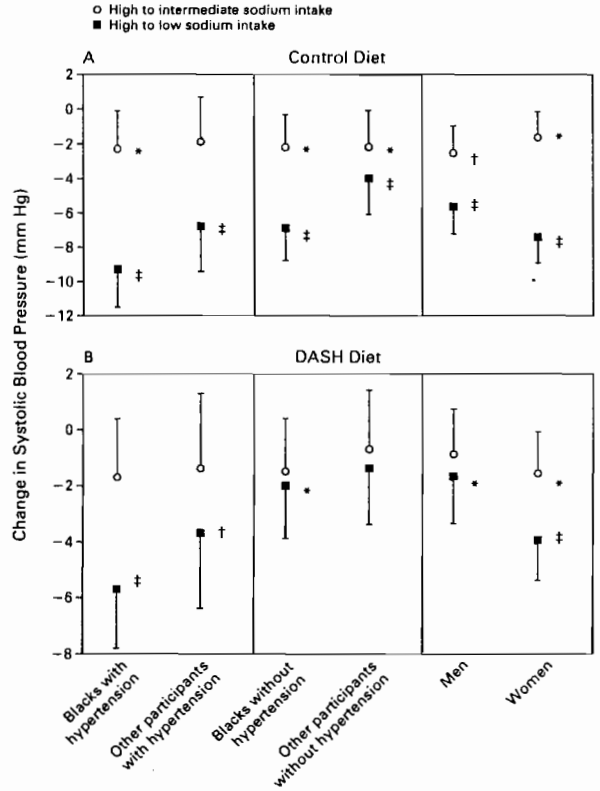


Figure 2. The Effect on Systolic Blood Pressure of Dietary Sodium Intake during the Control Diet (Panel A) and the DASH Diet (Panel B), According to Subgroup.

The error bars represent the 95 percent confidence limits of the changes in blood pressure for each subgroup. Hypertension was defined as an average systolic blood pressure of 140 to 159 mm Hg or an average diastolic blood pressure of 90 to 95 mm Hg during the three screening visits. The "other" category of race and ethnic group is composed primarily of non-Hispanic whites (see Table 1). Asterisks ($P\leq0.05$), daggers ($P<0.01$), and double daggers ($P<0.001$) indicate significant differences between levels of sodium inteke.

A systolic blood pressure of more than 170 mm Hg or a diastolic blood pressure of more than 105 mm Hg occurred in 36 participants in the control-diet group and in 7 in the DASH-diet group; in 18 participants during the period of high sodium intake, 22 during intermediate sodium intake, and 3 during low sodium intake; and in no participant during the low-sodium phase of the DASH diet. None of these participants reached the predefined threshold for sustained elevated blood pressure[16] that necessitated referral for antihypertensive pharmacologic therapy. The participants tended to report fewer symptoms during periods of reduced sodium intake. Headache was reported at least once by 47 percent of the participants during the high-sodium phase of the control diet, by 39 percent during the low-sodium phase of the control diet, and by 36 percent during the low-sodium phase of the DASH diet ($P<0.05$ for both comparisons with the high-

sodium phase of the control diet). The number of participants who did not complete an intervention period was similar during all three sodium levels (seven during the high-sodium phase, seven during the intermediate-sodium phase, and eight during the low-sodium phase).

DISCUSSION

This trial produced several key findings that are important for the prevention and treatment of hypertension. First, the DASH diet lowered blood pressure at high, intermediate, and low levels of sodium intake, confirming and extending the findings of the previous DASH study.[4] Thus, the benefits of following the DASH diet have now been shown to apply throughout the range of sodium intakes, including those recommended for the prevention and treatment of hypertension. Second, blood pressure can be lowered in the consumers of either a diet that is typical in the United States or the DASH diet by reducing the sodium intake from approximately 140 mmol per day (an average level in the United States) to an intermediate level of approximately 100 mmol per day (the currently recommended upper limit[1]), or from this level to a still lower level of 65 mmol per day. Moreover, reducing the sodium intake by approximately 40 mmol per day caused a greater decrease in blood pressure when the starting sodium intake was already at the recommended level than when it was at a higher level similar to the average in the United States. These results provide a scientific basis for a lower goal for dietary sodium than the level currently recommended.

Third, the combined effects on blood pressure of a low sodium intake and the DASH diet were greater than the effects of either intervention alone and were substantial. In participants with hypertension, the effects were equal to or greater than those of single-drug therapy.[21,22] The combined effects were not as great as would be estimated on the basis of strict additivity, perhaps because low levels of sodium attenuated the hypotensive effects of potassium in the DASH diet[23,24] or because the high potassium or calcium content of the DASH diet attenuated the effects of low levels of sodium.[25-27] Nevertheless, the combination of the two interventions achieved the greatest effect on blood pressure, and therefore, both — not just one or the other — merit recommendation. The DASH diet and the low sodium level were well tolerated, with no increase in symptoms or dropouts. However, long-term health benefits remain to be demonstrated and will depend on the ability of people to make long-lasting dietary changes, including the consistent choice of lower-sodium foods.

We found that the reduction of dietary sodium significantly lowered the blood pressure of persons without hypertension who were eating a diet that is typical in the United States. These results should settle the controversy over whether the reduction of sodium has a worthwhile effect on blood pressure in persons without hypertension. This controversy stemmed in part from the apparently divergent results and interpretations of individual trials and meta-analyses.[6-8] Because of differences in the designs, quality, and study populations of the trials and the subjectivity involved in judgments about which studies to include in meta-analyses, a single, large, well-controlled trial with a diverse population provides the most reliable estimates of the effects of treatments.

In our study the dietary intake was controlled and the influences of behavioral factors, programs of dietary education, and varying degrees of adherence to the diets were minimized, so that we measured only true biologic effects. This method offers the optimal approach for determining the effects of diet on blood pressure. The variation in the results in persons without hypertension among previous trials and meta-analyses were probably caused in large part by variable adherence to the prescribed reduction in sodium, inadequate trial design, small samples, or limitations in analysis and presentation, rather than by the lack of a biologic effect of sodium on blood pressure.

We found that the level of dietary sodium and assignment to the control or the DASH diet each had a substantial effect on the blood pressure of blacks, confirming previous findings.[28-33] Blacks have a higher rate of hypertension and the resulting cardiovascular disease than other racial and ethnic groups in the United States. We speculate that a greater sensitivity to the deleterious effects of diet could contribute to the high prevalence of hypertension in blacks. These findings justify the intensification of public health and therapeutic efforts to induce dietary change among blacks.

The attainment of a lower sodium level in the population as a whole presents challenges, since sodium is widely prevalent in the food supply, and since most of the daily sodium intake comes from salt in processed foods rather than from table salt.[34] The first report on U.S. dietary goals by the Senate Select Committee for Nutrition and Human Needs recommended a goal of 3 g of sodium chloride per day (52 mmol of sodium),[35] but concern about the feasibility of achieving this goal led to an increase of the goal to 5 g of sodium chloride.[36] Hence, efforts to reduce sodium intake must ultimately rely both on consumers' selection of low-sodium foods and, perhaps more important, on the increased availability of low-sodium products.

Our results should be applicable to most people in the United States. Approximately 50 percent of the adult population of the United States and 80 percent of those 50 years of age or older have a blood pressure of at least 120/80 mm Hg,[37] which is the upper limit of optimal blood pressure[1] and which was the lower limit of the eligibility requirements for blood pressure for our trial. Furthermore, epidemiologic studies suggest that diets low in sodium and high in potassium blunt the rise in blood pressure that normally occurs with age.[9] The intervention periods in our trial were, of necessity, brief — just 30 days. Still, the effect of the reduction in dietary sodium on blood pressure tends to persist over time to the extent that adherence to the lower-sodium diet is maintained.[7,15,38] In conclusion, our results provide support for a more aggressive target for reduced sodium intake, in combination with use of the DASH diet, for the prevention and treatment of elevated blood-pressure levels.

Supported by cooperative agreements and grants from the National Heart, Lung, and Blood Institute (U01-HL57173, to Brigham and Women's Hospital; U01-HL57114, to Duke University; U01-HL57190, to Pennington Biomedical Research Institute; U01-HL57139 and K08 HL03857-01, to Johns Hopkins University; and U01-HL57156, to Kaiser Permanente Center for Health Research) and by the General Clinical Research Center Program of the National Center for Research Resources (M01-RR02635, to Brigham and Women's Hospital, and M01-RR00722, to Johns Hopkins University).

We are indebted to the study participants for their sustained commitment to the DASH–Sodium Trial; to the Almond Board of California, Beatrice Foods, Bestfoods, Cabot Creamery, C.B. Foods, Dannon, Diamond Crystal Specialty Foods, Elwood International, Hershey Foods, Hormel Foods, Kellogg, Lipton, McCormick, Nabisco U.S. Foods Group, Procter & Gamble, Quaker Oats, and Sun-Maid Growers for donating food; to Frost Cold Storage for food storage; to the members of the external Protocol Review Committee — Janice A. Derr, Ph.D., Richard D. Mattes, Ph.D., Lemuel A. Moye, M.D., Ph.D., Jeremiah Stamler, M.D. (chair), and Jackson T. Wright, M.D., Ph.D.; and to the members of the Data and Safety Monitoring Board — Asrat Cnaan, Ph.D., Janice A. Derr, Ph.D., Richard Grimm, M.D. (chair), Richard D. Mattes, Ph.D., Jeremiah Stamler, M.D., and Jackson T. Wright, M.D., Ph.D.

APPENDIX

In addition to the authors, other members of the DASH–Sodium Collaborative Research Group include the following: *Brigham and Women's Hospital and Harvard Medical School, Boston (clinical center)* — N. Alexander, J. Belmonte, F. Bodden, L. Cashman, B. Cox, J. Dyer, A. Ghosh, J. Hackett, E. Hamilton, T. Holiday, J. Karimbakis, C. Larson, M. McCullough, D. McDonald, P. McVinney, D. Moeller, P. Morris, M. Myrie, K. Osborn, B. Penachio, S. Redican, J. Sales, J. Swain, Z. Than, K. Weed; *Duke University Medical Center, Durham, N.C. (clinical center)* — J. Abbott, K. Aicher, J. Ard, J. Baughman, M. Baughman, B. Brown, A. Bohannon, B. Graves, K. Hoben, J. Huang, L. Johnson-Pruden, T. Phelps, C. Plaisted, L. Reams, P. Reams, T. Ross, P. Rukenbrod, E. Ward; *Johns Hopkins University, Baltimore (clinical center)* — J. Abshere, D. Bengough, L. Bohlman, J. Charleston, L. Clement, C. Dahne, F. Dennis, S. Dobry, K. Eldridge, T. Erlinger, A. Fouts, C. Harris, B. Horseman, M. Jehn, S. Kritt, J. Lambert, E. Levitas, P. McCarron, N. Muhammad, M. Nagy, B. Peterson, D. Rhodes, V. Shank, T. Shields, T. Stanger, A. Thomas, E. Thomas, L. Thomas, R. Weiss, E. Wilke, W. Wong; *Kaiser Permanente Center for Health Research, Portland, Oreg. (coordinating center)* — M. Allison, S. Baxter, N. Becker, S. Craddick, B. Doster, C. Eddy, D. Ernst, A. Garrison, S. Gillespie, R. Gould, T. Haswell, L. Haworth, F. Heinith, M. Hornbrook, K. Kirk, P.A. LaChance, R. Laws, M. Leitch, W.R. Li, L. Massinger, M. McMurray, G. Meltesen, G. Miranda, S. Mitchell, N. Redmond, J. Reinhardt, J. Rice, P. Runk, R. Schuler, C. Souvanlasy, M. Sucec, T. Vogt; *National Heart, Lung, and Blood Institute, Bethesda, Md. (sponsor)* — C. Brown, M. Evans; *Pennington Biomedical Research Center, Baton Rouge, La. (clinical center)* — C. Champagne, S. Crawford, F. Greenway, J. Ihrig, B. Kennedy, J. Perault, D. Sanford, A. Sawyer, S. Smith, R. Tulley, J. Vaidyanathan; *Virginia Polytechnic Institute, Blacksburg, Va. (food-analysis coordinating center)* — K. Phillips; *Washington University School of Medicine, St. Louis (core laboratory)* — T. Cole.

REFERENCES

1. The sixth report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. Arch Intern Med 1997;157:2413-46. [Erratum, Arch Intern Med 1998;158:573.]
2. National High Blood Pressure Education Program Working Group report on primary prevention of hypertension. Arch Intern Med 1993;153:186-208.
3. Stamler J, Stamler R, Neaton JD. Blood pressure, systolic and diastolic, and cardiovascular risks: US population data. Arch Intern Med 1993;153:598-615.
4. Appel LJ, Moore TJ, Obarzanek E, et al. A clinical trial of the effects of dietary patterns on blood pressure. N Engl J Med 1997;336:1117-24.
5. Facts about the DASH diet. Bethesda, Md.: National Heart, Lung, and Blood Institute, 1998. (NIH publication no. 98-4082.)
6. Cutler JA, Follmann D, Allender PS. Randomized trials of sodium reduction: an overview. Am J Clin Nutr 1997;65:Suppl:643S-651S.
7. Law MR, Frost CD, Wald NJ. By how much does dietary salt reduction lower blood pressure? III. Analysis of data from trials of salt reduction. BMJ 1991;302:819-24. [Erratum, BMJ 1991;302:939.]
8. Graudal NA, Galloe AM, Garred P. Effects of sodium restriction on blood pressure, renin, aldosterone, catecholamines, cholesterols, and triglyceride: a meta-analysis. JAMA 1998;279:1383-91.
9. Intersalt Cooperative Research Group. Intersalt: an international study of electrolyte excretion and blood pressure: results for 24 hour urinary sodium and potassium excretion. BMJ 1988;297:319-28.
10. Hatch FT, Wertheim AR, Eutman GH, Watkin DM, Froeb HF, Epstein HA. Effects of diet in essential hypertension. III. Alterations in sodium chloride, protein and fat intake. Am J Med 1954;17:499-513.
11. Chapman CB, Gibbons T, Henschel A. The effect of the rice-fruit diet on the composition of the body. N Engl J Med 1950;243:899-905.
12. Parfrey PS, Markandu ND, Roulston JE, Jones BE, Jones JC, MacGregor GA. Relation between arterial pressure, dietary sodium intake, and renin system in essential hypertension. BMJ 1981;283:94-7.
13. Gill JR, Gullner HG, Lake CR, Lakatua DJ, Lan G. Plasma and urinary catecholamines in salt-sensitive idiopathic hypertension. Hypertension 1988;11:312-9.
14. Kawasaki T, Delea CS, Bartter FC, Smith H. The effect of high-sodium and low-sodium intakes on blood pressure and other related variables in human subjects with idiopathic hypertension. Am J Med 1978;64:193-8.
15. MacGregor GA, Markandu ND, Sagnella GA, Singer DR, Cappuccio FP. Double-blind study of three sodium intakes and long-term effects of sodium restriction in essential hypertension. Lancet 1989;2:1244-7.
16. Svetkey LP, Sacks FM, Obarzanek E, et al. The DASH diet, sodium intake and blood pressure trial (DASH-sodium): rationale and design. J Am Diet Assoc 1999;99:Suppl:S96-S104.
17. Karanja NM, Obarzanek E, Lin PH, et al. Descriptive characteristics of the dietary patterns used in the Dietary Approaches to Stop Hypertension Trial. J Am Diet Assoc 1999;99:Suppl:S19-S27.
18. Liang K-Y, Zeger SL. Longitudinal data analysis using generalized linear models. Biometrika 1986;73:13-22.
19. Holm S. A simple sequentially rejective multiple test procedure. Scand J Stat 1979;6:65-70.
20. Aickin M, Gensler H. Adjusting for multiple testing when reporting research results: Bonferroni vs Holm methods. Am J Public Health 1996;86:726-8.
21. Materson BJ, Reda DJ, Cushman WC, et al. Single-drug therapy for hypertension in men: a comparison of six antihypertensive agents with placebo. N Engl J Med 1993;328:914-21. [Erratum, N Engl J Med 1994;330:1689.]
22. The Treatment of Mild Hypertension Research Group. The treatment of mild hypertension study: a randomized, placebo-controlled trial of a nutritional-hygienic regimen along with various drug monotherapies. Arch Intern Med 1991;151:1413-23.
23. Grimm RH Jr, Neaton JD, Elmer PJ, et al. The influence of oral potassium chloride on blood pressure in hypertensive men on a low-sodium diet. N Engl J Med 1990;322:569-74.
24. Whelton PK, He J, Cutler JA, et al. The effects of oral potassium on blood pressure: meta-analysis of randomized controlled clinical trials. JAMA 1997;227:1624-32.
25. Krishna GG, Miller E, Kapoor S. Increased blood pressure during potassium depletion in normotensive men. N Engl J Med 1989;320:1177-82.
26. Saito K, Sano H, Furuta Y, Fukuzaki H. Effect of oral calcium on blood pressure response in salt-loaded borderline hypertensive patients. Hypertension 1989;13:219-26.
27. Rich GM, McCullough M, Olmedo A, Malarick C, Moore TJ. Blood pressure and renal blood flow responses to dietary calcium and sodium intake in humans. Am J Hypertens 1991;4:642S-645S.
28. Luft FC, Grim CE, Fineberg N, Weinberger MC. Effects of volume expansion and contraction in normotensive whites, blacks, and subjects of different ages. Circulation 1979;59:643-50.
29. Miller JZ, Weinberger MH, Daugherty SA, Fineberg NS, Christian JC, Grim CE. Heterogeneity of blood pressure response to dietary sodium restriction in normotensive adults. J Chronic Dis 1987;40:245-50.
30. Voors AW, Dalferes ER Jr, Frank GC, Aristimuno GG, Berenson GS.

29. Agmon Y, Khandheria BK, Meissner I, et al. Frequency of atrial septal aneurysm in patients with cerebral ischemic events. *Circulation*. 1999;99:1942–1944.
30. Mas JL, Zuber M, for the French Study Group on Patent Foramen Ovale and Atrial Septal Aneurysm. Recurrent cerebrovascular events in patients with patent foramen ovale, atrial septal aneurysm, or both and cryptogenic stroke or transient ischemic attack. *Am Heart J*. 1995;130:1083–1088.
31. Comess KA, De Rook FA, Beach KW, et al. Transesophageal echocardiography and carotid ultrasound in patients with cerebral ischemia: prevalence of findings and recurrent stroke risk. *J Am Coll Cardiol*. 1994;23:1598–1603.
32. Cabanas L, Mas JL, Cohen A, et al. Atrial septal aneurysm and patent foramen ovale as risk factors for cryptogenic stroke in patients less than 55 years of age. A study using transesophageal echocardiography. *Stroke*. 1993;24:1865–1873.
33. Rosenow EC, Osmundson PJ, Brown MJ. Pulmonary embolism. *Mayo Clin Proc*. 1981;56:161–178.
34. Hirsch J. Diagnosis of venous thrombosis and pulmonary embolism. *Am J Cardiol*. 1990;65:45C–49C.
35. Hull RD, Hirsch J, Carter CJ, et al. Pulmonary angiography, ventilation lung scanning, and venography for clinically suspected pulmonary embolism with abnormal perfusion lung scan. *Ann Intern Med*. 1983;98:891–899.
36. Mohr JP. Cryptogenic stroke. *N Engl J Med*. 1988;318:1197–1198.
37. Griffin JH, Evatt B, Wideman G, Fernandez JA. Anticoagulant protein C pathway defective in majority of thrombophilic patients. *Blood*. 1993;82:1989–1993.
38. Martinelli I, Sacchi E, Landi G, et al. High risk of cerebral-vein thrombosis in carriers of a prothrombin-gene mutation and in users of oral contraceptives. *N Engl J Med*. 1998;338:1793–1797.
39. Heijboer H, Brandjes DPM, Büller HR, et al. Deficiencies of coagulation-inhibiting and fibrinolytic proteins in outpatients with deep-vein thrombosis. *N Engl J Med*. 1990;323:1512–1516.
40. Pabinger I, Brucker S, Kyrle PA, et al. Hereditary deficiency of antithrombin III, protein C and protein S: prevalence in patients with a history of venous thrombosis and criteria for rational patient screening. *Blood Coagul Fibrinolysis*. 1992;3:547–553.
41. Lensen RP, Rosendaal FR, Koster T, et al. Apparent different thrombotic tendency in patients with factor V Leiden and protein C deficiency due to selection of patients. *Blood*. 1996;88:4205–4208.
42. Bertina RM, Rosendaal FR. Venous thrombosis—the interaction of genes and environment. *N Engl J Med*. 1998;338:1840–1841.
43. Mayer EI, Jacobson DW, Robinson K. Homocysteine and coronary atherosclerosis. *J Am Coll Cardiol*. 1996;27:517–527.
44. Brattström L, Tengborn L, Israelsson B, Hultberg B. Plasma homocysteine in venous thromboembolism. *Haemostasis*. 1991;21:51–57.
45. Dawson S, Henney A. The status of PAI-1 as a risk factor for arterial and thrombotic disease: a review. *Atherosclerosis*. 1992;95:105–117.
46. Alving BM, Comp PC. Recent advances in understanding clotting and evaluating patients with recurrent thrombosis. *Am J Obstet Gynecol*. 1992;167:1184–1191.
47. Kniffin WD, Baron JA, Barett J, et al. The epidemiology of diagnosed pulmonary embolism and deep venous thrombosis in the elderly. *Arch Intern Med*. 1994;154:861–866.

(323) The Effects of Transdermal Estradiol on the [377] Response to Mental Stress in Postmenopausal Women: A Randomized Trial

Graziano Ceresini, MD, Marilena Freddi, MD, Simonetta Morganti, MD, Isabella Rebecchi, MD, Alberto Bacchi Modena, MD, Maurizio Rinaldi, MD, Carlo Manca, MD, Alberto Amaducci, MD, Graziano Del Rio, MD, Giorgio Valenti, MD

PURPOSE: Estrogens inhibit adrenomedullary catecholamine release and catecholamine-mediated responses to stress. We examined whether estrogen supplementation reduces the sympathoadrenal response to mental stress in postmenopausal women.

MATERIALS AND METHODS: We compared the effects of 3-week treatment with transdermal 17-beta-estradiol and placebo in 10 postmenopausal women using a randomized, blinded, crossover design. We measured plasma catecholamine levels and the cardiovascular and metabolic responses to a 15-minute stress with mental arithmetic. Treatments were compared using repeated measures analysis of variance.

RESULTS: During placebo treatment, mean (± SD) epinephrine levels reached a peak of 431 ± 135 pmol/liter after 15 minutes of stress; the epinephrine response was blunted during estradiol treatment, with a peak of 357 ± 77 pmol/liter ($P < 0.05$). Estradiol also blunted the diastolic blood pressure response to stress (baseline levels of 78 ± 15 mm Hg vs peak of 90 ± 6 mm Hg during placebo; baseline of 80 ± 8 mm Hg vs peak of 84 ± 6 mm Hg during estradiol; $P < 0.05$). Estradiol treatment also blunted the decrease in the standard deviation of the mean of the electrocardiographic RR intervals and the increase in the ratio between the low-frequency and high-frequency bandwidths.

CONCLUSION: We observed a moderate, although significant, reduction in markers of the stress response to mental arithmetic in postmenopausal women treated with transdermal 17-beta-estradiol. **Am J Med. 2000;109:463–468.**

Sex hormones affect the activity of the sympathoadrenal system. The catecholamine responses to various stimuli, such as physical exercise (1,2), hypoglycemia (3), and mental stress (1), are lower in women than in men. There is a reduction in the stress-induced catecholamine response in men after estrogen administration (4). Estrogens reduce catecholamine release from adrenomedullary cells (5) and affect the enzymatic pathways regulating catecholamine synthesis and degradation (6,7). They may also modulate the activity of both alpha- and beta-adrenoreceptors (8–10).

The menopause-related decline in circulating levels of estrogens is associated with an increased cardiovascular risk (11), and postmenopausal women have a greater cardiovascular response to certain stressful stimuli than do premenopausal women (12). However, although observational studies have reported a significant decrease in cardiovascular events in women who use postmenopausal estrogen therapy (13,14), a randomized trial found that estrogens were not effective for the secondary prevention of coronary heart disease in postmenopausal women (15). Moreover, interim results from the Women's Health Initiative, a primary prevention trial, have suggested a potential increase in the risk of heart disease during postmenopausal estrogen therapy (16). Therefore, although estrogens have beneficial effects on serum lipid levels and endothelial and vasomotor function (17,18), the overall effects of postmenopausal hormonal replacement therapy are uncertain.

One unsolved question is whether estrogens affect the cardiovascular system by decreasing the sympathoadrenal response to acute stress. Investigators have reported that estrogens reduce both basal and stress-induced sympathetic tone (19,20), but others have reported no effects on cardiovascular responses to stress (21,22). Different regimens, including dose, duration, and route of administration, of estrogen were used in these studies. Therefore, we studied the effects of 17-beta-estradiol, administered transdermally for 3 weeks, on plasma catecholamine levels and the cardiovascular and metabolic responses to a mental stress challenge in postmenopausal women.

MATERIAL AND METHODS

The study was performed in 10 postmenopausal women, who had a mean age of 53.9 ± 1.1 years, a mean body

From the Department of Internal Medicine and Biomedical Sciences (GC, MF, SM, IR, CM, GV), and the Institute of Gynecology (ABM, MR), University of Parma, Parma; and the Division of Cardiology, Regina Margherita Hospital (AA), Torino; and the Department of Internal Medicine, University of Modena (GDR), Modena, Italy.
Requests for reprints should be addressed to Graziano Ceresini, MD, Department of Internal Medicine and Biochemical Sciences, Section of Geriatrics, University of Parma, Via don Bosco 2, 43100 Parma, Italy.
Manuscript submitted December 28, 1999, and accepted in revised form June 5, 2000.

0002-9343/00/$–see front matter
PII S0002-9343(00)00523-4

mass index of 25.2 ± 1.3 kg/m^2, and few or no vasomotor symptoms. All women had undergone natural (nonsurgical) menopause. Menopausal status was confirmed (serum levels of follicle-stimulating hormone greater than 40 mU/mL and estradiol less than 60 pmol/liter). All subjects were nonsmokers, had normal serum lipid levels, and had no history of cardiovascular disease. They had a low caffeine intake and took no medications. None had ever used estrogen replacement therapy. All of the subjects answered the Kellner Symptom Questionnaire (23), the Paykel Life Events Scale (24), and the Spielberg State-Trait Anxiety Inventory (25); those with identified psychological problems or who had experienced stressful life events in the previous 6 months were excluded. Subjects gave their informed consent to participate in the study. The protocol was approved by the Ethics Committee of the University of Parma.

Design

Subjects were treated with 17-beta-estradiol or placebo for 3 weeks, according to a randomized, double-blind, crossover design with a 3-week wash-out between the two treatments; 6 women were randomly assigned to receive estrogen as the first treatment. Both treatments were administered by skin patches (Rotta Research, Milan, Italy), which were renewed every 3 days. The release of 17-beta-estradiol was 50 μg per day. At the end of each treatment, the subjects underwent a mental stress test.

Mental Stress Test

In preliminary sessions, the subjects were familiarized with the serial subtraction test and the digit span test. Both sessions of mental stress were performed in the morning between 9 and 10 AM. The subjects ate a light breakfast at least 2 hours before the experiment, avoiding substances known to affect the sympathoadrenal system. They were seated in a comfortable chair in a quiet testing chamber. A 19-gauge intravenous catheter was placed in a superficial vein of the hand; the hand and the distal forearm were then placed in a box heated at a constant temperature of 55° C to obtain arterialized venous blood samples. The catheter was kept patent with a slow infusion of 0.9% saline. A cuff for automatic recording of blood pressure (A&D International, Ltd., Tokyo, Japan) was placed in the arm contralateral to the one used for blood sampling. In addition, continuous automatic electrocardiographic (ECG) monitoring was performed using a three-channel amplitude modulated tape recorder (Cardio Corder; Del Mar Avionics, Irvine, California). The stress test was divided into three consecutive parts: an initial 30-minute period of resting, a 15-minute mental stress, and a final 15-minute recovery period in resting conditions. During the last 15 minutes of the initial rest period, two (−15 and 0 minute) blood samples were collected, and systolic and diastolic blood pressures were recorded.

The 15-minute mental stress was divided into three 5-minute periods, each consisting of two different tasks: a 4-minute serial subtraction test and a 1-minute digit span test. The serial subtraction test required subjects to subtract 17 from a four-digit number. The digit span test required subjects to remember numbers in the correct sequence forward during the first and the third periods, and backward during the second period of the stress challenge. The length of the items was varied to make the difficulty of the task unpredictable. The numbers used were the same as those in the Digit Span Test of the Wechsler Adult Intelligent Scale (26). The subjects performed the tests as fast as possible and were accompanied by a metronome; they were not allowed to use paper or pencil during the procedure. If they made a mistake, they were asked to start again from the beginning. Blood samples were collected during the mental stress at 5, 10, and 15 minutes from the beginning of the challenge. At each sample time, systolic and diastolic blood pressures were recorded.

At the end of the stress challenge, subjects were left in resting conditions for 15 minutes, after which a blood sample was collected and blood pressure was recorded; soon after, the continuous automatic ECG monitoring was stopped. The subjects rated the overall difficulty of the tests using a five-point scale (with 1 indicating easy and 5 indicating very hard). To assess the level of anxiety caused by participation in the study, subjects completed the Spielberg State-Trait Anxiety Scale before and at the end of the stress challenge.

Measurements of Hormonal and Biochemical Variables in Blood Samples

Blood samples for measuring plasma epinephrine and norepinephrine levels were collected at each sample time in tubes containing glutathione and ethyleneglycol-bis-(beta-aminoethyl ether)-N,N,N′,N′-tetraacetic acid (EGTA) and immediately put on ice; after separation in a refrigerated centrifuge, plasma was stored at −80° C until assayed. Plasma catecholamine levels were evaluated using a reverse-phase high-performance liquid chromatography method with electrochemical detection (4). Blood glucose levels were determined with a glucose analyzer (Beckman, Palo Alto, California). Serum insulin levels were measured by radioimmunoassay (Sorin, Milan, Italy). Serum levels of free fatty acids were determined in basal conditions and at the end of both the 15-minute stress test and the recovery period by a spectrophotometer (Beckman, Palo Alto, California). Serum levels of 17-beta-estradiol were evaluated on basal samples that were collected in both sessions before the stress test by radioimmunoassay using a commercially available kit (Sorin, Milan, Italy). All samples were run in duplicate. The intra- and interassay coefficients of variation were less than 10% for all of the measurements.

Analysis of Continuous ECG Recording

Three leads were used for monitoring. The Strata Scan Holter Analysis System program (Del Mar Avionics, Irvine, California) was used for the tape analysis. Heart rate was expressed as the mean of the RR intervals. A histogram of the consecutive RR ratio was examined, and cycles outside 80% to 120% of preceding RR intervals were excluded to avoid the interference of artifacts, premature beats, or postextrasystolic pauses (27,28). All tapes were then analyzed to measure heart rate variability in both the time and frequency domains. Recordings were analyzed by one investigator who was unaware of treatment. In the time domain of heart rate variability, we calculated the standard deviation of the mean of all normal RR intervals in the following three periods: the final 15 minutes of the initial resting conditions, the 15-minute stress challenge, and the 15-minute recovery period. In the same intervals, we also evaluated the frequency domain of heart rate variability; spectral measurements were computed by fast-Fourier transform analysis. Spectral plots were used to identify two subsets of the frequency domain: low frequency (0.05 to 0.15 Hz) and high frequency (0.15 to 1.35 Hz). Spectral power was quantified in these two frequency bandwidths. Spectral plots were squared to quantify power in the two frequency bands (in milliseconds squared). The ratio of low-frequency to high-frequency power was calculated.

Data Analysis

A preliminary analysis with the Shapiro-Wilk W statistic test was performed to determine whether the data conformed to a normal distribution, and the homogeneity of variance was computed by Bartlett's test. Data were analyzed using repeated measures analysis of variance in which the effects of both time and treatment were evaluated. There were no statistically significant interactions between order of treatment and effects on stress response (all $P < 0.19$); thus results were combined. If an F value was significant ($P < 0.05$), Student's t test was used to compare means between the groups. When data were not normally distributed, analysis was performed using the Friedman rank test followed by the Wilcoxon signed rank test to identify differences between distributions; these tests were used to evaluate the effects of treatment on free fatty acid levels. All statistical calculations were made using SPSS software (SPSS, Chicago, Illinois) (29). Continuous values are expressed as mean ± SD.

RESULTS

Serum estradiol levels were increased at the end of the treatment with the 17-beta-estradiol patch (placebo 30 ± 8 pmol/liter vs estradiol 171 ± 25 pmol/liter, $P < 0.003$). Baseline Spielberg State Anxiety scores were similar during the two treatments (placebo 40 ± 3 vs estradiol 43 ± 5, $P = 0.72$) and were not affected by the mental stress challenge (placebo, 41 ± 4; estradiol 41 ± 5, $P = 0.61$). The subjects rated the test as being equally difficult (on a 1-to-5 scale) during both the placebo and the estradiol sessions (placebo 3 ± 0.6 vs estradiol 3 ± 1, $P = 0.21$).

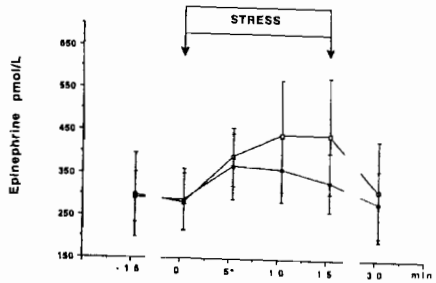


Figure 1. Mean (± SD) plasma epinephrine levels before, during, and after mental stress after 3-week treatment with placebo (open square) or estradiol (filled circle) in 10 postmenopausal women. Repeated measures analysis of variance revealed a significant ($P < 0.05$) difference between treatments.

Evaluations of Circulating Levels of Catecholamines and Other Hormonal and Biochemical Variables

Basal plasma epinephrine levels were similar during the estradiol and placebo treatments. In response to the mental stress (Figure 1), epinephrine levels increased significantly during placebo treatment, from basal values of 273 ± 67 pmol/liter to a peak of 431 ± 135 pmol/liter measured at the end of the stress challenge. The epinephrine response was less marked during estradiol treatment. From a baseline of 279 ± 72 pmol/liter, the values reached a peak of 357 ± 77 pmol/liter ($P < 0.05$ compared with baseline) 5 minutes after the beginning of stress, with a subsequent continuous decrease. There was a significant ($P < 0.05$) difference in the effects of placebo and estradiol on the stress-induced epinephrine responses using repeated measures analysis of variance.

There were no significant differences in basal levels of plasma norepinephrine during the two treatments. In response to the stress stimulus, norepinephrine levels increased during both treatments; however, these changes were not statistically significant (Figure 2).

Basal serum glucose and insulin levels were similar during the two treatments and were not significantly affected by the mental stress challenge during either placebo or 17-beta-estradiol treatment. Serum free fatty acid levels increased significantly in response to the stress challenge during the administration of both placebo [peak values of 667 ± 163 μM vs basal values of 520 ±

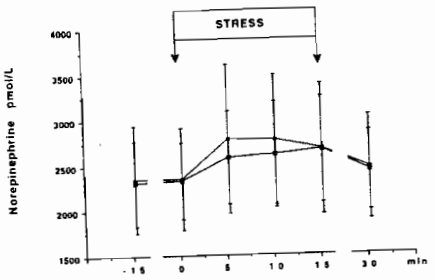


Figure 2. Mean (± SD) plasma norepinephrine levels before, during, and after mental stress after 3-week treatment with placebo (open square) or estradiol (filled circle).

136 μM (P <0.05)] and 17-beta-estradiol (peak values of 617 ± 123 μM vs basal values of 489 ± 108 μM (P <0.05); Figure 3]. There were no differences between treatments in either the absolute increments (P = 0.67) or the area under the curve responses for free fatty acids levels (P = 0.70).

Evaluations of Blood Pressure Levels and Continuous ECG Monitoring

Basal systolic and diastolic blood pressures were similar during the two treatments (Figure 4). In response to mental stress, systolic blood pressure increased significantly during placebo treatment, from 120 ± 9 mm Hg to a peak of 136 ± 15 mm Hg (P <0.001). During estradiol treatment, systolic blood pressure also increased significantly in response to the stress challenge, from 119 ± 9 mm Hg to a peak of 133 ± 15 mm Hg (P <0.01). There were no effects of treatment on the stress-induced systolic blood pressure changes. However, estradiol treatment

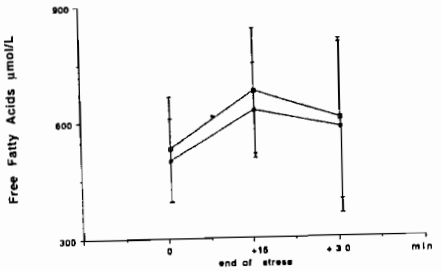


Figure 3. Mean (± SD) serum free fatty acid levels evaluated at baseline, at the end of mental stress, and at the end of the recovery period after 3-week treatment with placebo (open square) or estradiol (filled circle).

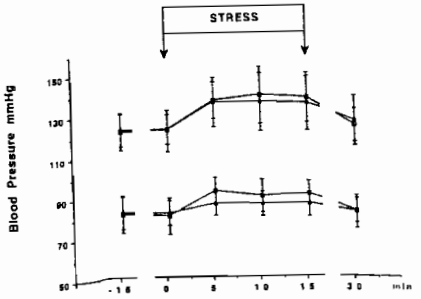


Figure 4. Mean (± SD) systolic blood pressure (above) and diastolic blood pressure (below) before, during, and after mental stress after 3-week treatment with placebo (open square) or estradiol (filled circle). Repeated measure analysis of variance revealed a significant (P <0.05) difference between treatments in the diastolic blood pressure response.

blunted the effects of stress on diastolic blood pressure (P <0.05). Mean diastolic blood pressure was significantly increased from baseline during placebo treatment (peak levels of 90 ± 6 mm Hg vs basal levels of 78 ± 15 mm Hg, P <0.002), but not during estrogen treatment (peak levels of 84 ± 6 mm Hg vs basal levels of 80 ± 8 mm Hg, P = 0.64).

Subjects had similar mean RR intervals, standard deviation of all mean RR intervals, and low-frequency and high-frequency bandwidths during the two treatments (Table). Mean RR intervals decreased significantly during the stress challenge during both treatments; the decrement was somewhat less pronounced during estradiol treatment, but the difference between treatments was not statistically significant (P = 0.51). In response to stress, there was a significant decrease in the standard deviation of all mean RR intervals values (P <0.01) during placebo treatment, but not during estrogen treatment (P = 0.23, P <0.05 for the comparison of placebo and estradiol treatments). We observed a stress-related increase in low-frequency bandwidths during both treatments; there were no effects of treatment on stress-induced high-frequency bandwidths. The ratio of low-frequency to high-frequency bandwidths increased after the stress challenge during both treatments but were significantly lower during estradiol treatment (P <0.05).

CONCLUSION

We found that transdermal administration of 17-beta-estradiol reduced the effects of mental arithmetic stress on plasma epinephrine levels and diastolic, but not systolic, blood pressure in postmenopausal women. These

Table. Indexes of Heart Rate and Heart Rate Variability before, during, and after Mental Stress after 3-Week Treatment with Either Placebo or Estradiol in 10 Postmenopausal Women

Measurement	Placebo			17-beta-Estradiol		
	Baseline	Stress	Recovery	Baseline	Stress	Recovery
	Mean ± SD					
Mean of all RR intervals (msec)	828 ± 39	776 ± 42	816 ± 56	824 ± 35	777 ± 37	831 ± 51
Standard deviation of all mean RR intervals (msec)	59 ± 11	47 ± 7	50 ± 9	54 ± 15	47 ± 12	52 ± 23
Low-frequency bandwidth ($msec^2$)	494 ± 172	1,562 ± 1,046	699 ± 269	543 ± 154	756 ± 152*	551 ± 250
High-frequency bandwidth ($msec^2$)	327 ± 133	309 ± 200	310 ± 120	320 ± 157	238 ± 121	276 ± 188
Low-frequency/high-frequency ratio	1.9 ± 0.6	5.9 ± 2.3	2.7 ± 1.4	2.1 ± 1	3.7 ± 1.4*	2.4 ± 0.9

*P <0.05 versus placebo, by analysis of variance.

effects were accompanied by a decrease in the responses of some measures of stress-induced cardiac sympathetic tone. The blunting effects of estrogen administration are unlikely to be the result of a learning effect on mental stress, given that we used a double-blind, crossover design. Moreover, the Spielberg State Anxiety scores and the degree of difficulty of the mental stress were similar during the two testing sessions. Thus, it seems unlikely that the reduced sympathoadrenal response to stress that we observed during the estradiol treatment was primarily the result of the stress being perceived as less severe.

Although estradiol had a moderate effect in reducing stress-induced plasma epinephrine levels, it did not have a significant effect on norepinephrine levels. However, norepinephrine levels were not significantly increased during mental stress with either treatment. These observations agree with previous findings that there is less activation of the sympathetic nervous system (ie, norepinephrine levels) than of the adrenal medulla (ie, epinephrine levels) during mental stress (30,31). In our study, subjects were tested at the end of 3 weeks of 17-beta-estradiol administration. Estrogens rapidly and directly inhibit catecholamine secretion from the adrenal medulla, likely through a nongenomic mechanism (32). Thus, we believe that 3 weeks of estradiol administration was sufficient to detect any effects of treatment on stress-induced adrenomedullary activity.

We did not find any significant effects of estrogen administration on stress-induced heart rate, as measured by mean RR intervals. However, when the heart rate variability was evaluated in the time domain, estradiol blunted the stress-related decrease in the standard deviation of all mean RR intervals. Similarly, estradiol reduced the response of the low-frequency/high-frequency ratio to stress, primarily because of a decrease in the values of the low-frequency bandwidth. The changes are directly related to the cardiac sympathetic tone, both in basal conditions and during psychologic stress (33,34). Our results, therefore, suggest that a decrease in the stress-induced cardiac sympathetic activity occurs during 17-beta-estradiol supplementation in postmenopausal women.

Recently, Komesaroff et al (35) reported a reduction in the blood pressure and catecholamine responses to mental stress during oral estrogen administration in perimenopausal women. We could only partially confirm those results, as we found that estradiol had a less potent blunting effect on the overall sympathoadrenal response to stress. Therefore, we hypothesize that a reduction in the sensitivity of the sympathoadrenal system to estrogen occurs during the transition from perimenopause to overt postmenopausal status. We studied postmenopausal women who had no or minor vasomotor symptoms, to avoid the potential bias of an estrogen-related relief in those symptoms. This may have led to a less severe perception of the stress stimulus during estrogen treatment. However, in the study by Komesaroff et al, estrogen administration resulted in circulating estradiol levels greater than those that occur during the standard transdermal estrogen administration, as in our study. Therefore, the more complete blunting effect of estrogen on mental stress–induced sympathoadrenal activity in the previous study could have been the result of greater circulating estrogen levels.

In agreement with a previous report (36), we did not find any stress-related changes in either blood glucose or insulin levels, but we did observe similar, statistically significant mental stress–induced increases in serum levels of free fatty acids during both placebo and estradiol treatment. Mental stress causes an increase in catecholamine-mediated lipolytic activity through an activation of beta-adrenoreceptors (37), and estradiol decreases the stress-induced activation of lipolysis in men (4). Data from experimental animals show that estrogens blunt the beta-

adrenoreceptor response to catecholamine stimulation (10). Thus, postmenopausal estrogen supplementation might be expected to reduce the effects of mental stress–induced sympathoadrenal activation on lipolysis; however, our results did not confirm this hypothesis.

In conclusion, we found that transdermal 17-beta-estradiol administration, at a dose of 50 μg per day, moderately, but significantly, reduces the response to mental stress, as measured by plasma epinephrine levels, diastolic blood pressure, and the overall cardiac sympathetic tone, in postmenopausal women. Studies are needed to determine the influence of estrogens, either alone or in combination with progestins, on sympathoadrenal function in basal conditions, during daily activities, and after challenges with several types of psychologic tests.

REFERENCES

1. Davidson L, Vandogen R, Rouse IL, et al. Sex-related differences in resting, and stimulated plasma noradrenaline, and adrenaline. *Clin Sci.* 1984;67:347–352.
2. Gustafson AB, Kalkhoff RK. Influence of sex and obesity on plasma catecholamine response to isometric exercise. *J Clin Endocrinol Metab.* 1982;55:703–708.
3. Amiel SA, Maran A, Powrie JK, et al. Gender differences in counterregulation to hypoglycemia. *Diabetologia.* 1993;36:460–464.
4. Del Rio G, Velardo A, Zizzo G, et al. Effect of estradiol on the sympathoadrenal response to mental stress in normal men. *J Clin Endocrinol Metab.* 1994;79:836–840.
5. Lopez MG, Abad F, Sancho C, et al. Membrane-mediated effects of the steroid 17 alpha-estradiol on adrenal catecholamine release. *J Pharmacol Exp Ther.* 1991;259:279–285.
6. Lloyd T, Weisz J. Direct inhibition of tyrosine hidroxylase activity by catechol estrogen. *J Biol Chem.* 1987;253:4841–4843.
7. Ball P, Knuppen R. Catecholoestrogens. *Acta Endocrinol (Copenh).* 1980;93:91–99.
8. Peters JR, Martin-Elliott J, Grahame-Smith DG. Effect of oral contraceptives on platelet noradrenaline and 5-hydroxytryptamine receptors and aggregation. *Lancet.* 1979;2:933–936.
9. Larsson B, Andersson KE, Batra S, et al. Effects of estradiol on norepinephrine-induced contraction, alpha adrenoceptor number and norepinephrine content in the female rabbit urethra. *J Pharmacol Exp Ther.* 1984;229:557–562.
10. Ball P, Knuppen R, Bruer H. Interactions between estrogens and catecholamines, III: studies on the methylation of catechol estrogens, catecholamines and other catechols by the catechol-o-methyltransferase of human liver. *J Clin Endocrinol Metab.* 1972;34:736–746.
11. Kannel W, Hjortland MC, McNamara PM, Gordon T. Menopause and risk of cardiovascular disease: the Framingham Study. *Ann Intern Med.* 1976;85:447–452.
12. Owens JF, Stoney CM, Matthews KA. Menopausal status influences ambulatory blood pressure levels and blood pressure changes during mental stress. *Circulation.* 1993;88:2794–2802.
13. Stampfer MJ, Colditz GA, Willett WC, et al. Postmenopausal estrogen therapy and cardiovascular disease: ten-year follow-up from the Nurses' Health Study. *N Engl J Med.* 1991;325:756–762.
14. Barrett-Connor E, Bush TL. Estrogen and coronary heart disease in women. *JAMA.* 1991;265:1861–1867.
15. Hulley S, Grady D, Bush T, et al. Randomized trial of estrogen plus progestin for secondary prevention of coronary heart disease in postmenopausal women. Heart and Estrogen/progestin Replacement Study (HERS) Research Group. *JAMA.* 1998;280:605–613.
16. McGowan JA, Potten L. Commentary on the Women's Health Initiative. *Maturitas.* 2000;34:109–112.
17. Bush TL, Barrett-Connor E, Cowan LD, et al. Cardiovascular mortality and non-contraceptive use of estrogen in women: results from the Lipid Research Clinics Program Follow-up Study. *Circulation.* 1987;75:1102–1109.
18. Gilligan DM, Quyyumi AA, Cannon RO. Effects of physiological levels of estrogen on coronary vasomotor function in postmenopausal women. *Circulation.* 1994;89:2545–2551.
19. Rosano GMC, Patrizi R, Leonardo F, et al. Effect of estrogen replacement therapy on heart rate variability and heart rate in healthy postmenopausal women. *Am J Cardiol.* 1997;80:815–817.
20. Del Rio G, Velardo A, Menozzi R, et al. Acute estradiol and progesterone administration reduced cardiovascular and catecholamine responses to mental stress in menopausal women. *Neuroendocrinology.* 1998;67:269–274.
21. Lee M, Giardina EGV, Homma S, et al. Lack of effect of estrogen on rest and treadmill exercise in postmenopausal women without known cardiac disease. *Am J Cardiol.* 1997;80:793–797.
22. Snabes MC, Herd A, Schuyler N, et al. In normal postmenopausal women physiologic estrogen replacement therapy fails to improve exercise tolerance: a randomized, double-blind, placebo-controlled, crossover trial. *Am J Obstet Gynecol.* 1996;175:110–114.
23. Kellner R. *Abridged Manual of the Symptom Questionnaire.* Albuquerque: University of New Mexico Press, 1976.
24. Paykell ES, McGuinness B, Gomez J. An Anglo-American comparison of the scaling of life events. *Br J Med Psychol.* 1976;49:237–247.
25. Spielberger CD, Gorsuch RL, Lushene RE. *Manual for the State-Trait Anxiety Inventory.* Palo Alto: Consulting Psychologists Press, 1970.
26. Wechsler D. *The Manual for the Wechsler Adult Intelligence Scale.* New York: Psychology Corporation, 1955.
27. Stein PK, Bosner M, Kleiger RE, Conger BM. Heart rate variability: a measure of cardiac autonomic tone. *Am Heart J.* 1994;127:1376–1381.
28. Maliani A, Lombardi F, Pagani M. Power spectrum analysis of heart rate variability: a tool to explore neural regulatory mechanism. *Br Heart J.* 1994;71:1–2.
29. Norusis MJ. *SPSS-X Advanced Statistic Guide.* New York. McGraw-Hill, 1985.
30. Hjemdahl P, Freyschuss U, Juhlin-Dannfelt A, Linde B. Plasma catecholamines and mental stress. In: Juel Christensen N, Henriksen O, Lassen NA, eds. *The Sympathoadrenal System.* Copenhaen: Raven, 1986:237–248.
31. Barnes FA, Raskind M, Gumbrecht G, Halter JB. The effects of age on the plasma catecholamine response to mental stress in man. *J Clin Endocrinol Metab.* 1982;54:64–69.
32. Schumacher M. Rapid membrane effects on steroid hormones: an emerging concept in neuroendocrinology. *Trends Neuroendocrinol Sci.* 1990;13:359–362.
33. Malliani A, Pagani M, Lombardi F, Cerruti S. Cardiovascular neural regulation explored in the frequency domain. *Circulation.* 1991;84:482–492.
34. Sgoifo A, De Boer SF, Buwalda B, et al. Vulnerability to arrhythmias during social stress in rats with different sympathovagal balance. *Am J Physiol.* 1998;275:H460–H466.
35. Komesaroff PA, Esler MD, Sudhir K. Estrogen supplementation attenuates glucocorticoid and catecholamine responses to mental stress in perimenopausal women. *J Clin Endocrinol Metab.* 1999;84:606–610.
36. Macdonald IA, Bennett T, Fellows IW. Catecholamines and the control of metabolism in man. *Clin Sci.* 1985;68:613–619.
37. Hagström-Toft E, Arner P, Wahrengerg H, et al. Adrenergic regulation of human adipose tissue metabolism in situ during mental stress. *J Clin Endocrinol Metab.* 1993;76:392–398.

An Intensive Communication Intervention for the Critically Ill

Craig M. Lilly, MD, Dawn L. De Meo, MD, Larry A. Sonna, MD, PhD, Kathleen J. Haley, MD, Anthony F. Massaro, MD, Robert F. Wallace, MPH, Sean Cody, RN, MSN, MBA

PURPOSE: We sought to determine the effects of a communication process that was designed to encourage the use of advanced supportive technology when it is of benefit, but to limit its burdens when it is ineffective. We compared usual care with a proactive, multidisciplinary method of communicating that prospectively identified for patients and families the criteria that would determine whether a care plan was effective at meeting the goals of the patient. This process allowed caregivers to be informed of patient preferences about continued advanced supportive technology when its continuation would result in a compromised functional outcome or death.

MATERIALS AND METHODS: We performed a before-and-after study in 530 adult medical patients who were consecutively admitted to a university tertiary care hospital for intensive care. Multidisciplinary meetings were held within 72 hours of critical care admission. Patients, families, and the critical care team discussed the care plan and the patients' goals and expectations for the outcome of critical care. Clinical "milestones" indicative of recovery were identified with time frames for their occurrence. Follow-up meetings were held to discuss palliative care options when continued advanced supportive technology was not achieving the patient's goals. We measured length of stay, mortality, and provider team and family consensus in 134 patients before the intensive communication intervention and in 396 patients after the intervention.

RESULTS: Intensive communication significantly reduced the median length of stay from 4 days (interquartile range, 2 to 11 days) to 3 days (2 to 6 days, $P = 0.01$ by survival analysis). This reduction remained significant after adjustment for acute physiology and chronic health evaluation (APACHE) 3 score [risk ratio (RR) = 0.81; 95% confidence interval (CI), 0.66 to 0.99; $P = 0.04$). Subgroup analysis revealed that this reduction occurred in our target group, patients with acuity scores in the highest quartile who died (RR = 0.60; 95% CI, 0.38 to 0.92; $P = 0.02$). The intervention, which allowed dying patients earlier access to palliative care, was not associated with increased mortality.

CONCLUSIONS: Intensive communication was associated with a reduction in critical care use by patients who died. Our multidisciplinary process targeted advanced supportive technology to patients who survived and allowed the earlier withdrawal of advanced supportive technology when it was ineffective. **Am J Med. 2000;109:469–475.**

As much as 1% of the gross national product in the United States is used to provide care for critically ill adults (1). Access to critical care is valued not only by survivors but by the families of most of those who do not survive (2). The wider respect for patient autonomy and the advent of advance directives have allowed some dying patients to receive palliative, rather than aggressive, critical care (3). Despite the alternatives, many patients choose aggressive care, and 2% to 40% do not survive (4). Access to palliative care is often delayed for dying, critically ill patients (5,6), because it is difficult for providers to articulate, and for patients and families to accept, that advanced supportive technology has been ineffective or will not result in a functional outcome that is acceptable to the patient. The optimal timing of the transition to palliative care affects patient comfort and critical care use.

Dying patients consume a disproportionate amount of intensive care resources. In one study, high-use patients who died comprised 13% of intensive care unit (ICU) patients, yet consumed 32% of the resources (7). In our medical ICU in 1996, more than 50% of patient-days were spent providing advanced supportive technology for patients who did not survive. More important, because the support of critically ill patients is accompanied by substantial discomfort and deprivation, it is not humane to continue burdensome technology when there is no reasonable hope that it will be effective. The intravascular lines, support tubes, and restraints used in this process are inherently uncomfortable and are associated with reduced mobility, autonomy, and ability to communicate. When advanced supportive care is associated with healing and the return of function, these discomforts are almost universally acceptable, but they are difficult to justify for patients who die. For this reason, as many as 75% of hospitals now work with patients and their families to withdraw advanced supportive technology when it is not thought to be of benefit (4,8–12).

Like others (13,14), we believe that improved commu-

From the Combined Program in Pulmonary and Critical Care Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts; and US Army Research Institute of Environmental Medicine, Natick, Massachusetts.

Supported by Brigham and Women's Hospital.

Requests for reprints should be addressed to Craig M. Lilly, MD, Respiratory Division, Brigham and Women's Hospital, 75 Francis Street, Boston, Massachusetts 02115.

Manuscript submitted November 22, 1999, and accepted in revised form June 5, 2000.

0002-9343/00/$–see front matter
PII S0002-9343(00)00524-6

(235) [268]

Mediterranean and Low-Fat Diets Improve Endothelial Function in Hypercholesterolemic Men

Francisco Fuentes, MD, PhD; José López-Miranda, MD, PhD; Elías Sánchez, MD, PhD; Francisco Sánchez, MD, PhD; José Paez, MD; Elier Paz-Rojas, BSc; Carmen Marín, PhD; Purificación Gómez, BSc; José Jimenez-Perepérez, MD, PhD; José M. Ordovás, MD, PhD; and Francisco Pérez-Jiménez, MD, PhD

Background: The regulatory function of the endothelium is altered in hypercholesterolemia, and the subsequent endothelial dysfunction plays a central role in the development of atherosclerosis.

Objective: To determine whether endothelial function in hypercholesterolemic patients is affected by replacing a saturated fat–enriched diet with a low-fat, low–saturated fat diet (the U.S. National Cholesterol Education Program stage 1 [NCEP-1] diet) or a diet rich in monounsaturated fat (such as that common in Mediterranean countries).

Design: Intervention dietary study with a baseline phase and two randomized crossover dietary periods.

Setting: Hospital Universitario Reina Sofía, Córdoba, Spain.

Patients: 22 hypercholesterolemic men.

Intervention: Patients followed a diet high in saturated fat, then were assigned in a crossover design to the NCEP-1 diet or a Mediterranean diet. Each dietary period lasted 28 days.

Measurements: Plasma P-selectin levels, lipid concentrations, and endothelial function.

Results: Compared with the saturated fat diet, flow-mediated dilatation increased during the Mediterranean diet but not during the NCEP-1 diet. In addition, levels of plasma cholesterol, low-density lipoprotein cholesterol, apolipoprotein B, and P-selectin decreased during the NCEP-1 and Mediterranean diets.

Conclusion: In hypercholesterolemic men, diets low in fat (especially saturated fat) and diets rich in monounsaturated fats improve endothelial function.

Ann Intern Med. 2001;134:1115-1119. www.annals.org

For author affiliations, current addresses, and contributions, see end of text.

The endothelium helps regulate vascular tone through the release of vasodilator and vasoconstrictor substances (1), but this regulatory function is altered in hypercholesterolemia (2). Endothelial dysfunction plays a central role in the development of atherosclerosis (3).

Diet is the cornerstone of hypercholesterolemia treatment. Studies in animals have shown that dietary therapies restore endothelium-dependent vasodilatation to normal (4). However, the information available in humans is scarce and was obtained mainly in studies of diets rich in polyunsaturated n-3 fatty acids (5). The U.S. National Cholesterol Education Program (NCEP) recommends a low-fat, low–saturated fat diet for the primary prevention of atherosclerosis (6); this diet is referred to as the stage 1 diet (NCEP-1). In addition, a diet high in monounsaturated fat, which is a common diet in Mediterranean countries, has been related to decreased risk for coronary heart disease (7), and the Lyon Diet Heart Study found that a Mediterranean diet rich in linolenic acid reduced cardiovascular events (8). At present, the effect of these diets on endothelial function in humans is unknown. Our aim was to investigate endothelial function after substituting each of these potentially cardioprotective diets for a saturated fat–enriched diet in hypercholesterolemic patients.

Methods

Patients

Twenty-two white hypercholesterolemic men attending the Lipid Clinic at the Reina Sofía University Hospital in Córdoba, Spain, volunteered to participate in the study. All patients were 18 to 65 years of age (mean [± SD], 40.5 ± 14.7 years) and had plasma cholesterol concentrations greater than 5.2 mmol/L (200 mg/dL). None had chronic illness, were extremely physically active, or had any known secondary cause of hypercholesterolemia. Patients were excluded if they had a family history of premature coronary heart disease or if they had taken vitamin supplements or medication known to affect plasma lipid levels in the 6 months before the study. Patients were encouraged to maintain their regular lifestyles and regular levels of physical activity.

Intervention

The study design included an initial 28-day period during which all patients consumed a diet high in saturated fat, with 15% protein, 47% carbohydrates, and 38% fat (20% saturated fat, 12% monounsaturated fat, and 6% polyunsaturated fat) constituting energy intake. Following this period, two groups of 10 and 12 patients each were randomly assigned in a crossover design to two diets for two 28-day periods: the NCEP-1 diet and a diet high in monounsaturated fat (a typical Mediterranean diet enriched with olive oil). Group 1 was placed on the Mediterranean diet followed by the NCEP-1 diet, while the order of the diets for patients in group 2 was reversed. The NCEP-1 diet consisted of 57% carbohydrates and 28% fat (<10% saturated fat, 12% monounsaturated fat, and 6% polyunsaturated fat). The Mediterranean diet consisted of 47% carbohydrates and 38% fat (<10% saturated fat, 22% monounsaturated fat, and 6% polyunsaturated fat). Olive oil provided 75% of the total monounsaturated fat consumed during the Mediterranean diet. Dietary cholesterol was maintained at a constant level (3 mmol/4800 kJ) during the three periods. We did not include washout periods between the three stages of dietary intervention because washout is unnecessary when dietary intervention lasts longer than 3 weeks (9). The Human Investigation Review Committee at Reina Sofía University Hospital approved the study, and informed consent was obtained from all patients.

The experimental diets followed the U.S. Department of Agriculture food tables and Spanish food composition tables. Twenty menus were prepared with common food items and were rotated during the experimental period. Dietary adherence was determined by measuring fatty acid enrichment of low-density lipoprotein (LDL) cholesterol esters at the end of each diet, since changes in dietary fatty acids induce a parallel change in the composition of LDL cholesterol esters.

Blood Sampling and Biochemical Determinations

At the end of each diet, after a 12-hour overnight fast, venous blood samples were collected in tubes containing EDTA (1 g/L). Cholesterol and triglycerides were assayed by using enzymatic procedures. High-density lipoprotein cholesterol level was measured after precipitation of apolipoprotein B–containing lipoproteins with phosphowolframic acid. Low-density lipoprotein cholesterol level was calculated by using the Friedewald formula. Apolipoprorein A-I and apolipoprotein B concentrations were determined by using turbidimetry. Plasma levels of cell adhesion molecules (vascular cell adhesion molecule-1 [VCAM-1] and intercellular cell adhesion molecule-1 [ICAM-1]) and P-selectin were determined by using enzyme-linked immunosorbent assay.

Endothelial Function

A noninvasive method described by Celermajer and colleagues (10) was used to evaluate endothelial function. This technique uses postischemic vasodilatation, causing enhanced flow in the proximal brachial artery and shear stress–induced vasodilatation considered endothelium dependent. The diameter of the right brachial artery was measured at the end of each dietary period by using two-dimensional ultrasonography images, with a 7-MHz linear array transducer and a standard 128XP/10 system (Acuson Corp., Mountain View, California). Scans were taken at rest, during reactive hyperemia (endothelium-dependent vasodilatation), again at rest (10 minutes), and after 400 μg of sublingual glyceryl trinitrate. Two independent observers who were unaware of clinical details and the stage of the experiment measured vessel diameter. The resistance index was also measured by each observer.

Statistical Analysis

The sample size calculations indicated that a minimum of 15 patients was required per group (based on a minimal expected increase in the brachial artery diameter [mean ± SD] of 0.08 mm ± 0.05 mm, an α value of 0.025, a power of 0.90, and a 10% estimated loss to follow-up). For each diet, we used analysis of variance for repeated measures to test for dietary effects on plasma lipid levels and endothelial function. When significant effects were detected ($P < 0.05$), the Bonferroni test was used for a post hoc comparison. For example, to determine whether consumption of the three diets modified levels of LDL cholesterol, we initially used the analysis of variance test, which indicated that the diets differ ($P < 0.001$). Subsequently, we performed a post hoc test using the Hochberg modification of the Bonferroni procedure (11) to establish in which pair of diets differences were produced, taking into consideration that each independent variable had three possible paired

Table 1. **Flow-Associated Vasodilatation (Endothelium-Dependent); Glyceryl Trinitrate–Induced Vasodilatation (Endothelium-Independent); and Plasma Levels of Soluble Intercellular Cell Adhesion Molecule-1, Vascular Cell Adhesion Molecule-1, and P-Selectin at the End of Each Dietary Period***

Variable	Saturated Fat Diet	Mediterranean Diet	NCEP-1 Diet	*P* Value†
Baseline vessel size, *mm*	4.14 ± 0.12	4.06 ± 0.09	4.04 ± 0.11	0.182
Flow-associated vasodilatation, *mm*	0.41 ± 0.05	0.54 ± 0.04‡	0.45 ± 0.03	0.04
Glyceryl trinitrate–induced vasodilatation, *mm*	0.58 ± 0.06	0.48 ± 0.05	0.49 ± 0.04	>0.2
Flow-associated vasodilatation, %	9.9 ± 0.8	13.5 ± 0.9‡	11.1 ± 1.1	0.05
Glyceryl trinitrate–induced vasodilatation, %	13.1 ± 1.6	11.3 ± 1.3	11.5 ± 1.1	>0.2
Flow-associated vasodilatation resistance index	0.944 ± 0.021	0.880 ± 0.017‡	0.907 ± 0.023‡	0.002
Glyceryl trinitrate–induced vasodilatation resistance index	0.992 ± 0.025	0.921 ± 0.027‡	0.916 ± 0.022‡	0.001
VCAM-1 level, *ng/mL*	402 ± 22	400 ± 20	395 ± 20	>0.2
ICAM-1 level, *ng/mL*	52.1 ± 4.1	50.5 ± 3.8	50.5 ± 3.9	>0.2
P-Selectin level, *ng/mL*	76.1 ± 6.2	57.4 ± 4.1‡	63.1 ± 5.6	0.001

* Values are expressed as the mean ± SD. ICAM-1 = intercellular cell adhesion molecule-1; NCEP-1 = National Cholesterol Education Program, stage 1; VCAM-1 = vascular cell adhesion molecule-1.
† Repeated-measures analysis of variance.
‡ $P < 0.05$ compared with the saturated fat diet.

comparisons. Triglyceride values were logarithmically transformed to achieve an approximately normal distribution, and statistical tests were performed on the transformed values.

Role of the Funding Sources

The funding sources had no role in the collection, analysis, or interpretation of the data or in the decision to submit the paper for publication.

Results

Adherence to the diets was excellent, as suggested by the absence of significant events recorded in the participants' diaries and the composition of plasma cholesteryl esters found in LDL cholesterol. Compared with the NCEP-1 diet, palmitic acid increased significantly during the saturated fat diet ($P = 0.01$) and oleic acid levels increased significantly during the Mediterranean diet ($P = 0.02$).

At the end of each dietary period, measurement of endothelial function showed no significant differences in the basal diameters of the brachial artery (Table 1) or in the glyceryl trinitrate–induced vasodilatation. Flow-associated vasodilatation of the brachial artery was significantly higher after the Mediterranean diet ($P = 0.027$) than after the saturated fat diet (13.5% vs. 9.9%). However, we did not observe significant differences in flow-associated vasodilatation between the saturated fat diet and the NCEP-1 diet (9.9% vs. 11.1%). The resistance index was higher after flow-associated vasodilatation (the endothelium-dependent value) ($P = 0.003$) and after glyceryl trinitrate–induced vasodilatation ($P = 0.002$) during the saturated fat diet than during the Mediterranean diet or the NCEP-1 diet ($P = 0.008$). P-Selectin levels were lower during the Mediterranean diet and the NCEP-1 diet ($P = 0.003$ and $P = 0.068$, respectively) than during the saturated fat diet. We did not find significant differences in plasma concentrations of VCAM-1 and ICAM-1 among the three diets.

Compared with the saturated fat diet (Table 2), the NCEP-1 and Mediterranean diets were associated with a decrease in plasma cholesterol level ($P = 0.001$), LDL cholesterol level ($P < 0.001$), and apolipoprotein B level ($P = 0.002$) but did not significantly change high-density lipoprotein cholesterol, apolipoprotein A-I, or triglyceride levels. In addition, we found a negative correlation between LDL cholesterol levels and flow-associated vasodilatation ($r = -0.23$; $P = 0.036$). Changes in LDL cholesterol levels were positively correlated with P-selectin levels ($r = 0.275$; $P = 0.030$), while flow-associated vasodilatation was negatively correlated with P-selectin levels ($r = -0.372$; $P = 0.018$).

Discussion

In our study, replacement of a saturated fat–enriched diet by a monounsaturated fat–enriched diet and an NCEP-1 diet increased endothelium-dependent vasodilatation. The increase in flow-associated vasodilatation observed with the consumption of a diet high in monounsaturated fat suggests that this diet improves endothelial function. Our patients had primary hypercholesterolemia. In hypercholesterolemic persons, the response of endothelium-dependent vasodilatation is altered, even in the absence of atherosclerotic lesions (12). Such alteration is reversible after treatment with lipid-lowering drugs (13). Our findings support this hypothesis, since we observed a correlation between a decrease in LDL cholesterol levels produced by dietary intervention and flow-associated vasodilatation. The mechanisms responsible for endothelial dysfunction in hypercholesterolemia are not completely understood but may be explained by decreased bioavailability of nitric oxide because of decreased production by endothelial cells or increased degradation by oxygen-derived free radicals (14). Our study was limited, however, because we intentionally included only men to exclude the effect of hormonal status in women.

In recent years, it has been demonstrated that dietary fat may affect the endothelium (5, 6) and factors related to the arterial wall, such as type 1 plasminogen activator inhibitor and von Willebrand factor (15). It has also been shown that the incubation of human endothelial cells with oleate and docosahexanoic acid (16) reduces the expression of adhesion molecules and monocyte adhesion (17). In addition, the antioxidant-rich foods of a Mediterranean diet benefit postprandial endothelial function after a fat-rich meal (18). Our data also support the notion that diet influences endothelial cells, improving vasomotor function and decreasing plasma levels of P-selectin by decreasing levels of LDL cholesterol (19). Previous studies have shown that endothelial dysfunction precedes atherosclerosis and is probably important in its pathogenesis (20). If atherosclerosis is part of a continuous spectrum from normal endothelial function to overt atherosclerosis, it may be possible to correct endothelial dysfunction, a useful early functional expression of the disease process, by prescribing a lipid-lowering diet, as we observed. The beneficial effects of dietary intervention are not limited solely to its action on plasma lipid levels but may also influence other pathogenic mechanisms, opening up new perspectives for its protective effect on atherosclerosis.

tradaments

Table 2. **Plasma Lipid and Apolipoprotein Concentrations at the End of Each Dietary Period***

Variable	Saturated Fat Diet	Mediterranean Diet	NCEP-1 Diet	*P* Value†
Total cholesterol level, *mmol/L (mg/dL)*	6.5 ± 1.5 (251 ± 58)	6.0 ± 1.3 (232 ± 50)	5.9 ± 1.3 (228 ± 50)	0.001
Triglyceride level, *mmol/L (mg/dL)*	1.9 ± 0.8 (168 ± 71)	2.0 ± 0.5 (177 ± 44)	2.1 ± 0.4 (186 ± 35)	>0.2
LDL cholesterol level, *mmol/L (mg/dL)*	4.4 ± 1.4 (170 ± 54)	3.8 ± 1.2 (147 ± 46)	3.7 ± 1.2 (143 ± 46)	<0.001
HDL cholesterol level, *mmol/L (mg/dL)*	1.2 ± 0.2 (46 ± 7.7)	1.2 ± 0.2 (45 ± 7.2)	1.2 ± 0.2 (44 ± 7.3)	>0.2
Apolipoprotein A-I level, *g/L*	1.47 ± 0.17	1.42 ± 0.17	1.40 ± 0.17	0.062
Apolipoprotein B level, *g/L*	1.21 ± 0.24	1.16 ± 0.24	1.14 ± 0.20	0.002
LDL–HDL ratio	3.52 ± 1.32	3.24 ± 1.18	3.08 ± 1.02	<0.001

* Values are expressed as the mean ± SD. HDL = high-density lipoprotein; LDL = low-density lipoprotein; NCEP-1 = National Cholesterol Education Program, stage 1.
† Repeated-measures analysis of variance.

From Hospital Universitario Reina Sofía, Córdoba, Spain, and Jean Mayer–U.S. Department of Agriculture Human Nutrition Research Center on Aging at Tufts University, Boston, Massachusetts.

Acknowledgments: The authors thank Beatriz Pérez for help in preparing the manuscript.

Grant Support: By research grants from the Comisión Interministerial de Ciencia y Tecnología (OLI 96/2146) (Dr. Pérez-Jiménez), the Spanish Ministry of Health (FIS 96/1540, 98/1531 [Dr. López-Miranda], FIS 99/0949 [Dr. Pérez-Jiménez]), Fundación Cultural "Hospital Reina Sofía-Cajasur" (Dr. Marín and Ms. Gómez), Agencia Española de Cooperación Internacional (Mr. Paz-Rojas), Consejería de Salud, Servicio Andaluz de Salud (PAI 97/58, 98/126, 99/116), Consejería de Agricultura y Pesca de la Junta de Andalucía, Patrimonio Comunal Olivarero (Dr. Pérez-Jiménez), and the National Institutes of Health (HL 54776) (Dr. Ordovás).

Requests for Reprints: Francisco Pérez-Jiménez, MD, PhD, Unidad de Lípidos y Arteriosclerosis, Hospital Universitario Reina Sofía, Avda Menéndez Pidal, s/n 14004 Córdoba, Spain; e-mail, fperez@sofia.hrt.sas.cica.es.

Current Author Addresses: Drs. Fuentes, López-Miranda, Sánchez, Marín, Jimenez-Perepérez, and Pérez-Jiménez, Mr. Paz-Rojas, and Ms. Gómez: Unidad de Lípidos y Arteriosclerosis, Hospital Universitario Reina Sofía, Avenida Menéndez Pidal, s/n 14004 Córdoba, Spain.
Drs. Sánchez and Paez: Servicio de Radiodiagnóstico, Hospital Universitario Reina Sofía, Avenida Menendez Pedal, s/n 14004 Córdoba, Spain.
Dr. Ordovás: Lipid Metabolism Laboratory, Jean Mayer–U.S. Department of Agriculture Human Nutrition Research Center on Aging at Tufts University, 711 Washington Street, Boston, MA 02111.

95 ?

Author Contributions: Conception and design: F. Fuentes, J. López-Miranda, J.M. Ordovás, F. Pérez-Jiménez.
Analysis and interpretation of the data: F. Fuentes, J. López-Miranda, E. Sánchez, E. Paz-Rojas, C. Marín, P. Gómez, J.M. Ordovás, F. Pérez-Jiménez.
Drafting of the article: F. Fuentes, J. López-Miranda, J. Jimenez-Perepérez, J.M. Ordovás, F. Pérez-Jiménez.
Critical revision of the article for important intellectual content: J. López-Miranda, J. Jimenez-Perepérez, J.M. Ordovás, F. Pérez-Jiménez.
Final approval of the article: J. López-Miranda, E. Paz-Rojas, C. Marín, P. Gómez, J.M. Ordovás, F. Pérez-Jiménez.
Provision of study materials or patients: F. Fuentes, J. López-Miranda, E. Sánchez, F. Sánchez, J. Paez.
Statistical expertise: J. López-Miranda, E. Paz-Rojas, P. Gómez.
Obtaining of funding: J. López-Miranda, F. Pérez-Jiménez.
Administrative, technical, or logistic support: F. Sánchez, J. Paez, E. Paz-Rojas, C. Marín, P. Gómez.
Collection and assembly of data: F. Fuentes, E. Sánchez, F. Sánchez, J. Paez, E. Paz-Rojas, C. Marín, P. Gómez.

References

1. Vane JR, Anggård EE, Botting RM. Regulatory functions of the vascular endothelium. N Engl J Med. 1990;323:27-36. [PMID: 2113184]

2. Sorensen KE, Celermajer DS, Georgakopoulos D, Hatcher G, Betteridge DJ, Deanfield JE. Impairment of endothelium-dependent dilation is an early event in children with familial hypercholesterolemia and is related to the lipoprotein(a) level. J Clin Invest. 1994;93:50-5. [PMID: 8282821]

3. Celermajer DS. Endothelial dysfunction: does it matter? Is it reversible? J Am Coll Cardiol. 1997;30:325-33. [PMID: 9247501]

4. Harrison DG, Armstrong ML, Freiman PC, Heistad DD. Restoration of endothelium-dependent relaxation by dietary treatment of atherosclerosis. J Clin Invest. 1987;80:1808-11. [PMID: 3680531]

5. Goode GK, Garcia S, Heagerty AM. Dietary supplementation with marine fish oil improves in vitro small artery endothelial function in hypercholesterolemic patients: a double-blind placebo-controlled study. Circulation. 1997;96:2802-7. [PMID: 9386141]

6. National Cholesterol Education Program. Second Report of the Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel II). Circulation. 1994;89:1333-445. [PMID: 8124825]

7. Keys A, Menotti A, Karvonen MJ, Aravanis C, Blackburn H, Buzina R, et al. The diet and 15-year death rate in the seven countries study. Am J Epidemiol. 1986;124:903-15. [PMID: 3776973]

8. de Lorgeril M, Salen P, Martin JL, Monjaud I, Delaye J, Mamelle N. Mediterranean diet, traditional risk factors, and the rate of cardiovascular complications after myocardial infarction: final report of the Lyon Diet Heart Study. Circulation. 1999;99:779-85. [PMID: 9989963]

9. Bonanome A, Grundy SM. Effect of dietary stearic acid on plasma cholesterol and lipoprotein levels. N Engl J Med. 1988;318:1244-8. [PMID: 3362176]

10. Celermajer DS, Sorensen KE, Gooch VM, Spiegelhalter DJ, Miller OI, Sullivan ID, et al. Non-invasive detection of endothelial dysfunction in children and adults at risk of atherosclerosis. Lancet. 1992;340:1111-5. [PMID: 1359209]

11. Hochberg Y. A sharper Bonferroni procedure for multiple tests of significance. Biometrika. 1988;75:800-2.

12. Creager MA, Cooke JP, Mendelsohn ME, Gallagher SJ, Coleman SM, Loscalzo J, et al. Impaired vasodilation of forearm resistance vessels in hypercholesterolemic humans. J Clin Invest. 1990;86:228-34. [PMID: 2195060]

13. Stroes ES, Koomans HA, de Bruin TW, Rabelink TJ. Vascular function in the forearm of hypercholesterolaemic patients off and on lipid-lowering medication. Lancet. 1995;346:467-71. [PMID: 7637480]

14. Ohara Y, Peterson TE, Harrison DG. Hypercholesterolemia increases endothelial superoxide anion production. J Clin Invest. 1993;91:2546-51. [PMID: 8390482]

15. Pérez-Jiménez F, Castro P, López-Miranda J, Paz-Rojas E, Blanco A, López-Segura F, et al. Circulating levels of endothelial function are modulated by dietary monounsaturated fat. Atherosclerosis. 1999;145:351-8. [PMID: 10488963]

16. De Caterina R, Liao JK, Libby P. Fatty acid modulation of endothelial activation. Am J Clin Nutr. 2000;71(1 Suppl):213S-235. [PMID: 10617974]

17. Mata P, Alonso R, Lopez-Farre A, Ordovas JM, Labot C, Garces C, et al. Effect of dietary fat saturation on LDL oxidation and monocyte adhesion to human endothelial cells in vitro. Arterioscler Thromb Vasc Biol. 1996;16:1347-55. [PMID: 8911273]

18. Vogel RA, Corretti MC, Plotnick GD. The postprandial effect of components of the Mediterranean diet on endothelial function. J Am Coll Cardiol. 2000;36:1455-60. [PMID: 11079642]

19. Gauthier TW, Scalia R, Murohara T, Guo JP, Lefer AM. Nitric oxide protects against leukocyte-endothelium interactions in the early stages of hypercholesterolemia. Arterioscler Thromb Vasc Biol. 1995;15:1652-9. [PMID: 7583540]

20. Yasue H, Matsuyama K, Matsuyama K, Okumura K, Morikami Y, Ogawa H. Responses of angiographically normal human coronary arteries to intracoronary injection of acetylcholine by age and segment. Possible role of early coronary atherosclerosis. Circulation. 1990;81:482-90. [PMID: 2105173]

(967)[532]

Effect of temazepam on ventilatory response at moderate altitude

EDITOR—Dubowitz's study of the effect of temazepam on oxygen saturation at high altitude found that benzodiazepines do not have a depressant effect.[1] He explains the discrepancy between his findings and those of previous studies by the fact that other studies have investigated the effect of long acting benzodiazepines.[2] Dubowitz's probands were investigated after altitude acclimatisation while walking to Everest base camp, whereas climbers in Europe mainly engage in short periods of mountaineering. We therefore evaluated the effect of 10 mg temazepam on respiration in non-acclimatised Alpine climbers at moderate altitude.

We performed a randomised, double blind, placebo controlled, crossover trial in seven healthy men aged 21 to 27. Participants at 171 m altitude were randomised to receive either 10 mg temazepam or placebo. Three days later the men were given the same medication and taken by cable car to 3000 m. The procedure was repeated after one week, with the men crossed to the other arm of the study. Arterial blood samples were obtained from the ear lobe before and one hour after temazepam or placebo was taken.[4] Arterial oxygen partial pressure and carbon dioxide partial pressure were analysed on an IL Synthesis 25 blood gas analyser (Instrumentation Laboratory, Milan, Italy). Differences in blood gas concentrations before and after temazepam or placebo at each altitude were analysed by paired *t* tests.

The table shows the results of blood gas analysis before and after temazepam. At 171 m blood gas concentrations did not change significantly after temazepam. At 3000 m the arterial oxygen pressure decreased and carbon dioxide pressure increased significantly after temazepam. The mean decrease in arterial oxygen concentration between altitudes was 0.77 (95% confidence interval −8.02 to −3.69) kPa ($P<0.01$) and the mean increase in arterial carbon dioxide concentration was 0.3 (0.46 to 4.11) kPa ($P<0.05$). Placebo did not affect blood gas concentrations at either altitude.

Although we did not measure respiration directly, our data indicate that a low dose of a short term benzodiazepine can impair respiration at moderate altitude. These findings seem to contradict Dubowitz's conclusion. Treatment with temazepam at stable conditions after altitude acclimatisation may not impair respiration, but initial stages of acute respiratory adaptation to hypoxia at altitude are inhibited. Similar results were found after 50 g alcohol at moderate altitude.[5] Caution in the use of benzodiazepines to treat sleep disorders at altitude is therefore necessary, especially in the initial stages of altitude acclimatisation.

Georg Röggla *head of department*
interne@khneunkirchen.at

Berthold Moser *student*
Department of Internal Medicine, Municipal Hospital of Neunkirchen, Neunkirchen, Austria

Martin Röggla *lecturer*
Department of Emergency Medicine, University of Vienna, Austria

1 Dubowitz G. Effect of temazepam on oxygen saturation and sleep quality at high altitude: randomised placebo controlled crossover trial. *BMJ* 1998;316:587-9.
2 Röggla G, Röggla M, Wagner A, Seidler D, Podolsky A. Effect of low dose sedation with diazepam on ventilatory response at moderate altitude. *Wien Klin Wochenschr* 1994;106:649-51.
3 Hills M, Armitage P. The two-period cross-over clinical trial. *Br J Clin Pharmacol* 1979;8:7-20.
4 Spiro SG, Dowdeswell IR. Arterialised ear lobe blood sampling for blood gas tensions. *Br J Dis Chest* 1976;70:263-8.
5 Röggla G, Röggla H, Röggla M, Binder M, Laggner AN. Effect of alcohol on acute ventilatory adaptation to mild hypoxia at moderate altitude. *Ann Intern Med* 1995;122:925-927.

Guidelines are needed if drug testing of those arrested by the police becomes compulsory

EDITOR—The prime minister's announcement at the last Labour Party conference that the government proposes to introduce compulsory drug and DNA testing for people arrested for indictable offences before they have been convicted will raise some important ethical issues for healthcare professionals if it does eventually become law.[1] As pointed out in the *Economist*,[2] the upshot of compulsory DNA testing might be that every alleged shoplifter could be held down forcibly while a mouth swab is taken.[2] Collecting sweat for a drug test by wiping the forehead of a restrained and resisting subject with a swab would be no more dignified.

Arterial oxygen (Pao_2) and carbon dioxide ($Paco_2$) concentrations (kPa) of seven men before and one hour after 10 mg temazepam at 171 and 3000 m

	Pao_2				$Paco_2$			
	171 m		3000 m		171 m		3000 m	
Case No	Before	After	Before	After	Before	After	Before	After
1	12.2	12.2	9.3	8.6	4.5	4.3	4.3	4.4
2	11.2	11.6	8.6	8.4	4.7	4.9	4.4	4.7
3	12.1	12	8.9	8.2	4.4	4.5	4.4	4.7
4	11	11.4	9.1	4.9	4.9	4.8	4.0	4.3
5	10.9	11	8.5	8.1	4.7	4.4	4.1	4.4
6	12.5	12.2	9.1	8.1	4.5	4.8	4.4	4.7
7	12	12.2	9.4	8.4	4.4	4.5	4.0	4.8
Mean (SD)	11.7 (0.63)	11.8 (0.48)	9 (0.29)	8.3 (0.2)	33.4 (1.4)	34.6 (1.8)	4.2 (0.19)	4.5 (0.2)

(335)[384]

Efficacy of 3 Commonly Used Hearing Aid Circuits

A Crossover Trial

Vernon D. Larson, PhD
David W. Williams, MS
William G. Henderson, PhD
Lynn E. Luethke, PhD
Lucille B. Beck, PhD
Douglas Noffsinger, PhD
Richard H. Wilson, PhD
Robert A. Dobie, MD
George B. Haskell, PhD
Gene W. Bratt, PhD
Janet E. Shanks, PhD
Patricia Stelmachowicz, PhD
Gerald A. Studebaker, PhD
Allen E. Boysen, PhD
Amy Donahue, PhD
Rinaldo Canalis, MD
Stephen A. Fausti, PhD
Bruce Z. Rappaport, PhD
for the Participants of the NIDCD/VA Hearing Aid Clinical Trial Group

Context Numerous studies have demonstrated that hearing aids provide significant benefit for a wide range of sensorineural hearing loss, but no carefully controlled, multicenter clinical trials comparing hearing aid efficacy have been conducted.

Objective To compare the benefits provided to patients with sensorineural hearing loss by 3 commonly used hearing aid circuits.

Design Double-blind, 3-period, 3-treatment crossover trial conducted from May 1996 to February 1998.

Setting Eight audiology laboratories at Department of Veterans Affairs medical centers across the United States.

Patients A sample of 360 patients with bilateral sensorineural hearing loss (mean age, 67.2 years; 57% male; 78.6% white).

Intervention Patients were randomly assigned to 1 of 6 sequences of linear peak clipper (PC), compression limiter (CL), and wide dynamic range compressor (WDRC) hearing aid circuits. All patients wore each of the 3 hearing aids, which were installed in identical casements, for 3 months.

Main Outcome Measures Results of tests of speech recognition, sound quality, and subjective hearing aid benefit, administered at baseline and after each 3-month intervention with and without a hearing aid. At the end of the experiment, patients ranked the 3 hearing aid circuits.

Results Each circuit markedly improved speech recognition, with greater improvement observed for soft and conversationally loud speech (all 52-dB and 62-dB conditions, $P \leq .001$). All 3 circuits significantly reduced the frequency of problems encountered in verbal communication. Some test results suggested that CL and WDRC circuits provided a significantly better listening experience than PC circuits in word recognition ($P = .002$), loudness ($P = .003$), overall liking ($P = .001$), aversiveness of environmental sounds ($P = .02$), and distortion ($P = .02$). In the rank-order ratings, patients preferred the CL hearing aid circuits more frequently (41.6%) than the WDRC (29.8%) and the PC (28.6%) ($P = .001$ for CL vs both WDRC and PC).

Conclusions Each circuit provided significant benefit in quiet and noisy listening situations. The CL and WDRC circuits appeared to provide superior benefits compared with the PC, although the differences between them were much less than the differences between the aided vs unaided conditions.

JAMA. 2000;284:1806-1813 www.jama.com

SENSORINEURAL HEARING LOSS IS one of the most prevalent disabling conditions reported in the United States, affecting some 20 million to 26 million people.[1-3] Hearing loss is present in 35% to 42% of individuals older than 65 years.[4-6] It adversely affects physical, cognitive, behavioral, and social function, as well as the general quality of life,[7] and has been linked to depression and dementia.[8-10]

While hearing aids are the main form of treatment, only about 20% of those who could benefit from hearing aids wear them.[2,3,11] Moreover, surveys have suggested that about 50% of hearing aid users are dissatisfied with their instruments.[12] A recent survey, however, ed that because of improved technol- , approximately 65% of hearing sers are satisfied with their de- .[13]

vast array of hearing aid technolo- s available, ranging from simple and ively inexpensive analog circuits to plex and expensive digital devices require sophisticated fitting proce- s. Whereas numerous studies have nstrated that hearing aids provide ficant benefit compared with un- d listening for persons with hear- losses ranging from mild to se- ,[14-17] carefully controlled, multicenter ical trials of the relative benefit pro- by different types of hearing aids not been conducted. Laboratory ies and small-scale field studies have designed in ways that make them ult to compare and have failed to consistent superiority for any type rial processing.[18,19]

hoices among available hearing aids be made without the benefit of data well-designed clinical trials. This re- presents the results of a double- d, multicenter clinical trial to com- the efficacy of 3 different hearing aid uits. Efficacy was measured in a va- of listening situations using tests of ch understanding, sound quality, patient rank-order ratings. The 3 ring aid circuits jointly account for of the US hearing aid market.

The clinical trial compared 3 com- ly used hearing aid circuits: the lin- eak clipper (PC), the compression iter (CL), and the wide dynamic range pressor (WDRC). The PC and CL uits amplify input sounds linearly up predetermined level (usually set rela- to loudness discomfort levels). Above level, the output is limited using 2 ferent electronic methods. FIGURE 1 strates the major difference among the circuits. For the PC, as the input sig- level increases by 10 dB, so does the tput level up to its maximum output abilities when the instrument is said e in "saturation." The CL operates ilarly in that the output increases lin- up to a certain point. After that, wever, the output is reduced by cir- ry that automatically turns down the gain of the hearing aid by a fixed ratio. In this instance, the output is allowed to increase by 1 dB for each 10-dB increase in the input sound level. Finally, the WDRC behaves similarly to the CL circuit except that the automatic gain function begins at lower input sound levels and allows, in this instance, the output to increase 1 dB for each 2 dB increase in input sound level up to its point of maximum output.

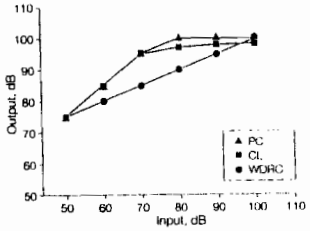
The PC removes the positive and/or negative peaks of the amplified signal, whereas the CL uses automatic volume control circuitry. A disadvantage of PC circuitry is that some acoustic distortion results when the output limit is exceeded.[20,21] Far less distortion is created by CL circuitry.[22] The WDRC circuit allows input signals that vary in level over a wide range to be amplified as a narrower range of output signals,[23,24] which is associated with the reduced dynamic range found in the majority of sensorineural hearing loss. Although theoretically beneficial to listener comfort and speech understanding, a disadvantage of compression circuits (eg, CL and WDRC) is that they alter the temporal characteristics of signals in a way that can be apparent to the listener.[25,26]

METHODS

Trial Design

Eight audiology laboratories located within Department of Veterans Affairs (VA) medical centers participated. The experimental design was a 3-period, 3-treatment crossover design. Baseline measurements were made using a battery of tests in the unaided condition (no hearing aids). Patients were then stratified by center and randomized to 1 of 6 sequences of the 3 hearing aid circuits. Six sequences were used so that each hearing aid circuit had approximately an equal number of patients who used the circuit first, second, and third. Each block of 6 consecutive patients within each center was balanced so that each sequence was represented once. The actual frequencies for the 6 sequences in the trial ranged from 59 to 61. In each of the 3 periods, the patients were fit binaurally and used 1 circuit (aided condition) for 3 months. At the end of each period, the battery of outcome tests was repeated in both the unaided and aided conditions. The protocol was conducted in double-blind fashion. Neither the audiologist who administered the tests nor the patient could identify the circuit being worn because all 3 hearing aid circuits resided in the same casement and because a different audiologist programmed the device.

Figure 1. Performance Characteristics of 3 Hearing Aid Circuits



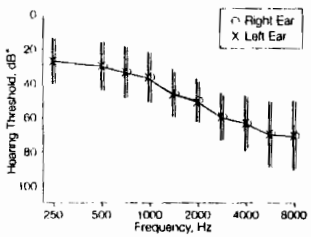
PC indicates linear peak clipper; CL, compression limiter; and WDRC, wide dynamic range compressor.

Patient Sample

The protocol was approved by the National Institute on Deafness and Other Communication Disorders (NIDCD) Hearing Aid Advisory Committee, the Hines VA Cooperative Studies Program Coordinating Center's Human Rights Committee, and by the institutional review board of each participating center. All patients provided informed consent, were fluent speakers of English, and had bilaterally symmetrical sensorineural losses with no evidence of retrocochlear pathology. Average audiometric thresholds for 500, 1000, 2000, 3000, and 4000 Hz were no better than 25-dB hearing level in either ear, with no threshold from 500 to 2000 Hz exceeding 70-dB hearing level. FIGURE 2 shows the mean (±1 SD) audiogram. To ensure that the sample included patients who were typical of the majority of adult hearing aid users, monaural word recognition scores

Author Affiliations and a complete list of the members of the NIDCD/VA Hearing Aid Clinical Trial group are listed at the end of this article.
Corresponding Author and Reprints: Vernon D. Larson, PhD, Howard Leight Industries, 7828 Waterville Rd, San Diego, CA 92154 (e-mail: vlarson@howardleight.com).

Figure 2. Mean Audiogram of Patient Sample (N=360)



Error bars for each ear indicate the range from 1 to −1 SD; the darker error bar on the left in each pair corresponds to the left ear. Asterisk indicates values based on American National Standards Institute guidelines for 1996.

on a recorded version of the Central Institute for the Deaf W-22 test[27] were required to be at least 28%, with a difference no greater than 26% between ears.

Experimental Apparatus

Each participant was fit binaurally with single channel, programmable, full-concha in-the-ear hearing aids (Dyna P2, Phonak, Stäfa, Switzerland) that contained all 3 circuit options. The National Institute of Standards and Technology evaluated prototypes to ensure that characteristics of the hearing aid conformed to the manufacturer's specifications.

The 3 programmable options were PC, CL, and WDRC. The CL had an 8:1 compression ratio (above compression threshold, an 8-dB increase in the input level resulted in only a 1-dB increase in the output) and duration-dependent release time capability. The WDRC had a fixed-compression threshold (approximately 52-dB input sound pressure level [SPL]), a compression ratio that ranged from 1.1:1 to 2.7:1 and a short, fixed release time (50 milliseconds). The maximum output levels of the 3 circuits were programmable over approximately the same range of SPLs.

Electroacoustic measurements[28] were made at each visit to ensure that hearing aid characteristics remained stable. Acoustic gain targets were established using the NAL-R method,[29] and probe microphone procedures were used to verify that targets had been achieved. Maximum output targets were obtained using loudness discomfort levels[30] and were subsequently held constant across visits and circuit types.

Outcome Test Battery

All testing was carried out in audiometric test rooms using identical equipment for test presentation and data collection at each site. Three categories of outcome measures were used: speech recognition tests, category ratings of perceived sound quality, and self-assessed subjective ratings of hearing aid benefit.

Speech Recognition

Two tests of speech recognition were used. A recorded version of a monosyllabic word-recognition test, the NU-6,[27] was presented using a single loudspeaker (positioned at 0°; ie, patients faced the loudspeaker) at an SPL of 62 dB (conversational speech level). Each of the 4 NU-6 lists contains 50 scoreable items with each item having a value of 2%. At conversational speech levels, listeners with normal hearing obtain perfect monosyllabic word recognition scores. The second test, a recorded version of the Connected Speech Test (CST),[31,32] consists of 48 passages with 8 to 10 sentences that approximate everyday, connected discourse. Because it was unlikely that a single laboratory condition could represent the range of possible listening conditions, we conducted this test in a variety of presentation and background-noise levels. The CST was presented via the loudspeaker (located at 0° azimuth) at a level of 74-dB SPL (loud speech) in quiet and then again at 74 dB in 3 background noise conditions. For SPLs of 52 dB (soft) and 62 dB (conversational loudness), the speech materials were presented in 3 conditions of background noise.

The background noise used was an uncorrelated multitalker babble,[31] which was delivered from loudspeakers located at azimuths 45° left and right at nominal signal-to-babble (S/B) ratios of −3 dB, 0 dB, and 3 dB. The S/B ratio refers to the relationship of the SPL of the speech to the SPL of the background babble. The nominal 0-dB S/B condition was estimated during the baseline visit prior to conducting tests for each patient by presenting CST practice materials at 62-dB SPL in the "unaided" condition using a bracketing procedure in which the binaural babble level was varied for each subject to produce 50% intelligibility. (The mean [SD] level of the babble was 55 [5.4] dB.) This relationship for each subject was designated as the 0-dB condition. The same S/B ratio for each subject was used for the −3-dB and 3-dB conditions, and for all tests conducted over the 9-month protocol, the same ratios were used. Normal listeners typically receive perfect scores at loud and conversational levels in a quiet background, but their performance at softer levels and in the presence of background noise varies as a function of the difficulty of the listening situation.

Category Ratings

The Quality Rating Test, was used to assess 3 aspects of patients' perception of sound quality: loudness, noise interference, and overall liking of the listening experience. The patients rated each dimension on a 10-point scale. On the loudness scale, 1 was too soft; 10, too loud; and 5, comfortably loud. For overall liking, 1 was very poor or terrible and 10 was excellent. In this task, the patients were instructed to ignore the loudness of the speech and consider only the overall sound quality. For noise interference, they assigned a 1 if noisiness completely interfered with quality and understanding of the speaker and 10 if noisiness did not interfere. Intermediate integer ratings could be assigned for all tests. Sentences designated as practice sentences of the CST[31] served as the stimuli for the Quality Rating Test. Patients were presented 5 different sentences and provided a rating after each presentation, which were then averaged. The sentences were presented at an SPL of 52 dB, 62 dB, and 74 dB in a quiet background and then in the multitalker babble (S/B ratio, 10 dB).

Subjective Ratings

Two measures were used to elicit expressions of the quality of hearing aid performance from the patients. One was the Profile of Hearing Aid Performance/Profile of Hearing Aid Benefit (PHAP/PHAB), which quantifies 2 major aspects of hearing aid performance: speech communication in a variety of daily life situations and reactions to the loudness and quality of environmental sound.[33] For the PHAP/PHAB, 7 subscale scores were derived from the 66 items of the inventory that were completed by the patient in written format. The scales quantify problems in communication in favorable and unfavorable listening conditions as well as the aversiveness and distortion of a variety of sounds. The 7 subscales include communication with familiar talkers, ease of communication, reverberation, reduced cues, background noise, aversiveness of sound, and distortion of sound. At the end of each of the 3 trial periods, the patients completed the PHAP/PHAB inventory in the unaided and aided conditions using a 4-point scale that ranged from always to never. The PHAP scores quantify the scale scores in terms of aided performance, while PHAB scores quantify the scale scores in terms of benefit (ie, the difference between the aided scores and unaided scores). Hence, in the PHAP, scores for all subscales are reported in terms of percentage of time a problem is experienced and scores for the PHAB are reported in terms of the change in percentage of time a problem is experienced.

The second subjective assessment procedure was used at the final visit only. After having completed each of the 3 treatments, the patients provided, from memory, a rank-order rating of the 3.

Statistical Methods

A crossover design was chosen for this study instead of the more traditional randomized, parallel group design because it required fewer patients, eliminated between-patient variation, and it increased power for other objectives of the trial (eg, to determine which patient characteristics predict success with the different hearing aid circuits). In addition, some of the known disadvantages of the crossover design (eg, large dropout rate, instability of the patient's condition, and a large carryover effect) were not expected in this study. The 3 circuits were compared using aided scores and aided minus unaided scores (benefit scores) with a repeated measures analysis.

The sample of 360 patients provided at least 80% power to detect a small-to-medium effect size for the patients' rank-order rating among the 3 circuits. This sample size also provided greater than 95% power to detect a 7.2% difference in the NU-6 test, greater than 95% power to detect a 3.6% difference in the CST, greater than 90% power to detect a 20% difference in the Quality Rating Test, and 90% power to detect a 16.6% difference in the PHAB.

A mixed, repeated measures model was used to compare the 3 hearing aid circuits for the individual outcome variables. If the overall test was statistically significant, then pairwise comparisons were made between the groups using the Bonferroni procedure to adjust the α level for multiple tests. No adjustment was made for multiple outcomes. For this reason, P values close to .05 should be interpreted with caution. Sample sizes reported for specific tests and conditions departed somewhat from 360 because some patients did not complete the study, some were unable to perform the task, or, occasionally, the examiner was unable to follow the study's protocol.

RESULTS

Patient Sample

Four hundred forty-six patients were screened for inclusion in the trial and 360 (80.7%) were randomized. Of the patients who were not randomized, 15 were excluded on the basis of a single criterion, but most failed to meet 2 or more of the inclusion criteria. The main reasons included: air conduction thresholds exceeded 70 dB in either ear (20); a difference in pure tone averages between ears of more than 10 dB (17); mean air-bone gap exceeded 5 dB in either ear (13); or routine otoscopy did not reveal clear ear canals (13). In 7 instances, the audiologist did not feel that the patient was capable of performing the tasks required by the trial.

Of the 360 patients enrolled, 69.7% were military veterans. The mean age of the group was 67.2 years (range, 29-91 years). The racial/ethnic distribution approximated that of the US population: 78.6% were white; 12.2% black; 6.1% Hispanic; 1.9% Asian; and 1.1% Native American. Fifty-seven percent were men; women were mainly nonveteran patients who were authorized to be treated at VA medical centers for the purposes of this trial because the study grant funded the cost of the hearing aids and the time of the treating and evaluating audiologists. The most common self-reported causes of the patient's hearing loss were noise exposure and aging. About half (46.7%) had never used a hearing aid.

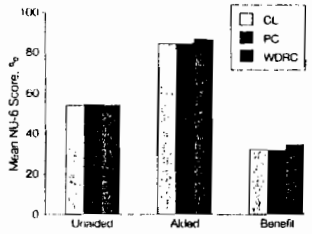
The number of patients from each center was nearly equal (range, 44-46). None of the groups representing the 6 randomized orders were statistically different in terms of age, age at onset of hearing loss, sex, race, previous hearing aid usage, and degree of hearing loss ($P \geq .11$ for all comparisons). Twenty-nine of the 360 patients did not complete the trial due to illness, relocation of residence, or other reasons (eg, withdrawal of patient consent, illness unrelated to hearing, death, sudden change in hearing). Three hundred thirty-seven patients completed the 90-day trial with the PC circuit, 338 with the CL, and 333 with the WDRC. The average reported hearing aid use time for the 3 circuits did not differ significantly and averaged about 9.8 (SD, 4) hours per day.

Speech Recognition Tests

FIGURE 3 provides a summary of the mean percentage correct results for the unaided and aided conditions for the

NU-6 test for each of the 3 circuits together with the benefit scores (aided minus unaided). For statistical testing, percentage correct scores for the NU-6 test were arcsine transformed to stabilize the error variance.[14] Comparison of the unaided means with the aided means showed that each of the 3 circuits improved the mean word recognition score by a substantial amount (approximately 29% in absolute score differences; $P<.001$). The overall statistical test comparing the 3 circuits was significant ($P=.002$ for the aided scores and $P=.002$ for the benefit scores). Pairwise comparison tests showed that the WDRC circuit was superior to the other 2 circuits for the aided scores and superior to the PC circuit for the benefit scores.

Figure 3. Mean Percentage of Correctly Recognized Monosyllabic Words



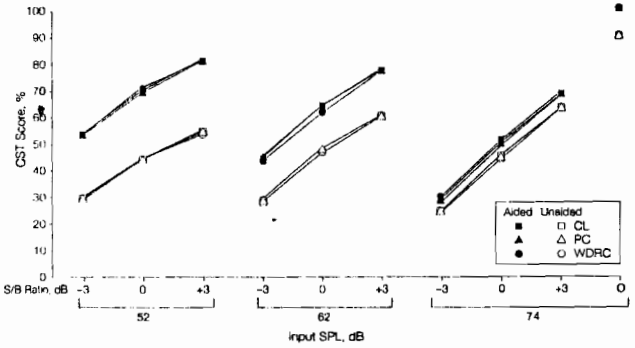
CL indicates compression limiter; PC, linear peak clipper; and WDRC, wide dynamic range compressor. The unaided values were obtained in the same test session as the aided values. Sample sizes varied slightly across conditions (320 to 324 for the aided conditions and 288 to 291 for the unaided conditions).

FIGURE 4 summarizes the findings for the aided and unaided CST results. Percentage correct scores were arcsine transformed to stabilize the error variance.[14] As expected, there were no differences among unaided means. However, significantly higher CST scores ($P<.001$) were achieved for all aided conditions relative to the unaided conditions. The overall statistical test comparing the 3 circuits for aided CST scores was significant for 1 condition (62/0; $P=.006$). Pairwise comparisons showed that the WDRC circuit was inferior to the CL and PC circuits.

The mean CST benefit scores (aided minus unaided) are shown in FIGURE 5. Comparison of the 3 circuits showed significant differences for the 62/0 ($P=.04$) and 74/0 conditions ($P=.02$). Pairwise comparison tests showed that for the 62/0 condition, the WDRC circuit was inferior to the CL circuit; and for the 74/0 condition, the WDRC circuit was superior to the CL circuit.

The data presented in Figures 4 and 5 also show that the 3 circuits provided similar amounts of improvement in test scores, but all showed successively less benefit as a function of signal level when background noise was present. A marked decrease in CST benefit scores from about 26% for the 52-dB conditions to approximately 6% for the 74-dB conditions was observed, suggesting that the hearing aids were less helpful at higher than at lower and moderate input levels. Furthermore, the Figures show that all 3 circuits provide measurable benefit in noisy conditions.

Figure 4. Connected Speech Recognition of Different Ratios of Signal to Ambient Noise



CL indicates compression limiter; PC, linear peak clipper; WDRC, wide dynamic range compressor; and SPL, sound pressure level. Mean unaided and aided scores on the Connected Speech Test (CST) shown for 10 test conditions for the 3 circuits. The abscissa labels the conditions by signal/babble levels (dB) for 9 conditions (52/-3 through 74/+3). S/B indicates signal-to-babble; Q, that the test was performed in quiet. Sample sizes varied across conditions (from 280 to 322 for unaided conditions and 320 to 336 for aided conditions).

Quality Rating Test

The Quality Rating Test was administered at 3 signal levels in quiet (designated as 52Q, 62Q, 74Q) and elicited ratings of loudness, noise interference, and overall liking. It was also administered at the same signal levels with an absolute S/B ratio of 10 dB (designated as 52N, 62N, and 74N), which means that the level of the speech was 10 dB greater than the level of the multitalker babble.

TABLE 1 shows no differences in the loudness ratings between the unaided means for each condition. Significant differences were observed, however, for the aided condition across the 3 circuits for both the quiet and background noise conditions for the lowest (52-dB SPL) and for the highest signal levels (74-dB SPL) ($P<.001$). The WDRC circuit was rated as being more comfortably loud (ie, a rating closer to 5) than the other 2 circuits for the 52-dB SPL conditions ($P=.003$) and 74-dB SPL conditions ($P=.003$). The CL circuit was more comfortably loud compared with the PC circuit for the 74-dB SPL condition.

A summary of data for the noise interference task is shown in TABLE 2. Analysis of the mean unaided data revealed no differences. For the aided data, the analysis also showed no significant differences among circuit types, except for the 62N condition ($P=.01$). Pairwise comparison revealed that the PC circuit scored higher (less noise interference) than the WDRC circuit.

A summary of data for the overall liking task is shown in TABLE 3. There were no significant differences between the unaided means at each condition. The analyses of the data among circuit types for the aided condition showed significance for the 74Q condition ($P=.001$). Pairwise comparisons across circuits showed that the PC was less liked than both the CL and the WDRC.

Finally, for each circuit, significant improvement in overall liking was observed for soft and conversational speech levels ($P\leq.05$). For the loud conditions (74Q, 74N), however, negative average benefit ratings were observed ($P\leq.01$) except for the 74N condition for the WDRC ($P=.39$), suggesting that the aided experience was rated as being less liked than the unaided experience for loud sounds.

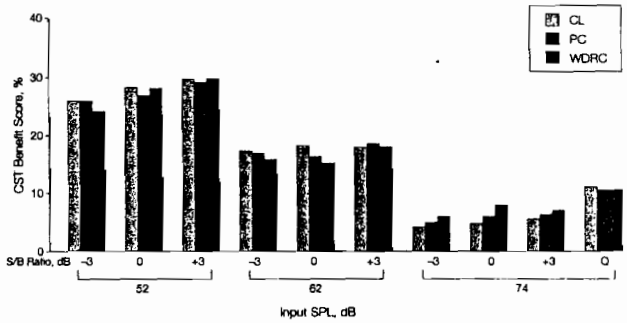
Objective Assessment

No differences were observed among the unaided means for the PHAP. For the aided means, the analysis showed statistical significance ($P<.001$) for 2 of the 7 scales: distortion of sounds and aversiveness of environmental sounds. Pairwise comparisons showed that the scores for the PC were significantly different (ie, higher frequency of problems) than both the CL and WDRC circuits on the aversion and distortion scales ($P\leq.02$). The mean values for the CL circuit were 4% to 5% higher (ie, lower frequency of problems) for aversion and were 2% to 3% higher for distortion.

The PHAB scores also showed that the CL circuit significantly reduced the frequency of problems reported on 6 of the 7 scales ($P<.001$). For aversion, however, all circuits produced a significantly higher frequency of problems ($P<.001$) than in the unaided condition. In the analysis comparing circuit types, significant differences were observed for aversion ($P<.001$) and distortion ($P=.02$). Pairwise comparisons for aversion showed that the PC circuit was more aversive than both of the other circuits ($P=.003$) with the mean frequency of problems being 4% to 5% higher. The PHAB scores also showed that the PC score was significantly higher than the WDRC for distortion ($P=.02$) with the mean difference between PC and WDRC being 3%.

Finally, on completion of the study, the patients provided, from memory, a ranking of the 3 hearing aid circuits. The CL circuit received the highest percentage of first rankings (41.6%), followed by the WDRC (29.8%), and the PC (28.6%). In addition, the CL circuit was ranked third by the lowest percentage of patients (25.4% for the CL vs 36.2% for the PC and 38.4% for the WDRC).

Figure 5. Mean Benefit Scores on the CST of Different Ratios of Signal to Ambient Noise



Abbreviations are expanded in the legend of Figure 4. The abscissa labels the conditions as signal/babble level (dB) for 9 conditions (52/−3 through 74/+3). S/B indicates signal-to-babble; Q that the test was performed in quiet. Sample sizes across conditions varied from 277 to 321.

Table 1. Loudness Ratings Obtained in the Aided and Unaided Conditions*

Condition, dB	Type of Circuit	Aided			Unaided		
		No. of Subjects	Mean (SD)	*P* Value	No. of Subjects	Mean (SD)	*P* Value
52Q	CL	335	4.14 (0.94)	<.001†	318	2.77 (1.24)	.97
	PC	333	4.15 (1.00)		314	2.77 (1.31)	
	WDRC	330	4.43 (0.99)		312	2.78 (1.26)	
52N	CL	335	3.92 (1.06)	<.001†	318	2.46 (1.23)	.92
	PC	333	3.81 (1.08)		316	2.45 (1.24)	
	WDRC	328	4.21 (1.07)		312	2.43 (1.22)	
62Q	CL	335	5.33 (0.90)	.26	325	4.36 (1.21)	.48
	PC	333	5.34 (0.90)		325	4.31 (1.22)	
	WDRC	330	5.41 (1.00)		320	4.29 (1.16)	
62N	CL	334	5.31 (1.19)	.35	326	3.85 (1.27)	.73
	PC	333	5.20 (1.09)		325	3.86 (1.33)	
	WDRC	330	5.25 (1.31)		320	3.90 (1.37)	
74Q	CL	335	7.96 (1.55)	<.001‡	335	6.30 (1.63)	.84
	PC	333	8.31 (1.49)		332	6.28 (1.64)	
	WDRC	330	7.73 (1.59)		330	6.30 (1.67)	
74N	CL	335	7.57 (1.89)	<.001‡	335	5.88 (1.72)	.72
	PC	333	8.11 (1.69)		334	5.81 (1.74)	
	WDRC	330	7.26 (1.94)		330	5.87 (1.84)	

*CL indicates compression limiter; PC, linear peak clipper; and WDRC, wide dynamic range compressor.
†Pairwise comparison reveals PC and WDRC, and CL and WDRC are statistically significant.
‡Pairwise comparison reveals all 3 circuits are statistically significantly different.

Statistical analysis using the Friedman test showed a significant overall difference among the rankings (P=.002). Subsequent analyses using the Wilcoxon test showed that, overall, the CL was preferred more frequently than the PC (P=.001) and the WDRC (P=.001) and that there were no significant differences between the rankings for the PC and the WDRC (P=.86).

Table 2. Noise Interference Ratings Obtained in the Aided and Unaided Conditions*

		Aided			Unaided		
Condition, dB	Type of Circuit	No. of Subjects	Mean (SD)	P Value	No. of Subjects	Mean (SD)	P Value
52Q	CL	335	9.74 (0.93)		317	9.53 (1.33)	
	PC	332	9.80 (0.79)	.42	314	9.57 (1.15)	.16
	WDRC	329	9.74 (0.86)		309	9.45 (1.46)	
52N	CL	336	7.10 (2.24)		316	6.96 (2.43)	
	PC	332	7.18 (2.24)	.56	315	6.93 (2.41)	.99
	WDRC	330	7.07 (2.19)		309	6.95 (2.47)	
62Q	CL	336	9.83 (0.66)		323	9.66 (1.25)	
	PC	332	9.85 (0.53)	.86	324	9.74 (0.95)	.22
	WDRC	328	9.83 (0.57)		318	9.66 (1.15)	
62N	CL	336	6.41 (2.28)		320	6.36 (2.28)	
	PC	333	6.55 (2.34)	.01†	322	6.37 (2.29)	.62
	WDRC	329	6.17 (2.38)		317	6.47 (2.23)	
74Q	CL	336	9.74 (0.73)		332	9.67 (1.22)	
	PC	331	9.73 (0.85)	.92	330	9.74 (1.07)	.22
	WDRC	329	9.72 (0.77)		327	9.66 (1.20)	
74N	CL	336	5.13 (2.47)		334	5.48 (2.31)	
	PC	333	5.18 (2.48)	.69	333	5.37 (2.36)	.80
	WDRC	329	5.25 (2.65)		329	5.43 (2.34)	

*CL indicates compression limiter; PC, linear peak clipper; and WDRC, wide dynamic range compressor.
†Comparison of PC circuit with WDRC circuit, which is statistically significantly different in pairwise comparison tests.

Table 3. Overall Liking Ratings Obtained in the Aided and Unaided Conditions*

		Aided			Unaided		
Condition, dB	Type of Circuit	No. of Subjects	Mean (SD)	P Value	No. of Subjects	Mean (SD)	P Value
52Q	CL	335	7.70 (2.36)		316	5.63 (2.93)	
	PC	333	7.78 (2.39)	.14	314	5.56 (2.80)	.88
	WDRC	330	7.96 (2.21)		308	5.57 (2.81)	
52N	CL	335	6.77 (2.38)		316	4.86 (2.77)	
	PC	332	6.66 (2.45)	.64	314	4.87 (2.74)	.72
	WDRC	330	6.76 (2.33)		309	4.75 (2.76)	
62Q	CL	336	8.44 (1.85)		322	7.28 (2.53)	
	PC	331	8.43 (1.84)	.97	323	7.31 (2.55)	.28
	WDRC	330	8.45 (1.78)		318	7.14 (2.51)	
62N	CL	335	6.85 (2.27)		321	5.86 (2.55)	
	PC	330	6.97 (2.30)	.07	323	5.82 (2.51)	.90
	WDRC	330	6.70 (2.30)		318	5.79 (2.60)	
74Q	CL	336	7.43 (2.60)		331	7.73 (2.31)	
	PC	332	6.91 (2.76)	.001†	330	7.94 (2.23)	.08
	WDRC	329	7.49 (2.54)		325	7.72 (2.24)	
74N	CL	336	5.80 (2.57)		334	6.04 (2.39)	
	PC	331	5.59 (2.77)	.12	333	6.05 (2.55)	.70
	WDRC	330	5.92 (2.64)		329	5.97 (2.47)	

*CL indicates compression limiter; PC, linear peak clipper; and WDRC, wide dynamic range compression.
†Comparison of CL circuit with PC circuit and comparison of PC circuit with WDRC circuit are statistically different in pairwise comparison tests.

COMMENT

Each of the 3 hearing aid circuits provides substantial benefit over unaided listening. Benefit was observed for measures of speech recognition and ratings of speech quality in a variety of noisy and quiet conditions as well as for subjective measures. Each circuit improved monosyllabic word recognition scores in a quiet background at conversational levels by an average of 29% (absolute score improvement). Speech recognition ability, as shown by the CST, in noise was improved by each circuit by amounts ranging from 10% to 30% (absolute score improvement), with greater improvement observed for speech at soft and conversational levels. Loudness rating data suggested that all 3 circuits amplified soft and conversationally loud speech to comfortable levels. The noise interference ratings showed that none of the circuits had a deleterious effect. For soft and conversational speech levels, each circuit improved the overall quality of the listening experience. For loud speech, the overall quality of listening was not significantly degraded. The results of 6 of the 7 subscales of the subjective measure of hearing aid benefit (PHAB) showed a significant reduction in the frequency of problems associated with communication in everyday environments.

Statistically significant differences (small in comparison with the benefits seen with each of the circuits) were noted among the circuits on several components of the outcome measures. The results of the loudness rating suggest that the WDRC circuit was more comfortably loud than the other 2 circuits for soft and for loud speech input conditions. Because of its operating characteristics, the WDRC was expected to produce a more comfortable listening experience for the soft and loud input levels. Differences in scores on the PHAP/PHAB for 2 subscales were statistically significant among circuits, with the PC rated as 4.5% more aversive than the other 2 circuits and producing an average of 3% more problems for distortion of sounds compared with the WDRC circuit. The preference rankings provided at the end of the trial favored the CL circuit. Because the differences between the hearing aid circuits were small in most cases, dispensing decisions should take into account cost vs benefit considerations for individual patients. In this regard, many programmable hearing aids (such as the one used in this trial) may be configured to function as a PC, CL, or WDRC and, as such, there are no cost differences between circuit options; however, for conventional, nonprogrammable devices, compression circuitry (either CL or WDRC) adds significantly to the single-unit price of the device.

Because concerted efforts were made to recruit patients into the study from both sexes and all racial groups, the study sample was a good representation of US adults who are candidates for hearing aids. We believe, therefore, that the study results are generalizable to the US population with sensorineural hearing loss. One limitation of the trial is that it did not measure other domains, such as affect and cognition, which are influenced by hearing loss.

Author Affiliations: Department of Veterans Affairs Medical Center, Washington, DC (Drs Larson, Beck, and Boysen); Hines VA Cooperative Studies Program Coordinating Center, Hines, Ill (Mr Williams and Dr Henderson); National Institute on Deafness and Other Communication Disorders, Bethesda, Md (Drs Luethke and Donahue); Greater Los Angeles VA Health Care System, Los Angeles, Calif (Dr Noffsinger); Department of Veterans Affairs Medical Center, Mountain Home, Tenn (Dr Wilson); University of Texas Health Sciences Center, San Antonio (Dr Dobie); Department of Veterans Affairs Medical Center, Iowa City, Iowa (Dr Haskell); Department of Veterans Affairs Medical Center, Nashville, Tenn (Dr Bratt); Department of Veterans Affairs Medical Center, Long Beach, Calif (Dr Shanks); Boys Town National Research Hospital, Omaha, Neb (Dr Stelmachowicz); University of Memphis, Memphis, Tenn (Dr Studebaker); Harbor-UCLA/University of California, Los Angeles, Medical Center, Torrance (Dr Canalis); Department of Veterans Affairs Medical Center, Portland, Ore (Dr Fausti); and Department of Veterans Affairs Medical Center, Albuquerque, NM (Dr Rappaport).

NIDCD/VA Hearing Aid Clinical Trial Group: Vernon D. Larson, PhD,* chairman and principal investigator; Jillyn Roxberg, MA, clinical coordinator, and Mary Lee Jacob, program assistant, Veterans Affairs Medical Center, Mountain Home, Tenn; Joyce S. Kang, MA, clinical coordinator, and Megan M. Wilkinson, BA, program assistant, Veterans Affairs Medical Center, Washington, DC.

Participating Veterans Affairs Medical Centers, Participating Investigators and Audiologists: Albuquerque, NM: Bruce Z. Rappaport, PhD, Karen K. Stewart, MA, Sheree Smith, MA, and Jonni K. Dressendorfer, MA; Iowa City, Iowa: George B. Haskell, PhD, Kathy Vander Werff, MA, Adriana Rodriguez, MA, and Louise L. Rogers, MA; Long Beach, Calif: Janet E. Shanks, PhD, Charlene M. Mikami, MA, and Donna Schmitt, MA; Mountain Home, Tenn: Richard H. Wilson, PhD, Anne L. Oyler, MA, Kristen L. Bauer, MA, Colleen Noe, PhD, and Jennifer Crocket, MA; Nashville, Tenn: Gene W. Bratt, PhD,* Mia A. L. Rosenfeld, MS, Kjersten Branscome, MS, Mary Beth Trine, MS, and Barbara Peek, PhD; Portland, Ore: Stephen A. Fausti, PhD, James Henry, PhD, Wendy Helt, MA, Craig Larson, MA, and Alison Gilbert, MS; Washington, DC: Lucille B. Beck, PhD,* Gerald Schuchman, PhD, and Holly A. Sandvik, MS; and West Los Angeles, Calif: Douglas Noffsinger, PhD,* Donald D. Dirks, PhD, Diane Kenworthy, MA, Sheril A. Plunkett, MS, and Eleanor L. Wilson, MS.

Hines VA Cooperative Studies Program Coordinating Center, Hines, Ill: William G. Henderson, PhD,* director, Domenic Reda, MS, associate director, David Semlow, BS, assistant director, David W. Williams, MS,* study biostatistician, Mary Biondic, BA, programmer, and Barbara J. Christine, statistical assistant.

National Institute on Deafness and Other Communication Disorders, Bethesda, Md: James B. Snow, Jr, MD, retired director; Lynn E. Luethke, PhD,* health scientist administrator, and Amy Donahue, PhD,* chief, Hearing and Balance/Vestibular Sciences Branch.

Department of Veterans Affairs Headquarters, Washington, DC: John R. Feussner, MD, chief research and development officer, Joseph Gough, MA, program manager, and Ping C. Huang, PhD, staff assistant, Cooperative Studies Program; Lucille B. Beck, PhD,* director, and Allen E. Boysen, PhD,* retired director, Audiology and Speech Pathology Service.

Executive Committee: Rinaldo Canalis, MD, Torrance, Calif; Robert A. Dobie, MD, San Antonio, Tex; Patricia Stelmachowicz, PhD, Omaha, Neb; and Gerald A. Studebaker, PhD, Memphis, Tenn.

An asterisk indicates that a person is an Executive Committee member in the other sections above.

Data and Safety Monitoring Board: Noel Cohen, MD, chairman, New York, NY; Marilyn Demorest, PhD, Baltimore, Md; G. Pamela Burch-Sims, PhD, Nashville, Tenn; Wayne O. Olsen, PhD, Rochester, Minn; Richard A. Chole, MD, PhD, St Louis, Mo; Julianna Gulya, MD, Washington, DC; Debra L. Tucci, MD, Durham, NC; and Marie Diener-West, PhD, Baltimore, Md.

Funding/Support: Financial support was provided by the National Institute on Deafness and Other Communication Disorders and the Cooperative Studies Program of the Department of VA Research and Development Service.

REFERENCES

1. Collins JG. Prevalence of selected chronic conditions. *Vital Health Stat 10.* 1997;194:1-89.
2. Ries PW. Prevalence and characteristics of persons with hearing trouble: United States, 1990-1991. *Vital Health Stat* 10. 1994;188:1-75.
3. Kochkin S. MarkeTrak III. *Hear J.* 1992;45:47-54.
4. Goldstein DP. Hearing impairment, hearing aids and audiology. *ASHA.* 1984;26:24-35,38.
5. Leske MC. Prevalence estimates of communicative disorders in the US. *ASHA.* 1981;23:229-237.
6. Havlik RJ. Aging in the eighties, impaired senses for sound and light in persons age 65 and over. *Vital Health Stat 125.* 1986;125:1-8.
7. Mulrow CD, Aguilar C, Endicott JE, et al. Quality-of-life changes and hearing impairment: a randomized trial. *Ann Intern Med.* 1990;113:188-194.
8. Mulrow CD, Aguilar C, Endicott JE, et al. Association between hearing impairment and the quality of life of elderly individuals. *J Am Geriatr Soc.* 1990;38:45-50.
9. Uhlmann RF, Larson EB, Koepsell TD. Hearing impairment and cognitive decline in senile dementia of the Alzheimer's type. *J Am Geriatr Soc.* 1986;34:207-210.
10. Palmer CV, Adams SW, Bourgeois M, Durrant J, Rossi M. Reduction in caregiver-identified problem behaviors in patients with Alzheimer disease post-hearing-aid fitting. *J Speech Lang Hear Res.* 1999;42:312-328.
11. Russell JN, Hendershot GE, LeClere F, et al. Trends and differential use of assistive technology devices. *Vital Health Stat* 292. 1997;292:1-9.
12. Kochkin S. MarkeTrak III identifies key factors in determining consumer satisfaction. *Hear J.* 1992;45:39-44.
13. Kochkin S. Customer satisfaction and subjective benefit with high performance hearing aids. *Hear J* 1996;49:16-26.
14. Cox RM, Alexander GC. The abbreviated Profile of Hearing Aid Benefit. *Ear Hear.* 1995;16:176-186.
15. Humes LE, Christensen LA, Bess FH, Hedley-Williams A. A comparison of the benefit provided by well-fit linear hearing aids and instruments with automatic reductions of low-frequency gain. *J Speech Lang Hear Res.* 1997;40:666-685.
16. Humes LE, Christensen L, Thomas T, Bess FH, Hedley-Williams A, Bentler R. A comparison of the aided performance and benefit provided by a linear and a two-channel wide dynamic range compression hearing aid. *J Speech Lang Hear Res.* 1999;42:65-79.
17. Dillon H, Birtles G, Lovegrove R. Measuring the outcomes of a national rehabilitation program. *J Am Acad Audiol.* 1999;10:67-79.
18. Hickson L. Compression amplification in hearing aids. *Am J Audiol.* 1994;3:51-65.
19. Dillon H. Compression? *Ear Hear.* 1996;17:4,287-307.
20. Anderson BA, Hagen LT, Peterson TS, et al. Hearing instrument distortion. *Hear Rev.* 1996;69:18-26.
21. Hawkins DB, Naidoo SV. Comparison of sound quality and clarity with asymmetrical peak clipping and output limiting compression. *J Am Acad Audiol.* 1993;4:221-228.
22. Krebs D. Distortion and methods of output limiting. *Scand Audiol.* 1972;1:167-175.
23. Braida LD, Durlach NI, Lippmann RP, et al. Hearing aids—a review of past research on linear amplification, amplitude compression, and frequency lowering. *ASHA Monogr.* 1979;19:1-114.
24. Moore BCJ. Signal processing to compensate for reduced dynamic range. In: Beilin J, Jensen GR, eds. *Recent Developments in Hearing Instrument Technology, 15th Danavox Symposium, Kolding, Denmark, Mar 30-Apr 2, 1993.* Scanticon, Denmark: Danavox Jubilee Foundation; 1993:147-165.
25. Teder H. Compression in the time domain. *Am J Audiol.* 1993;2:41-46.
26. Plomp R. The negative effect of amplitude compression in multichannel hearing aids in the light of the modulation-transfer function. *J Acoust Soc Am.* 1988;83:2322-2327.
27. Wilson RH. Development and use of auditory compact discs in auditory evaluation. *J Rehabil Res Dev.* 1993;30:342-351.
28. American National Standards Institute. *Specification of Hearing Aid Characteristics.* New York, NY: Acoustical Society of America; 1996. ANSI S3.22-1987.
29. Byrne D, Dillon H. The National Acoustic Laboratories' (NAL) new procedure for selecting the gain and frequency response of a hearing aid. *Ear Hear.* 1986;7:257-265.
30. Hawkins DB, Walden BE, Montgomery A, et al. Description and validation of an LDL procedure designed to select SSPL90. *Ear Hear.* 1987;8:162-169.
31. Cox RM, Alexander GC, Gilmore C. Development of the Connected Speech Test (CST). *Ear Hear.* 1987;8(5 suppl):119S-126S.
32. Cox RM, Alexander GC, Gilmore C, Pusakulich KM. Use of the Connected Speech Test (CST) with hearing-impaired listeners. *Ear Hear.* 1988;9:198-207.
33. Cox RM, Gilmore C. Development of the Profile of Hearing Aid Performance (PHAP). *J Speech Lang Hear Res.* 1990;33:343-357.
34. Studebaker GA. A "rationalized" arcsine transform. *J Speech Lang Hear Res.* 1985;28:455-462.

Topical butyrate for acute radiation proctitis: randomised, crossover trial (320) [3702]

P Vernia, P L Fracasso, V Casale, G Villotti, A Marcheggiano, V Stigliano, P Pinnaro, V Bagnardi, R Caprilli

Summary

Background No available therapy has, as yet, proven effective to treat acute radiation proctitis (ARP) following radiation therapy for malignant pelvic disease. We assessed whether sodium butyrate enemas, at a dose of 80 mmol/L (80 mL/24 h), might offer effective treatment for this condition.

Methods 20 patients presenting with ARP after completing a cycle of 35–52 Gy external-beam radiation therapy for pelvic malignant disease, were treated for 3 weeks with topical sodium butyrate and saline enemas according to a randomised, double-blind, crossover protocol. Clinical, endoscopic, and histological findings were assessed at enrolment, at week 3, and then at the end of the study. Data were analysed by two-tailed *t* test for paired data (continuous variables) and a logistic-regression model with variable multiple response for ordered categorical data.

Findings Topical butyrate, but not saline, led to remission of symptoms (clinical score from 8·2 [SE 1·6] to 1·5 [0·7] *vs* 7·9 [1·8] to 8·1 [3·4]). When the treatment regimen was switched, eight out of nine of the previously placebo-treated patients went into remission, whereas three patients relapsed when switched to saline. The advantage of butyrate over placebo, expressed as CI, odds ratio, and p value was significant for almost all the clinical, endoscopic and histological factors taken into consideration.

Interpretation Topical sodium butyrate, unlike other therapeutic regimens used so far, proved effective in the treatment of ARP.

Lancet 2000; **356:** 1232–35

Introduction

Radiotherapy for malignant pelvic disease is often followed by acute treatment-induced bowel injury, the prevalence being related to the total radiation dose and therapeutic protocol.[1,2] The risk of chronic proctitis is lower, ranging between 5% and 20% at 5 years, after pelvic radiation therapy.[3,4]

Although most patients have only mild side-effects during treatment and heal spontaneously after completing therapy, no proven effective treatment of acute radiation proctitis (ARP) is available as yet. Topical steroids and sucralfate, in some instances, improve mucous and bloody diarrhoea, tenesmus, and pain,[5,6] but true efficacy is still debated. The crucial issue seems to be, however, the prevention of acute damage, since clinical and experimental data indicate that early damage to the intestinal mucosal lining and the vascular endothelium leads to late mucosal injury.[7] Acute toxicity is, indeed, the only factor, besides the total radiation dose, significantly correlated with increased risk of chronic proctitis,[8] but again, no effective therapeutic regimen has so far proved effective in the prevention of ARP.

Among new therapeutic approaches, short-chain fatty acids (SCFA) and butyrate are likely candidates. Derived from the bacterial fermentation of unabsorbed carbohydrates within the colonic lumen, they have a pivotal role in the regulation of mucosal proliferation and provide over half the energy requirements of the mucosa. Butyrate, besides being the main contributor to these processes, favours mucosal repair through transglutaminase-mediated and non-transglutaminase-mediated pathways,[9] dilates resistance arteries increasing mucosal blood flow and oxygen uptake,[10,11] reduces mucosal permeability, and enhances production and release of mucus.[12] Some preliminary data indicate that SCFA may be effective in acute, but not in chronic, radiation proctitis,[13] but no hard data are available.

The present study was thus aimed at assessing the efficacy of topical butyrate in the management of ARP, by a randomised, double-blind, placebo-controlled crossover protocol.

Methods

Patients

Over a period of 13 months, 58 patients completed a cycle of external-beam radiation therapy to the pelvis (35–52 Gy) for prostate or cervical cancer. Every week, patients had received five daily fractions of 1·8–2 Gy, with a four-field box centered on the pelvis. Enrolled in the study were 20 patients (11 male and nine female, mean age 57·0 years [SD 6·8]) out of 25 who had ARP within 3 weeks of completing radiation therapy. Five patients refused to enter the study. ARP was defined as the occurrence of loose stools and tenesmus lasting more than 7 days, associated with at least two of the following symptoms: greater than four bowel movements per day, rectal bleeding, night bowel movements, and abdominal/rectal pain. Patients with a history of chronic diarrhoea or inflammatory bowel diseases, or who had used topical treatment for haemorrhoids in the 4 weeks preceding the study, were excluded. All patients gave written informed consent to take part in the study, which had been approved by the ethics committee of our institution.

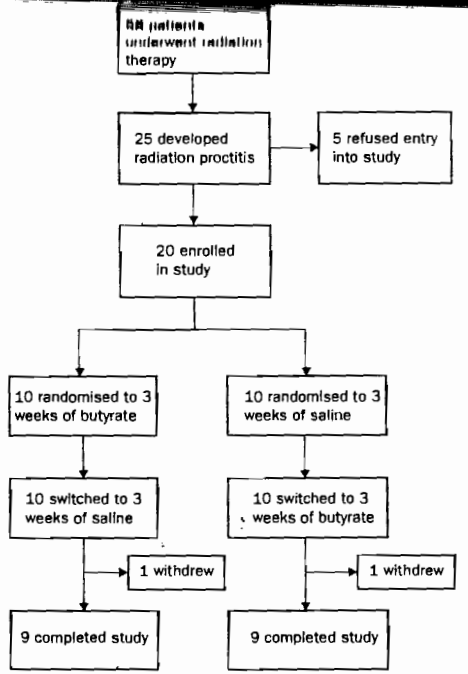


Figure 1: **Trial profile**

Study design

In accordance with a double-blind, placebo-controlled, crossover protocol, the patients were submitted to a 3 week treatment period with either one 80 mL enema per day of sodium butyrate (80 mmol/L, ten patients) or a sodium chloride (placebo) solution (ten patients) once a day, in random order. One drop of butyric acid was applied to the cap of the saline-containing bottles to mimic the smell of the butyrate solution. Patients were allocated to the two treatment arms according to a computer-generated randomisation list and were instructed to self-administer enemas each evening before sleeping. Patients were assessed clinically, endoscopically, and histologically before entry to the study, at week 3, and at the end of the study. The physicians were unaware of the treatment regimen received by the individual patients. To assess compliance, patients were asked to return unused enemas during the visit at week 3 and at the end of the study.

Stool consistency	40·8 (3·1–528·3)	0·005
Bloody stools	1·7 (1·0–194·4)	0·01
Abdominal pain	4·8 (0·7–32·1)	0·10
Burning/tenesmus	16·5 (1·5–181·3)	0·04
Self evaluation	23·4 (1·9–276·4)	0·01
Endoscopy	10·4 (0·9–114·2)	0·05
Histology (inflammation)	6·2 (0·9–42·2)	0·06
Histology (epithelial damage)	1·7 (0·3–9·1)	0·52
Bowel movements	−0·95 (−0·83 to −0·16)	0·008
Night bowel movements	−0·35 (−0·57 to −0·50)	0·0006

Data refer to first 3 weeks of treatment. OR=odds ratio.

Advantage of butyrate over placebo in the treatment of ARP

The clinical activity of ARP was assessed by recording bowel movements and stool consistency, rectal bleeding, night bowel movements, abdominal pain, rectal burning/tenesmus, and from a self-rating based on the impact of symptoms on normal life activities. For each factor, a scale from 0 (normal) to 3 (highly abnormal) was used. The maximum overall score was 21 and the minimum score required for enrolment was 6. The endoscopic assessment was graded as follows: 0=normal mucosal pattern; 1=mucosal erythema, slight oedema; 2=pronounced oedema, superficial erosions, friable mucosa with bleeding on touch; 3=mucosal ulcerations, spontaneous bleeding. The histological scoring system was also graded from 0 (normal mucosa) to 3 (defined as severe inflammation/damage), according to the degree of superficial epithelium impairment/goblet cell depletion, and inflammatory cell infiltration. 18 patients completed the study, and two did not complete the second 3 weeks period.

Statistical evaluation

Student's *t* test for paired data was used in the comparison between groups for continuous outcome variables (number of bowel movements). All other variables were analysed by logistic regression for ordered categorical data.[14] Since a considerable carryover effect was observed when butyrate preceded the placebo treatment, an analysis was also done to compare data from the first 3 weeks of treatment.

Results

Figure 1 shows the trial profile. No important differences were observed between groups before treatment. All patients first treated with butyrate became symptom-free or improved greatly (clinical score from 8·2 [SE 1·6] to 1·5 [0·7]) within the first 3 weeks of treatment. In the placebo group three patients had some improvement whereas three deteriorated slightly. Thus, the overall score value was unchanged (7·9 [1·8] *vs* 8·1 [3·4]). The number of bowel movements decreased in the butyrate group from 3·7 (0·6) to 1·7 (0·2), but remained unchanged in the placebo group (from 2·8 [0·3] to 2·6 [0·2]).

When the treatment regimen was switched to butyrate, clinical remission occurred in eight of nine of the previously placebo-treated patients. Three of the patients treated with butyrate had a relapse of ARP symptoms when switched to saline (clinical score from 1·5 [0·7] to 2·6 [2·1]). Pretreatment endoscopy showed, in most cases, only mild lesions, consisting of

Cattedra di Gastroenterologia 1, Università La Sapienza, Roma (P Vernia MD, G Villotti MD, A Marcheggiano MD, R Caprilli MD); **Istituto Regina Elena, Roma** (P L Fracasso MD, V Casale MD, V Stigliano MD, P Pinnaro MD); **and Dipartimento di Statistica, Università di Milano, Italy** (V Bagnardi MD)

Correspondence to: Dr Piero Vernia, Cattedra di Gastroenterologia 1, Dipartimento di Scienze Cliniche, Università di Roma "La Sapienza", 00161 Roma, Italy

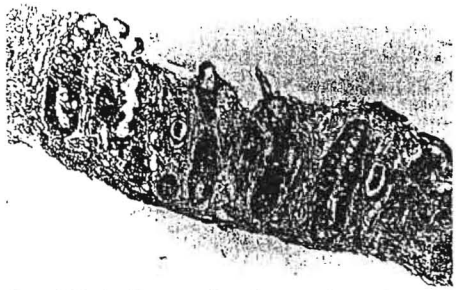


Figure 2: **Rectal biopsy specimen from a patient with acute radiation proctitis**

Three crypts show necrotic epithelium and invasion by neutrophils and eosinophils with formation of crypt abscesses. Focal mucin depletion of crypt epithelium and damage in the superficial epithelium are also visible. Chronic inflammatory infiltrate in the lamina propria is negligible.

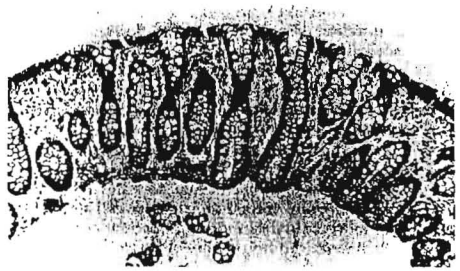


Figure 3: **Rectal biopsy specimen from same patient as figure 2 after treatment**

After a cycle of topical butyrate, appearance of mucosa has returned to normal.

oedema, erythema, and erosions. Butyrate led to improvement or remission in all but one patient, whereas only four patients on placebo improved, and one worsened.

Superficial epithelial damage and inflammation were negligible before treatment in both study groups. Again, improvement was observed in most patients after butyrate, but only in three patients after saline (figures 2 and 3).

Compliance was good because no participant returned more than two unused enemas at the end of treatment. Assessment of the data showed that the advantage of butyrate over placebo and the carryover effect were so marked that the odds ratio, in some instances, could not be quantified, inasmuch as the odds ratio approached infinite values. An analysis was thus performed comparing only data from the first 3 weeks of treatment. The advantage of butyrate over placebo, expressed as 95% CI, odds ratio, and p value for clinical endoscopic and histological findings is reported in the table.

Discussion

Early radiation-induced symptoms are frequent in patients submitted to radiation therapy for cancers in the pelvic region and have a further negative effect on the quality of life in these patients. No proven effective therapy is, at present, available.

This double-blind placebo-controlled, crossover trial indicates that topical butyrate at a dose of 80 mmol/L (80 mL/day) is effective in ARP. Relapse of symptoms was observed in three patients after withdrawal of treatment, suggesting that the optimum course of butyrate is longer than 3 weeks. Pretreatment endoscopic and histological abnormalities were not distributed homogeneously, and were, in most instances, minor. Despite the possibility of sampling errors, the level of significance, after butyrate therapy, was reached for endoscopy and approached it for histological findings.

ARP symptoms, occurring during or soon after treatment, spontaneously resolve, in most instances, within weeks or a few months. The risk of developing chronic proctitis besides depending upon the total radiation dose, is greatly increased by concomitant pathological conditions, such as diabetes and vascular damage, or old age and the occurrence of early radiation toxicity.[1,8] Of these various factors, only ARP is potentially preventable, but preliminary findings were that putative drugs, such as 5-aminosalicylic acid, epidermal growth factor, and prostaglandins were not fully satisfactory.[15-17] The present study provides strong but preliminary evidence that sodium butyrate is effective in the treatment of ARP. Whether early administraiton of this SCFA also has an effect on late intestinal radiation damage remains to be established.

Finally, ARP represents a unique human model in which a noxious agent is applied, for a short time, on a previously normal mucosa. The efficacy of sodium butyrate in enhancing the healing of mucosal lesions and reducing related symptoms further supports the role of this agent in the treatment of ulcerative colitis and other mucosal diseases of the colon.[18,19]

Contributors

All investigators took part in the planning of the design of the study and were involved in writing and revising the paper. P Vernia conceived the study and was the principal trial investigator. V Casale, P L Fracasso, V Stigliano were responsible for recruitment, follow-up, and endoscopy of patients. P Pinnarò was responsible for the recruitment of patients and radiotherapy. A Marcheggiano was responsible for histological data. G Villotti blindly collected and elaborated data obtained from multiple sources. V Bagnardi was responsible for statistical data. R Caprilli contributed by conceiving the protocol and coordinating many aspects of the study.

References

1 Zimmermann FB, Feldmann HJ. Radiation proctitis. Clinical and pathological manifestations, therapy and prophylaxis of acute and late injurious effects of radiation on the rectal mucosa. *Strahlenther Onkol* 1998; 174: 85–89.

2 Nussbaum ML, Capana TJ, Weese JL. Radiation-induced intestinal injury. *Clin Plast Surg* 1993; 20: 573–80.

3 Babb RR. Radiation proctitis: a review. *Am J Gastroenterol* 1996; 91: 1309–11.

4 Miller AR, Martenson JA, Nelson H, et al. The incidence and clinical consequences of treatment-related bowel injury. *Int J Radiat Oncol Biol Phys* 1999; 43: 817–25.

5 Kochhar R, Sriram PV, Sharma SC, Goel RC, Patel F. Natural history of late radiation proctosigmoiditis treated with topical sucralphate suspension. *Dig Dis Sci* 1999; 44: 973–78.

6 Casale V, Fracasso P, Grassi A, Lapenta R, Stigliano V. Sucralfate versus hydrocortisone enemas in the treatment of radiation-induced proctitis. *Ital J Gastroenterol* 1993; 25: 41.

7 Dubray BM, Thames HD. Chronic radiation damage in the rat rectum: an analysis of the fractionation, time and volume. *Radiother Oncol* 1994; 33: 41–47.

8 Wang CJ, Leung SW, Chen HC, et al. The correlation of acute toxicity and late rectal injury in radiotherapy for cervical carcinoma: evidence suggestive of consequential late effect (CQLE). *Int J Radiat Biol Phys* 1998; 40: 85–91.

9 D'Argenio G, Cosenza V, Sorrentini I, et al. Butyrate, mesalamine, and factor XIII in experimental colitis in the rat: effects on transglutaminase activity. *Gastroenterology* 1994; 106: 399–404.

10 Mortensen FV, Nielsen H, Mulvany MJ, et al. Short chain fatty acids dilate isolated human colonic resistance arteries. *Gut* 1990; 31: 1391–94.

11 Kvietys PR, Granger DN. Effect of volatile fatty acids on blood flow and oxygen uptake by the dog colon. *Gastroenterology* 1981; 80: 962–69.

12 Sakata T. Effects of short-chain fatty acids on epithelial cells proliferation and mucus release in the intestine. In: Roche AF, ed Short-chain fatty acids: metabolism and clinical importance. Report of the 10th Ross Conference on Medical Research. Columbus, Ohio: Ross Laboratories, 1991: 63–67.

13 al-Sabbagh R, Sinicrope FA, Sellin JH, Shen Y, Roubein L. Evaluatinn of short-chain fatty acid enemas: treatment of radiatinn proctitis. *Am J Gastroenterol* 1996; 91: 1814–16.

14 Senn S. Cross-over trials in clinical research. New York: John Wiley and Sons, 1993.

15 Baughan CA, Canney PA, Buchanan RB, et al. A randomised trial to assess the efficacy of 5-ASA for the prevention of radiation enteritis. *Clin Oncol* 1993; 5: 19–24.

16 McKenna KJ, Ligato S, Kauffman GL, et al. Epidermal growth factor enhances intestinal mytotic activity and DNA content after acute abdominal radiation. *Surgery* 1994; 115: 626–32.

17 Delaney JP, Bonsac ME. Felemivicus I. Misoprostol in the intestinal lumen protects against radiation injury of the mucosa of the small bowel. *Radiat Res* 1994; 137: 405–09.

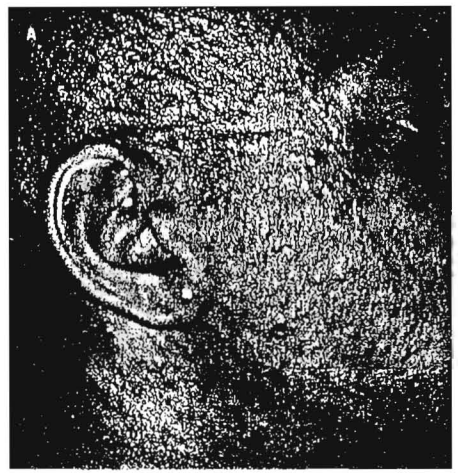
18 Vernia P, Marcheggiano A, Caprilli R, Frieri G, Corrao G, et al. Short-chain fatty acid topical treatment in distal ulcerative colitis. *Aliment Pharmacol Ther* 1995; 9: 309–13.

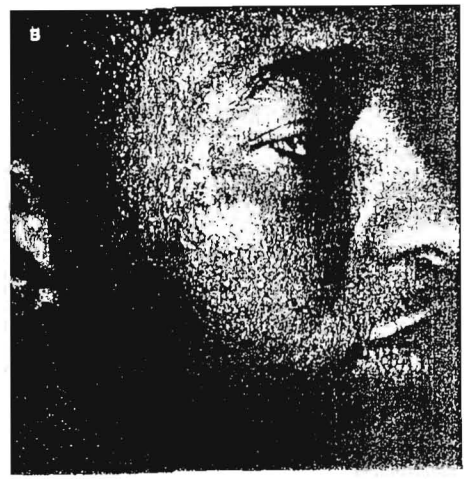
19 Kim YI. Short-chain fatty acids in ulcerative colitis. *Nutr Rev* 1998; 56: 17–24.

Clinical picture

Eosinophilic pustular folliculitis

Scott Dale, James Shaw





A 29-year-old HIV-negative man presented with a 1-year history of a widespread pruritic eruption. Unsuccessful prior treatments had included isotretinoin, fluocinonide gel, indomethacin, erythromycin, dicloxacillin, fluconazole, loratidine, cetirizine, doxepin, and ultraviolet B phototherapy. Physical examination revealed clusters of erythematous follicular papules and pustules primarily on the head, neck, and upper body (figure A). A biopsy specimen showed a mixed dermal perifollicular infiltrate of lymphocytes, eosinophils, and histiocytes with necrosis of follicular epithelium confirming the diagnosis of esoinophilic pustular folliculitis (Ofuji's disease). After unsuccessful trials with isotretinoin, antibiotics, ciclosporin, and dapsone, 0·1% tacrolimus ointment (topical FK-506) applied twice daily resulted in near complete resolution of lesions and pruritus in 10 days (figure B). During an inadvertant 5-day period without treatment, all lesions recurred, but resolved rapidly upon resumption of treatment. Tacrolimus ointment, by inhibiting interleukin-2 formation and T-cell activation, is a potent immunosuppressant with potential for topical use in Ofuji's disease.

Section of Dermatology, University of Chicago Hospitals, Chicago, IL 60637, USA (S Dale MD, J Shaw MD)

Early report

Comparison of pharmacokinetics and systemic effects of inhaled fluticasone propionate in patients with asthma and healthy volunteers: a randomised crossover study (165)[421]

Martin H Brutsche, Ingrid Carlen Brutsche, Mohamed Munavvar, Stephen J Langley, Catherine M Masterson, Peter T Daley-Yates, Ronan Brown, Adnan Custovic, Ashley Woodcock

Summary

Background Inhaled corticosteroids are currently the cornerstone of asthma treatment. Some studies of high-dose fluticasone propionate in patients with no or mild asthma have, however, suggested substantial systemic absorption. We investigated the pharmacokinetics of fluticasone propionate in patients with asthma receiving appropriate doses for severity.

Methods We did a double-blind, randomised, crossover study in 11 patients with asthma and 13 matched healthy controls (age 20–65 years; asthma patients forced expiratory volume in 1 s <75% and stable on high-dose inhaled corticosteroids). Patients received one 1000 μg intravenous dose or 1000 μg daily for 7 days inhaled (via spacer device) fluticasone propionate. In the 12 h after dosing, we monitored plasma fluticasone propionate and cortisol concentrations by mass spectrometry and competitive immunoassay with use of direct chemiluminescence. Analysis was by intention to treat.

Findings After inhalation, geometric mean values were significantly lower in the asthma group than in controls for fluticasone propionate plasma area under curve (1082 [95% CI 850–1451] vs 2815 pg mL^{-1} h^{-1} [2262–3949], −62% difference [45–72]; p<0·001), maximum concentrations (117 [91–159] vs 383 pg/mL [302–546], −68% [−50 to −81]; p<0·001), and systemic bioavailability (10·1 [7·9–14·0] vs 21·4% [15·4–32·2], −54% [−27 to −70]; p=0·001). Intravenous-dose clearance, volume of distribution at steady state, plasma half-life, and mean residence time, were similar in the two groups. Less suppression of plasma cortisol concentrations was seen in the asthma group than in controls 4–12 h after inhaled fluticasone propionate.

Interpretation Systemic availability of fluticasone propionate is substantially less in patients with moderate to severe asthma than in healthy controls. Inhaled corticosteroids that are absorbed through the lungs need to be assessed in patients who are receiving doses appropriate for disease severity, and not in normal volunteers.

Lancet 2000; **356:** 556–61
See *Commentary* page 527

North West Lung Centre, Wythenshawe Hospital, Manchester M23 9LT, UK (M H Brutsche MD, I Carlen Brutsche PhD, M Munavvar MB, S J Langley MB, C M Masterson RGN, A Custovic MD, Prof A Woodcock MD); **Glaxo Wellcome R&D, Stockley Park West, Uxbridge, Middlesex, UK** (R Brown PhD); **and Clinical Pharmacology, Glaxo Wellcome R&D, Greenford, UK** (P T Daley-Yates PhD)

Correspondence to: Prof Ashley Woodcock
(e-mail: ashley@nwlung.u-net.com)

Introduction

Since their introduction almost 30 years ago, inhaled corticosteroids have played a central part in the treatment of asthma. These drugs have almost entirely replaced maintenance oral corticosteroids, which had major adverse effects. Inhaled corticosteroids, because of topical application, have a substantially better therapeutic index than oral steroids.[1,2] Further improvements have been seen because of reduced oral bioavailability for newer inhaled coricosteroids. About 20% of the total inhaled dose from most metered-dose inhalers is deposited in the lungs and 80% stays in the oropharynx and is swallowed.[3] Molecules with a high hepatic first-pass metabolism and low oral bioavailability, such as fluticasone propionate, therefore, have lower systemic exposure than other inhaled corticosteroids.[4–6] For fluticasone propionate, any systemic activity results from absorption of the drug deposited in the lungs and its oral bioavailability is negligible.

Fluticasone propionate has a high therapeutic index and efficacy.[7,8] The drug has been used successfully for several years for all severities of asthma and has proved to be well tolerated. No clinically important systemic effects are reported for the normal therapeutic dose range.[7–10] By contrast, pharmacokinetic studies have suggested hypothalamic-pituitary-adrenal suppression with higher doses. However, those studies involved normal volunteers[5,11–15] or patients who had mild asthma and were receiving inappropriately high doses, well in excess of those needed to control their disease.[16–18] In patients with moderate or severe asthma requiring higher doses of inhaled corticosteroids, factors such as airflow obstruction and ventilation-perfusion mismatch could alter drug deposition in the lung and change systemic absorption.

To clarify the safety of higher doses of fluticasone propionate in asthma, we studied the pharmacokinetics and pharmacodynamics of the drug in patients with moderately severe asthma compared with normal controls in a randomised double-blind, double-dummy, crossover design (figure 1).

Methods

Study population

We recruited individuals from outpatient clinics at the North West Lung Centre, who had physician-diagnosed asthma, gave written informed consent, and were aged between 20 years and 65 years. The inclusion criteria included forced expiratory volume in 1 s (FEV_1) lower than 75% at screening, previous bronchodilator use, and a stable condition on high-dose inhaled corticosteroids (beclomethasone dipropionate [BDP] 2000 μg/day or budesonide 1600 μg/day). For each patient, we selected

healthy volunteers (generally staff and relatives of staff at the North West Lung Centre), matched for sex, age, and body-mass index. All participants were non-smokers (≥6 months). Exclusion criteria were: clinically important disease, systemic disease other than asthma, or both, pregnancy or lactation in women, suspected hypersensitivity to inhaled corticosteroids, treatment with oral or parenteral corticosteroids in the past 6 weeks, or inhaled fluticasone propionate in the past 2 months.

Study design

Participants were trained in an optimum inhalation technique before entering a 1-week run-in period. They were randomly assigned 1000 μg fluticasone propionate (500 μg twice daily), or a therapeutically equivalent dose of 2000 μg of BDP daily (1000 μg twice daily), inhaled at 0800 h and 2000 h from metered-dose inhalers with use of spacer devices (Volumatic, Glaxo Wellcome, UK; figure 1).

We assessed adherence by weighing the aerosol canisters before and after each treatment period. On the day before the pharmacokinetic sampling, participants were telephoned at 2000 h to remind them to take the last inhalation at the precise time. We asked participants to refrain from strenuous exercise. The next day all participants attended the study-day visit at 0700 h after fasting overnight and abstaining from alcohol for 24 h. We measured baseline spirometry and blood pressure. FEV_1 had to be within 15% of the value obtained during the screening visit. Participants had a standard breakfast 30 min before drug administration, and ate a standard diet for the rest of the day. A venous catheter was inserted into a forearm vein of each arm in all participants. At 0800 h, we administered 1000 μg inhaled fluticasone propionate and intravenous placebo (sodium chloride 0·9%) or 2000 μg inhaled BDP and 1000 μg intravenous fluticasone propionate (in 2 mL of propylene glycol). The methods of administration were identical for intravenous treatment and placebo or inhaled treatment and BDP. The infusion was administered over 10 min with the aid of a syringe driver. Inhalation treatment was taken in a sitting position. During the first hour after dosing, participants remained in bed. We took venous blood samples at baseline and 10 min, 20 min, 30 min, and 45 min, and at hours 1, 2, 3, 4, 6, 8, 10, and 12 after dosing to measure plasma concentrations of fluticasone propionate and cortisol. For the next week, participants crossed over to the other run-in treatment, after which they returned for a second pharmacokinetic-sampling day.

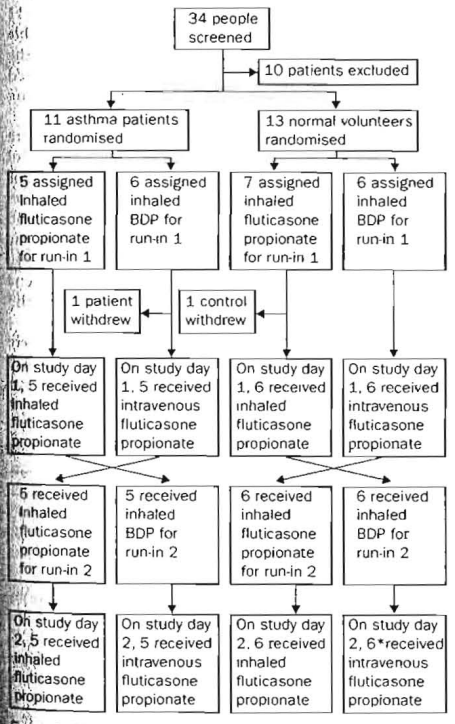


Figure 1: **Trial profile**
*Includes patient with extravasation of intravenous dose (data not included in analysis).

Characteristics	Controls (n=13)	Asthma group (n=11)
Demography		
Age (years)	42 (12)	45 (10)
Sex (M/F)	5 (38%)/8 (62%)	4 (36%)/7 (64%)
Height (m)	1·70 (0·09)	1·69 (0·10)
Weight (kg)	77 (15)	84 (14)
Body mass index (kg/m^2)	26·4 (3·7)	29·7 (7·2)
Spirometry		
FEV_1 (L/s)	3·52 (0·80)	1·73 (0·66)
FEV_1 (% predicted)	108 (13)	54 (16)
FVC (L)	4·43 (1·16)	3·25 (1·12)
FVC (% predicted)	111 (13)	83 (19)
FEV_1/FVC (%)	80 (5)	53 (9)
Total lung capacity (L)	6·06 (1·33)	6·23 (1·31)
Total lung capacity (% predicted)	103 (10)	107 (10)
Residual volume (% predicted)	95 (11)	161 (36)
R_{aw} (kPaL^{-1} s^{-1})	0·18 (0·04)	0·52 (0·22)
G_{aw} (kPa/s)	1·50 (0·29)	0·49 (0·29)
NO_{aw} (parts per billion)	5·7 (2·9)	6·1 (2·6)
Gas transfer		
DLCO (% predicted)	96 (11)	97 (14)
KCO (% predicted)	97 (15)	104 (20)

FVC=forced expiratory vital capacity. Data are mean (SD) except where indicated otherwise.

Table 1: **Baseline characteristics**

All static and dynamic pulmonary-function tests were measured with a VMAX22 spirometer (Sensor Medics BV, Bilthoven, Netherlands) and a body box (Autobox 6200 DL, Sensor Medics BV), according to ATS recommendations. Carbon-monoxide transfer factor was assessed with Transfer Test (Morgan Ltd, Chatham, Kent, UK).

In addition to establishing plasma cortisol profiles on sampling days, we measured plasma cortisol concentrations at 0800 h and 24 h urinary cortisol concentrations at the screening visit and on the two sampling days (ie, at a steady state for fluticasone propionate 500 μg and BDP 1000 μg).

All blood samples were drawn into heparinised tubes. They were immediately placed on ice and centrifuged within 30 min at 1500 rpm for 10 min at 4°C. Plasma and urine samples were immediately frozen at −70°C until assay. Masked analysis of plasma and urine samples was done at the Department of International Bioanalysis, Glaxo Wellcome Research and Development, Ware, UK. Fluticasone propionate was isolated from plasma by solid-phase extraction liquid chromatography—tandem mass spectrometry (LC-MS-MS) that uses thermally and

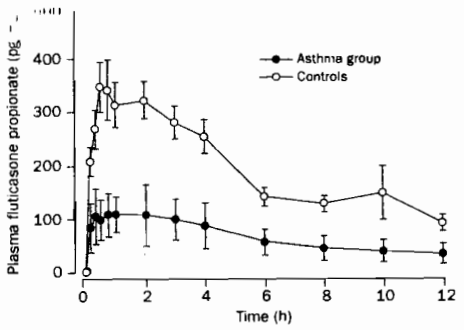


Figure 2: **Area under curve for plasma fluticasone propionate concentration after inhalation**
Group mean values (SE).

pneumatically assisted electrospray ionisation.[19] The assay has a lower limit of detection of 20 pg/mL, with between-assay and within-assay coefficients of variance at less than 6%. Plasma and urinary cortisol concentrations were measured by a competitive immunoassay with use of direct chemiluminescence (Product 672303 and ACS:180SE, Chiron Diagnostics, Harefield, Middlesex, UK). For this assay, the lower limit of detection is 6 nmol/L and the coefficient of variance is less than 3%.

We did the study according to the 1995 Declaration of Helsinki, and the design was approved by the local ethics committee.

Statistical analysis

We calculated sample size based on the main outcome variable, the area under the curve for plasma fluticasone propionate concentrations. According to previous studies, we assumed a within-participant variability of 10%. To detect a significant difference of 15% with a statistical power of 90%, we would need to enrol ten participants in each group.

We analysed data by intention to treat. Pharmacokinetic data were analysed by a conventional non-compartmental approach with WinNonLin Pro software (version 1.5).[20] Systemic availability was calculated, with reference to the nominal dose, as the area under the plasma concentration time curve after inhalation divided by the area under the curve after intravenous administration. Plasma clearance and apparent volume of distribution at steady-state were calculated by conventional equations. We analysed 24 h urinary cortisol concentrations as a cortisol to creatinine ratio and as total free cortisol excretion. Pharmacokinetic and cortisol data were log-normalised for group comparisons and are presented as geometric means with 95% CI. The lung-function variables followed a normal distribution and are presented as arithmetic mean (SD). We compared the results for asthma and control groups by two-tailed Student's *t* test. One-way and multiple ANOVA was applied to identify significant differences between the three different cortisol time points—screening, and sampling days 1 and 2—and the plasma cortisol profiles on the kinetic sampling days. Non-parametric data were compared by Mann-Whitney *U* and 2×2 χ^2 tests. We used bivariate correlations with Pearson's correlation coefficient for the systemic availability and the area under the curve for fluticasone propionate after inhalation against FEV_1, carbon monoxide transfer, and body-mass index. We set significance at 0·05. Statistical analyses were performed with SPSS software (version 7.5).

Results

Of 18 patients with asthma and 16 healthy volunteers screened, 11 and 13, respectively, entered the study (figure 1). One patient with asthma did not fulfil the lung-function entry criteria on the first sampling day and did not continue. One control attended the sampling day for inhaled fluticasone propionate, but refused to attend for the intravenous study day. One control had an extravasation at the intravenous site and the data were omitted from analysis. Ten patients with asthma and 11 controls completed the study, and had full data available for analysis. Baseline characteristics were similar in the two groups (table 1). All treatments were well tolerated and no serious adverse events occurred at any time during the study.

The pharmacokinetics of fluticasone propionate differed significantly between the asthma and control groups, seen in plasma area-under-curve values for inhaled fluticasone proprionate (1082 [850–1451] *vs* 2815 pg mL^{-1} h^{-1} [2262–3949], mean difference 0·38 [95% CI 0·28–0·55]; p<0·001; figure 2), systemic availability (10·1 [7·9–14·0] *vs* 21·4% [15·4–32·2], 0·46 [0·30–0·73]; p=0·001), and maximum fluticasone propionate concentration (117 [91–159] *vs* 383 [302–546] pg/mL, 0·31 [0·19–0·50]; p<0·001).

All intravenous pharmacokinetic parameters were similar in the two groups (clearance, time at which maximum fluticasone propionate concentration was reached, area under curve, volume of distribution at steady state, fluticasone propionate plasma half-life, and mean residence time, table 2).

	Controls (n=11)	Asthma group (n=10)	p
Pharmacokinetic data			
Inhaled area under curve (pg mL^{-1} h^{-1})	2815 (2262–3949)	1082 (850–1451)	<0·001
C_{max} (pg/mL)	383 (302–546)	117 (91–159)	<0·001
Systemic availability	21·4 (15·4–32·2)	10·1 (7·9–14·0)	0·001
Intravenous area under curve (pg mL^{-1} h^{-1})	12 357 (10 048–15 970)	10 731 (9120–12 893)	0·29
Volume of distribution at steady state (L)	253 (181–387)	282 (181–456)	0·63
Time to C_{max} (h)	1·4 (0·6–3·7)	1·0 (0·7–1·6)	0·24
Plasma half-life (h)	5·6 (4·8–6·7)	6·1 (4·1–9·9)	0·65
Mean residence time (h)	3·13 (2·69–3·73)	3·02 (2·28–4·15)	0·79
Clearance (mL/min)	1349 (1103–1735)	1553 (1319–1868)	0·29
Mean (SD) fluticasone propionate intravenous dose (µg)	1021 (23)	1018 (41)	0·86
Adherence			
Mean (SD) fluticasone propionate per day (g)	0·313 (0·046)	0·305 (0·049)	0·68
Mean (SD) BDP per day (g)	0·671 (0·042)	0·694 (0·170)	0·62

C_{max}=maximum plasma concentration of fluticasone propionate after inhalation. Data are geometric mean (95% CI) except where otherwise indicated.

Table 2: **Inhaled and intravenous pharmacokinetic parameters**

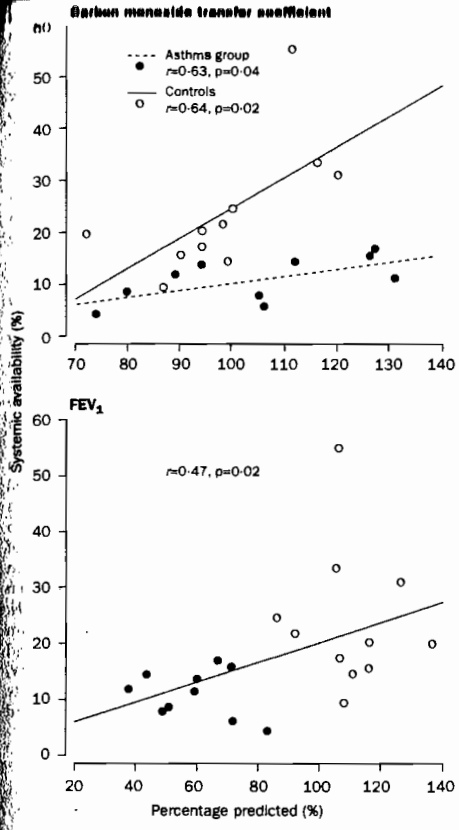


Figure 3: **Scatter plot of the systemic availability with different baseline variables**
Linear-regression line and *r* are presented for significant Pearson's correlations.

The area-under-curve and systemic availability values were positively correlated with baseline transfer of carbon monoxide (figure 3). Variability in systemic exposure within groups was explained partly by carbon monoxide transfer, but not the difference between groups. Baseline FEV_1 did not correlate with pharmacokinetic parameters (figure 3).

Controls had higher 24 h urinary cortisol concentrations at screening than after steady-state fluticasone propionate and BDP, measured at the start of the two sampling days, compared with the asthma group at any time point, as would be expected (table 3). The plasma cortisol concentrations at 0800 h changed little at all time points and were similar in the two groups. Suppression of cortisol concentrations was lower in the asthma group than in the control group at all time points after inhaled fluticasone propionate (estimated mean group difference −19·73 [95% CI −41·26 to 2·16], p=0·076); the difference between groups was significant at 4–12 h after dosing (estimated mean group difference −31·6 [−54·02 to −9·18], p=0·006; figure 4).

	Controls (n=11)	Asthma group (n=10)	p
24 h urinary cortisol			
Baseline corrected for creatinine (nmol/mmol)	7·88 (6·24–11·10)	3·22 (1·05–6·19)	0·01
Fluticasone propionate (500 µg twice daily)	3·65 (2·15–5·71)	4·01 (1·49–6·18)	0·82
BDP (1000 µg twice daily)	4·20 (2·49–6·18)	3·66 (0·97–5·02)	0·43
Within-group p	0·01*	0·83*	0·01†
Baseline total excretion (nmol/24 h)	63·84 (53·90–89·18)	24·89 (1·80–64·55)	0·15
Fluticasone propionate (1000 µg twice daily)	47·48 (28·55–72·91)	43·75 (13·98–74·82)	0·90
BDP (2000 µg twice daily)	53·17 (24·51–97·01)	37·15 (7·19–56·83)	0·39
Within-group p	0·79*	0·76*	0·37†
Plasma cortisol (nmol/L)			
Baseline	379 (330–462)	303 (216–468)	0·20
Fluticasone propionate (1000 µg twice daily)	363 (311–432)	323 (237–476)	0·43
BDP (2000 µg twice daily)	335 (280–418)	326 (256–438)	0·85
Within-group p	0·70*	0·98*	0·85†

Data are geometric mean (95% CI). *One-way ANOVA. †Multifactorial ANOVA.

Table 3: **Results of cortisol measurements as measure of effects on the hypothalamic-pituitary-adrenal axis function**

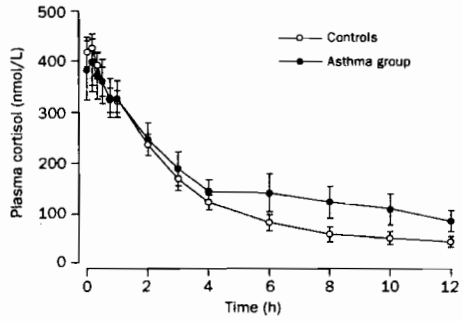


Figure 4: **Area under curve for plasma cortisol concentrations after inhalation of fluticasone propionate**
Group mean values (SE).

Discussion

The systemic availability of inhaled fluticasone propionate was more than halved in asthma patients compared with healthy controls. As a consequence, less hypothalamic-pituitary-adrenal suppression was seen after high doses at steady-state.

Investigations into the potential systemic effects of inhaled corticosteroids have concentrated on four main areas: hypothalamic-pituitary-adrenal-axis function, bone density, growth in children, and cataracts. Several confounding factors make a precise assessment of the risk-benefit ratio difficult, especially for high doses. First, significant effects on short-term indices are not necessarily clinically relevant (eg, knemometry might not be related to long-term growth). Second, chronic asthma might have some adverse effects directly, such as on growth, or indirectly, such as on bone density, through restriction of exercise. Third, patients with more severe asthma taking higher doses of inhaled corticosteroids frequently need rescue treatment with oral steroids, and the impact of this approach is unknown. If there were minor long-term effects of very-high-dose inhaled coricosteroids, the most severe patients would, however, still benefit more by taking this treatment than by taking maintenance oral steroids. Fourth, group data might hide substantial variation between patients in the degree of systemic availability, which has yet to be explained.

Long-term studies on potential systemic effects of inhaled corticosteroids have been generally reassuring.[21] Short-term effects on bone metabolism, especially on osteocalcin concentrations, are seen with these drugs,[22,23] but cross-sectional studies on bone density suggest no effect or only small changes. Such studies are, however, confounded by the use of rescue steroids.[24,25] A longitudinal study on bone density in patients with moderate or severe asthma taking 1000 μg fluticasone proprionate daily for 2 years showed no adverse effect, assessed by computed tomography, on trabecular bone density, which remained almost twice that of patients taking maintenance oral steroids.[26]

All inhaled corticosteroids, if given in high enough doses, lead to hypothalamic-pituitary-adrenal suppression. No thresholds have yet been proposed to define high risk and low risk of excessive metabolic effects. We saw less hypothalamic-pituitary-adrenal suppression in the asthma group than in the control group, which parallels reduced systemic bioavailability. These findings underscore pharmacokinetic differences between patients with asthma and healthy controls, and the related impact on the hypothalamic-pituitary-adrenal axis. Baseline urinary cortisol excretion among controls was significantly higher than in the asthma group, which suggests some suppression of the hypothalamic-pituitary-adrenal axis after the high doses of fluticasone, budesonide, and beclomethasone administered in the run-in period.

Our results agree with other comparisons of the systemic effects of fluticasone propionate and other inhaled corticosteroids, which have been done at different doses in healthy individuals and patients with mild asthma. In a meta-analysis, the systemic effects of fluticasone propionate and budesonide were compared, and results differed between patients with asthma and healthy people.[27] At roughly equal doses, in healthy volunteers, fluticasone propionate increases suppression of cortisol concentrations more than budesonide (budesonide/fluticasone prioprionate suppression ratio 3·3 for the residual cortisol concentration at the end of treatment). In patients with asthma, however, fluticasone priopionate and budesonide have an equal effect on the hypothalamic-pituitary-adrenal axis (budesonide/fluticasone propionate suppression ratio 1·0) for equivalent doses.

There is little information on factors that influence the systemic bioavailability of inhaled corticosteroids, especially for drugs with minimum oral bioavailability, such as fluticasone propionate, in which drug delivery, and pulmonary deposition, have key roles. Almost all fluticasone propionate present in the systemic circulation has been absorbed in an unchanged active form via the lungs.[28] Once present in the bloodstream, this drug remains potent, with a high binding affinity to the corticosteroid receptor. Until now, despite the characteristic differences in the airways in asthma, the assumption had been made that drug deposition, pulmonary drug absorption, and systemic effects measured in healthy individuals can predict the outcomes in people with asthma. Asthma is, however, characterised by reversible and non-homogeneous airflow obstruction, leading to ventilation-perfusion mismatch. Several studies, some by use of three-dimensional imaging techniques,[29] have shown that the uniformity of deposition in the lungs is greater in healthy individuals than in patients with airway disease. It is suggested that the narrowing of airways in asthma results in less penetration of the drug particles and, consequently more central-airway deposition.[30] Drug particles, which deposit on the airways, are more prone to clearance by mucociliary action than those that deposit in alveoli, which will be totally absorbed. Therefore, for a drug with a relatively slow dissolution rate in the lung, such as fluticasone propionate,[28] there is more potential for drug depositing in the airways of patients to be removed from the lung by mucociliary clearance and swallowed, thereby not giving rise to a similar degree of systemic exposure to that seen in healthy volunteers. For other drugs, the magnitude of the effect might be less and deserves further investigation.

The positive correlation between the carbon monoxide transfer coefficient and the systemic bioavailability we saw suggests that ventilation-perfusion mismatch is important. This phenomenon was seen in the absence of pathologically low transfer factors in either group. Given the potential differences in the pattern of deposition, the role of carbon monoxide might be confounded by differences in mucociliary clearance between patients with asthma and normal individuals.

Relevant pharmacokinetic and pharmacodynamic data on inhaled corticosteroids with minimum oral bioavailability can be derived only when appropriate doses are given for severity of asthma. Studies are needed to define the factors that influence pulmonary absorption of inhaled corticosteroids in patients with lung disease. In severe asthma, even higher inhaled doses might be given safely, and with more safety than oral corticosteroids. Patients with asthma should be managed on the minimum dose of inhaled corticosteroids to control their disease and this should be regularly reviewed. Combination therapy, such as with a long-acting β-agonist, might be appropriate in some patients rather than increasing the dose of inhaled corticosteroids.

Contributors

Martin Brutsche was involved in study concept, design, execution, analysis, and in writing the paper. Ingrid Carlen Brutsche, Mohamed Munnaver, Stephen Langley, and Catherine Masterson were involved in study design and execution. Peter Daley-Yates and Ronan Brown were involved in study concept design and analysis. Adnan Custovic was involved in study design and analysis and writing the paper. Ashley Woodcock was the principal investigator and was involved in study concept and design, analysis, and writing of the paper.

Acknowledgments

We thank Andrew Williams, Naomi Houghton, Bridget Simpson, and the nurses of the Clinical Research Unit; David Richards and John Efthimiou from Glaxo Wellcome R&D, Uxbridge; Sheryl Callejas and Helen Billings for fluticasone and cortisol assays (Glaxo Wellcome, Department of International Bioanalysis, Ware, UK); and John Pritchard (Glaxo Wellcome, Respiratory Therapeutic Development, Stockley Park, UK). The study was supported by Glaxo Wellcome R&D International (FAS40022). ICB was sponsored by the Swiss National Science Foundation, Uarda Frutiger Foundation.and the Novartis Foundation.

References

1 Barnes PJ, Pedersen S. Efficacy and safety of inhaled corticosteroids in asthma. *Am Rev Respir Dis* 1993; **148**: S1–26.

2 Lipworth BJ. Pharmacokinetics of inhaled drugs. *Clin Pharmacol* 1996; **42**: 697–705.

3 Borgström L. Methodological studies on lung depostion: evaluation of inhaler devices and absorption mechanisms. In: Comprehensive summaries form the Uppsala Faculty of Pharmacy. Uppsala: University of Uppsala, 1993: 105.

4 Derendorf H. Pharmacokinetic and pharmacodynamic properties of inhaled corticosteroids in relation to efficacy and safety. *Respir Med* 1997; **91** (Suppl A): 22–28.

5 Harding SM. The human pharmacology of fluticasone propionate. *Respir Med* 1990; **84** (Suppl A): 25–29.

6 Falcoz C, Mackie AE, McDowall J, et al. Oral bioavailability of fluticasone propionare in healthy subjects. *Br J Clin Pharmacol* 1996; **41**: 459–60.

7 Fabbri L, Burge PS, Croonenborgh L, et al. Comparison of fluticasone propionate with beclomethasone dipropionate in moderate to severe asthma treated for one year. *Thorax* 1993; **48**: 817–23.

8 Ayres JG, Bateman ED, Lundback B, Harris TA. High dose fluticasone propionate, 1 mg daily versus fluticasone propionate, 2 mg daily, or budesonide, 1.6 mg daily, in patients with chronic severe asthma. *Eur Respir J* 1995; **8**: 579–86.

9 Agertoft L, Pedersen S. A randomised, double-blind dose reduction study to compare the minimal effective dose of budesonide Turbohaler and fluticasone propionate Diskhaler. *J Allergy Clin Immunol* 1997; **99**: 773–80.

10 Yiallouros PK, Milner AD, Conway E, Honour JW. Adrenal function and high dose inhaled corticosteroids for asthma. *Arch Dis Child* 1997; **76**: 405–10.

11 Lönnebo A, Grahnén A, Jansson B, Brundin RM, Ling-Andersson A, Eckernäs SA. An assessment of the systemic effects of single and repeated doses of inhaled fluticasone propionate and inhaled budesonide in healthy volunteers. *Eur J Clin Pharmacol* 1996; **49**: 459–63.

12 Grahnén A, Eckernäs SA, Brundin RM, Ling-Andersson A. An assessment of the systemic activity of single doses of inhaled fluticasone propionate in healthy volunteers. *Br J Clin Pharmacol* 1994; **38**: 521–25.

13 Grahnén A, Jansson B, Brundin RM, et al. A dose-response study comparing suppression of plasma cortisol induced by fluticasone propionate from Diskhaler and budesonide from Turbohaler. *Eur J Clin Pharmacol* 1997; **52**: 261–67.

14 Thorsson L, Dahlström K, Edsbäcker A, Källén A, Paulson J, Wirén JE. Pharmacokinetics and systemic effects of inhaled fluticasone propionate in healthy subjects. *Br J Clin Pharmacol* 1997; **43**: 155–61.

15 Boorsma M, Andersson N, Larsson P, Ullman A. Assessment of the relative systemic potency of inhaled fluticasone and budesonide. *Eur Respir J* 1996; **9**: 1427–32.

16 Clark DJ, Grove A, Cargill RI, Lipworth BJ. Comparative adrenal suppression with inhaled budesonide and fluticasone propionate in adult asthmatic patients. *Thorax* 1996; **51**: 262–66.

17 Clark DJ, Lipworth BJ. Adrenal suppression with chronic dosing of fluticasone propionate compared with budesonide in adult asthmatic patients. *Thorax* 1997; **52**: 55–58.

18 Derom E, Van Schoor J, Vincken W, et al. A comparison of the systemic activity of two doses of inhaled fluticasone propionate and budesonide in asthmatic patients. *Am J Respir Crit Care Med* 1997; **155**: A355.

19 Callejas SL, Biddlecombe RA, Jones AE, Joyce KB, Pereira AI, Pleasance S. Determination of the glucocorticoid flutiasone propionate in plasma by automated solid-phase extraction and liquid chromatography-tandem mass spectrometry. *J Chromatogr B Biomed Sci Appl* 1998; **718**: 243–50.

20 Gibaldi M. Biopharmaceutics and Clinical Pharmacokinetics, 4th edn. London: Lea and Febiger, 1991.

21 Efthimiou J, Barnes PJ. Effect of inhaled corticosteroids on bones and growth. *Eur Respir J* 1998; **11**: 1167–77.

22 Wolthers OD, Hansen M, Juul A, Nielsen HK, Pedersen S. Knemometry, urine cortisol excretion, and measures of the insulin-like growth factor axis and collagen turnover in children treated with inhaled glucocorticosteroids. *Pediatr Res* 1997; **41**: 44–50.

23 Hodsman AB, Toogood JH, Jennings B, Fraher LJ, Baskerville JC. Differential effects of inhaled budesonide and oral prednisolone on serum osteocalcin. *J Clin Endocrinol Metab* 1991; **72**: 530–40.

24 Packe GE, Douglas JG, McDonald AF, Robbins SP, Reid DM. Bone density in asthmatic patients taking high dose inhaled beclomethasone dipropionate and intermittent systemic corticosteroids. *Thorax* 1992; **47**: 415–17.

25 Stead RJ, Horsman A, Cooke NJ, Belchetz P. Bone mineral density in women taking inhaled corticosteroids. *Thorax* 1990; **45**: 792.

26 Egan JJ, Kalra S, Adams J, Eastell R, Maden C, Woodcock A. A randomized double blind study comparing the effects of beclomethasone 2000mg/day versus fluticasone propionate 1000mg/day on bone density over 2 years. *Eur Respir J* 1999; **13**: 1267–75.

27 Hallett C. Corticosteroid treatment of asthma: now at the crossroads. *Respir Med* 1999; **93**: 292–94.

28 Mollman H, Wagner M, Meibohm B, et al. Pharmacokinetic and pharmacodynamic evaluation of fluticasone propionate after inhaled administration. *Eur J Clin Pharmacol* 1998; **53**: 459–67.

29 Yanai M, Hatazawa J, Ojima F, Sasaki H, Itoh M, Ido T. Deposition and clearance of inhaled 18FDG powder in patients with chronic obstructive pulmonary disease. *Eur Respir J* 1998; **11**: 1342–48.

30 Melchor R, Biddlecombe M, Mak V, Short M, Spiro S. Lung deposition patterns of directly labelled salbutamol in normal subjects and in patients with reversible airflow obstruction. *Thorax* 1993; **48**: 506–11.

Effect of consumption of red wine, spirits, and beer on serum homocysteine (419) [480]

Martijn S van der Gaag, Johan B Ubbink, Pekka Sillanaukee, Seppo Nikkari, Henk F J Hendriks

Serum homocysteine increases after moderate consumption of red wine and spirits, but not after moderate consumption of beer. Vitamin B_6 in beer seems to prevent the alcohol-induced rise in serum homocysteine.

Homocysteine concentrations are affected by lifestyle factors such as diet—eg, inadequate intake of B vitamins involved in the homocysteine breakdown. Alcoholics have raised homocysteine concentrations that are either caused by low vitamin B intake or chronic excessive alcohol consumption. Beer is a rich source of folate and vitamin B_6, whereas red wine and spirits contain negligible amounts of these vitamins. We postulated that moderate alcohol consumption could affect homocysteine metabolism, and that these effects are beverage specific. In a randomised, diet-controlled, crossover trial,[1] 11 healthy, non-smoking men (aged 44–59 years) who were moderate alcohol drinkers, consumed four glasses of red wine, beer, or spirits (Dutch gin), or sparkling mineral water (control) with dinner. Beverages were switched every 3 weeks in a randomised order, according to a Latin square design. All food and drink was supplied for 12 weeks. The diet, which was essentially the same during all four periods, contained adequate amounts of macronutrients and micronutrients. Alcohol intake equalled 40 g daily (with exception of the water period), which did not affect activities of the liver enzymes γ-glutamyltransferase, alanine aminotransferase, and aspartate aminotransferase. Treatment effects were assessed by analysis of variance, by use of general linear modelling, in which homocysteine measurements were log transformed. No carry-over effects were seen.

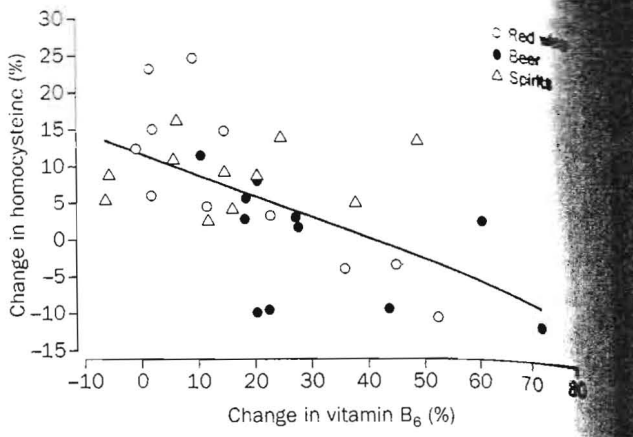
Homocysteine concentrations were raised after 3 weeks' consumption of red wine and spirits by 8% and 9%, respectively, as compared with water consumption, whereas no increase was recorded after beer consumption (table). Such an increase in homocysteine coincides with a 10–20% increase in cardiovascular disease (CVD) risk.[2] However, moderate alcohol consumption is associated with a lowered CVD risk. The cardioprotective effects of moderate drinking could exceed the increase in risk by higher homocysteine concentrations. Alternatively, slightly and transiently raised homocysteine concentrations might be beneficial. Homocysteine could act as a mediator of tissue repair and as a regulator of blood cells and cells of the vascular wall.[3]

Homocysteine concentrations could rise by inhibition of its two major breakdown pathways, both dependent on B-vitamins. The remethylation pathway depends on folate and vitamin B_{12}, whereas vitamin B_6 is essential in the breakdown via trans-sulphuration.[4] We assessed beverage specific effects on plasma values of these vitamins. No significant differences in vitamin B_{12} were reported. A 10% fall in folate occurred after spirits consumption only (table), and no correlation was found between changes in homocysteine values and changes in folate (Pearson correlation coefficient, p=0·99). Plasma vitamin B_6, analysed as pyridoxal-5′-phosphate, was increased after beer consumption by about 30%. Surprisingly, vitamin B_6 concentrations were also higher after intake of wine and spirits—17% and 15%, respectively (table). Changes in vitamin B_6 showed a significant inverse correlation with changes in homocysteine ($r=-0·58$, p=0·0004; figure), suggesting that vitamin B_6 might be a rate-limiting factor for homocysteine breakdown after moderate alcohol consumption. Interestingly, prospective data from the Atherosclerosis Risk in Communities study suggest that vitamin B_6 itself is inversely associated with CVD risk, independently of homocysteine.[5] So, the increase in plasma vitamin B_6, as seen after beer and to a lesser extent after red wine and spirits consumption might even contribute to a lower CVD risk.

	Red wine	Beer	Spirits	Water
Homocysteine (μmol/L)	14·2¶**	12·9††‖	14·2†‡	13·0††‖
Folate (nmol/L)	18·4‖	17·8	16·5¶§	18·3‡‡
Vitamin B_{12} (pmol/L)	223	218	209	232
Vitamin B_6 (nmol/L)	56·5†‡	62·9†§‖	55·6†‡	48·3§‡‖

*Median homocysteine concentrations in serum samples. Vitamins have been measured in plasma; different from †water, ‡beer, §red wine, ‖spirits (p<0·01), and from ¶water, **beer, ††red wine, ‡‡spirits (p<0·02).

Mean blood concentrations of homocysteine,* folate, vitamins B_6, and B_{12} after 3 weeks consumption of red wine, beer, spirits and water



Relation between changes in homocysteine and changes in vitamin B_6 concentrations

Individual changes were computed by subtracting outcome after 3 weeks' consumption of each alcoholic beverage from outcome after 3 weeks' water consumption and expressed as percentage of outcome after water consumption, per alcoholic beverage.

This study was funded by the Dutch Foundation for Alcohol Research. Martijn S van der Gaag died in January, 2000.

1 Van der Gaag MS, Van Tol A, Scheek LM, et al. Daily moderate alcohol consumption increases serum paraoxonase activity: a diet-controlled randomised intervention study in middle-aged men. *Atherosclerosis* 1999; **147:** 405–10.

2 Verhoef P, Stampfer MJ, Rimm EB. Folate and coronary heart disease. *Curr Opin Lipidol* 1998; **9:** 17–22.

3 Dudman NPB. An alternative view of homocysteine. *Lancet* 1999; **354:** 2072–74.

4 Hankey GJ, Eikelboom JW. Homocysteine and vascular disease. *Lancet* 1999; **354:** 407–13.

5 Folsom AR, Nieto FJ, McGovern PG, et al. Prospective study of coronary heart disease incidence in relation to fasting total homocysteine, related genetic polymorphisms, and B vitamins: the Atherosclerosis Risk in Communities (ARIC) study. *Circulation* 1998; **98:** 204–10.

Department of Nutritional Physiology, TNO Nutrition and Food Research Institute, PO Box 360, 3700 AJ Zeist, The Netherlands (M S van der Gaag MSc, H F J Hendriks PhD); **Department of Chemical Pathology, University of Pretoria, South Africa** (J B Ubbink DSc); **and Tampere University Medical School, Tampere, Finland** (P Sillanaukee PhD, S Nikkari PhD, MD)

Correspondence to: Dr Henk Hendriks
(e-mail: Hendriks@voeding.tno.nl)

investigation for a heterozygous effect was done on the variants Arg142His, Arg151Cys vs. Arg160Trp, Asp294His (which were chosen because they had been associated with red hair in previous case-control studies, and because they show a diminished ability to activate cyclic AMP in transfection assays,[4] and on an insertion at codon 179 (ins179), which results in a predicted premature stop codon and is likely to be non-functional. Other (pseudo wild-type) variants were treated as wild type. Detection of a heterozygote effect was sought by fitting a cumulative proportional odds model by Stata. This process automatically takes into account the ordering in the response. The effect of genotype was assessed by inclusion of genotype as a factor in the analysis, firstly as a linear trend from wild type through heterozygous to homozygous, and secondly as the remaining non-linear trend.

Individuals with one variant allele were intermediate with regard to skin type and the ability to tan after repeated sun exposure between those with two variant alleles and those with none of the variants (figure). Analysis for trend from zero to two variants was highly significant ($p<0·001$), for both depth of tanning and skin type, with little evidence of any non-linear trend ($p=0·844$ and $p=0·624$ for depth of tanning and skin type, respectively). The odds of having skin type below any given category were 3·82 times greater in heterozygotes than in individuals with wild-type alleles (95% CI 2·22–6·59), and 14·6 times greater in homozygotes or compound heterozygotes than in wild-type individuals (4·9–43·4). For tanning, the respective odds ratios were 4·5 for heterozygotes (2·54–7·99) and 20·3 for homozygotes or compound heterozygotes (6·4–63·9).

The identification of a dosage effect of *MC1R* variant alleles on sensitivity to ultraviolet radiation, and the large attributable risk for heterozygotes (28% of the study population were heterozygous for Arg142His, Arg151Cys, Arg160Trp, or Asp294His variant alleles) suggests that the *MC1R* gene is closely associated with variation in the skin's response to ultraviolet radiation in most of the population who do not have red hair. Furthermore, because *MC1R* heterozygotes are common, the results suggest that the *MC1R* gene is of substantial importance as a susceptibility gene for sunburn, photoageing, and skin cancer.

Pigmentation is a complex genetic trait, and our results have implications for studies on other complex genetic disease states. Because a large number of loci can affect pigmentation in mice, and a large number of rare Mendelian disorders affect pigmentation in man, a large number of loci may mistakenly be assumed to underlie "physiological" differences in pigmentation within human populations; in this scenario each locus would contribute, on average, a small fraction of the overall variance. However, for *MC1R*, the effect of one locus in man is substantial, and our results show that large risk ratios can be detected for common alleles where an adequate phenotypic classification of the disease or trait is available. Such variation is likely to outweigh the effect of rare Mendelian disorders on population attributable risk.

Eugene Healy is a Clinician Scientist of the UK Medical Research Council (MRC). This study was supported in part by the Leech Trust, CRC, and the MRC. Amanda Ray is a Leech postdoctoral fellow.

1 Gallagher RP, Ho VC. Environmental and Host Risk Factors. In: Grob JJ, Stern RS, MacKie RM, Weinstock MA, ed. Epidemiology, causes and prevention of skin diseases. London: Blackwell, 1998: 235–42.
2 Rees JL. Flanagan N. Pigmentation, melanocortins and red hair. *Q J Med* 1999; 92: 125–31.
3 Smith R, Healy E, Siddiqui S, Flanagan N, Steijlen PM, Rosdahl I, et al. Melanocortin 1 receptor variants in an Irish population. *J Invest Dermatol* 1998; 111: 119–22.
4 Schioth HB, Phillips S, Rudzish R, Birch-Machin M, Wikberg J, Rees JL. Loss of function mutations of the human melanocortin 1 receptor are common and associated with red hair. *Biochem Biophys Res Commun* 1999; 260: 488–91.

Departments of Dermatology (E Healy and N Flannagan) and Statistics (J N S Matthews PhD), University of Newcastle upon Tyne; Dermatopharmacology Unit, University of Southampton (E Healy); Department of Dermatology, University of Edinburgh (J L Rees); and Medical Research Council Human Genetics Unit, Edinburgh (I J Jackson PhD), UK

Correspondence to: Prof Jonathan L Rees, Department of Dermatology, University of Edinburgh, Royal Infirmary, Lauriston Place, Edinburgh EH3 9YW, UK
(e-mail: jonathan.rees@ed.ac.uk)

[4933]

Transcranial magnetic stimulation and auditory hallucinations in schizophrenia

Ralph E Hoffman, Nashaat N Boutros, Sylvia Hu, Robert M Berman, John H Krystal, Dennis S Charney

12 patients with schizophrenia and auditory hallucinations received 1 Hz transcranial magnetic stimulation of left temporoparietal cortex. In a double-blind crossover trial, active stimulation significantly reduced hallucinations relative to sham stimulation.

Auditory hallucinations are reported by 50–70% of patients with schizophrenia and generally consist of spoken speech or voices. Response to drug treatment is often incomplete or non-existent, and these hallucinations can cause great distress, functional disability, and lack of behavioural control.

Silbersweig and colleagues[1] described regional brain activation by use of ^{15}O positron emission tomography when auditory hallucinations occurred in six patients with schizophrenia. Blood flow activation was detected in left temporoparietal auditory-linguistic association cortex and in thalamic, hippocampal, and striatal regions. Low frequency (1 Hz), extended duration (15–30 min), repetitive transcranial magnetic stimulation (rTMS) reduces activation in the brain area directly stimulated as well as in other functionally connected brain areas.[2,3] We postulated that low frequency rTMS delivered to the left temporoparietal cortex would curtail auditory hallucinations by reduction of excitability of distributed neurocircuitry that produce these experiences.

12 right-handed patients with auditory hallucinations who met Diagnostic and Statistical Manual IV (DSM-IV) diagnostic criteria for schizophrenia (eight paranoid type, four schizoaffective type; ten men; mean age 41·8 years [SD 8·6]) were included. Education level of the participants in grades was mean 14·2 (SD 1·8); a level of 14 grades corresponds to 2 years of college. All patients received antipsychotic drugs and were maintained on these drugs without change in dose throughout the study period. Five patients received concomitant anticonvulsant drugs (four valproate semisodium, one carbamazepine). All patients had daily auditory hallucinations without remission for at least 6 months. Auditory hallucinations were either continuous (three) or intermittent (nine). Each patient had normal routine laboratory studies, electrocardiogram, and electroencephalogram.

Motor threshold was identified as the minimum magnetic field strength required to produce left thenar muscle activation by single transcranial magnetic stimulation pulses delivered to the motor cortex, confirmed by electromyographic monitoring, for at least four of eight trials. Site and strength of the motor threshold was redefined each session. 1 Hz stimulation at 80% motor threshold was then given midway between the left temporal (T_3) and left parietal (P_3) electroencephalogram electrode sites on the basis of the international 10–20 electrode placement system. Sham stimulation was given at the same location, strength, and frequency with the coil angled 45° away from the skull to induce scalp stimulation but curtailing brain stimulation. A Cadwell magnetic stimulator system (Cadwell Inc, Kennewick, WA, USA) with a water-cooled, handheld figure-eight coil was used to deliver rTMS. Stimulation was initiated at 4 min for each condition and built up on successive days by 4 min to 16 min on day 4. Psychiatric symptoms were assessed daily with the Positive and Negative Symptom Scale (PANSS).

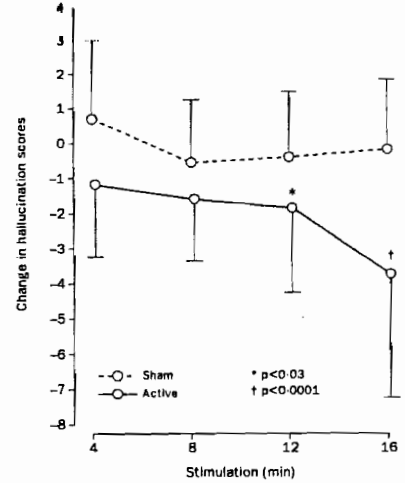


Figure 1: **Hallucination severity ratings and repetitive transcranial magnetic stimulation duration in active and sham stimulation trials**
Data analysed with a random effects model with Dunnett's criterion to adjust for multiple comparisons.

An earlier study showed that factors contributing to severity of auditory hallucinations varied between patients (for instance, frequency, loudness, content, number of voices, emotional distress, and level of distraction). Consequently, auditory hallucinations were assessed with an individualised, composite scale. A score of ten corresponded to a narrative description of the patient's hallucinations at the time of study entry, with zero corresponding to no hallucinations. For reassessments the patient produced a severity rating on the basis of these individualised anchor points. Higher scores were permitted if the patient's hallucination severity exceeded that at study initiation. Trials of active versus sham stimulation took place on separate weeks with 2–3 days separating each trial. Baseline hallucination assessments for each trial were done just before initiation of each stimulation condition and reflected the 24 h before. Reassessments took place the morning after each of the four rTMS sessions and indicated overall hallucination severity since the last rTMS session. Patients, clinical interviewers, and clinical staff were unaware of stimulation condition. Patients randomised to receive sham stimulation first received active stimulation the second week and viceversa.

Besides complaints of mild headache in two cases after active stimulation, patients tolerated rTMS without difficulty. Mean (SD) baseline hallucination score was 8·5 (2·2) for the active rTMS trial and 7·5 (2·6) for the sham rTMS trial. Symptom improvements relative to baseline were significant following 12 and 16 minutes of active stimulation but not for any duration of sham stimulation (figure 1). In all but one case, hallucination severity was lower after the active stimulation sequence than the sham stimulation sequence (figure 2). Endpoint hallucination ratings were analysed by use of a repeated measure ANOVA with two additional factors; order of stimulation (active or sham first), and concomitant treatment with anticonvulsant drugs. Reductions in hallucination severity after active compared with sham stimulation were significant ($p<0·006$), as was the interaction between change in hallucination severity and anticonvulsant drugs ($p<0·02$) showing reduced treatment effects with these drugs (figure 2). No effect of order of stimulation was seen.

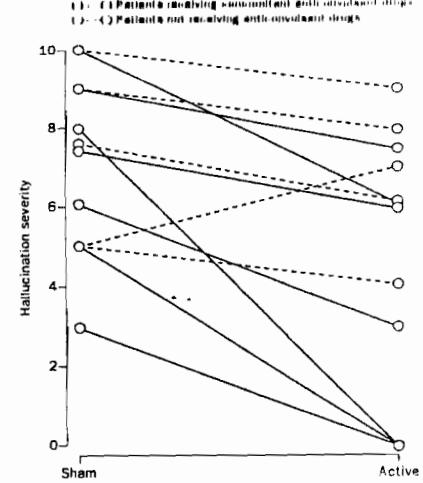


Figure 2: **Endpoint auditory hallucination severity for sham versus active repetitive transcranial magnetic stimulation**

Other positive symptoms and negative symptoms did not change much after rTMS. Follow-up assessments of the eight patients classified as responders (ie, hallucination severity improved for active relative to sham stimulation by score >1) indicated that auditory hallucinations returned roughly to baseline 1 day after the course of active rTMS in two patients, 4 days in two, 5 days in one, 2 weeks in one, 3 weeks, and 2 months in one. Left temporoparietal cortex, the site of rTMS in this study, is a brain area critical in perceiving spoken speech.[4] Our findings therefore support the hypothesis that speech perception neurocircuitry plays a part in the generation of hallucinated speech.

Not all patients showed robust improvements in hallucinations after active rTMS. One factor contributing to variable response was concurrent anticonvulsant drug treatment, which seemed to reduce rTMS effects. This observation suggests that higher levels of signal propagation are required for rTMS to curtail auditory hallucinations, or that symptoms prompting administration of anticonvulsant drugs (eg, mood lability) are negative predictors of rTMS response. Other factors that might contribute to the variability of rTMS effects include individual differences in anatomical location of speech processing areas,[5] variable location of cortical activation producing auditory hallucinations,[1] and differences in skull-brain relation, and baseline physiology.

Research supported by [illegible] grant [illegible], [illegible] program grant MH01118, a [illegible] Independent Investigator Award, and the Cheryle Fund

1 Silbersweig DA, Stern E, Frith C, et al. A functional neuroanatomy of hallucinations in schizophrenia. *Nature* 1995; 378: 176–79.
2 Siebner HR, Tormos JM, Ceballos-Baumann AO, et al. Low-frequency repetitive transcranial magnetic stimulation of the motor cortex in writer's cramp. *Neurology* 1999; 52: 529–37.
3 Wasserman EM, Wedegaertner FR, Ziemann U, et al. Crossed reduction of human motor cortex excitability by 1 Hz transcranial magnetic stimulation. *Neurosci Lett* 1998; 250: 141–44.
4 Fiez JA, Raichle ME, Balota DA, et al. PET activation of posterior temporal regions during auditory word presentation and verb generation. *Cereb Cortex* 1996; 6: 1–10.
5 Ojemann GA. Cortical organization of language. *J Neurosci* 1991; 11: 2281–87.

Department of Psychiatry, Yale University School of Medicine, Yale Psychiatric Institute, PO Box 208038, New Haven, CT 06520-8038, USA (R E Hoffman MD, N N Boutros MD, S Hu PhD, R M Berman MD, J H Krystal MD, D S Charney MD)

Correspondence to: Dr Ralph Hoffman
(e-mail: ralph.hoffman@yale.edu

Circulating plasma platinum more than 10 years after cisplatin treatment for testicular cancer

J A Gietema, M T Meinardi, J Messerschmidt, T Gelevert, F Alt, D R A Uges, D Th Sleijfer

We have shown in patients cured from metastatic testicular cancer that up to 20 years after administration of cisplatin-containing chemotherapy, circulating platinum is still detectable in plasma. This finding may influence the development of long-term, treatment-related side-effects.

More than two decades after the introduction of cisplatin-containing chemotherapy for metastatic testicular cancer, this treatment remains one of the few with a high curative potential in disseminated cancer. The success of this therapy has resulted in increased survival in patients who have had testicular cancer. Therefore, the long-term sequelae of cytostatic treatment, such as cardiovascular problems or secondary malignancies, is becoming increasingly relevant.[1,2] The pathogenesis of the long-term sequelae of cisplatin combination chemotherapy in patients with testicular cancer has not yet been fully elucidated but it has been associated with prolonged retention of platinum in the body.

During a long-term follow-up investigation in 61 testicular-cancer patients cured with cisplatin combination chemotherapy more than 10 years previously, we measured plasma platinum concentrations using a sensitive assay. The median age of the patients at the time of chemotherapy was 27 years (range 17–36 years). The median age at the time of follow-up investigation was 42 years (30–50 years), with a median follow-up of 14 years (10–20 years). Patients were treated with four courses of cisplatin, bleomycin, and vinblastine or etoposide, every 3 weeks. 17 patients additionally received maintenance therapy with vinblastine and cisplatin for a maximum of 1 year. The total amount of administered cisplatin per patient ranged from 350 to 950 mg/m² (663–1846 mg). The 44 patients without maintenance chemotherapy received a mean cisplatin dose of 400 mg/m² (SD 14; range 350–450 mg/m² [663–987 mg]), and the 17 patients who also were treated with maintenance chemotherapy received a mean cisplatin dose of 801 mg/m² (SD 99, range 600–950 mg/m² [1191–1846]). Plasma platinum concentrations in these patients were compared with those of 20 control patients who were cured from stage I testicular cancer by orchidectomy without chemotherapy. The median age of the control patients at the time of orchidectomy was 26 years (18–38 years). The median age at follow-up was 42 years (30–50 years), with a median follow-up duration of 14 years after orchidectomy (10–20 years). Platinum concentrations were measured in masked plasma samples by a sensitive procedure during which high-pressure decomposition of samples is followed by an adsorptive voltammetric measurement.[3] The limit of quantification of platinum was 6 pg/g plasma. Measurements were done in duplicate; the coefficient of variation and day-to-day variation were 6% and 5%, respectively.

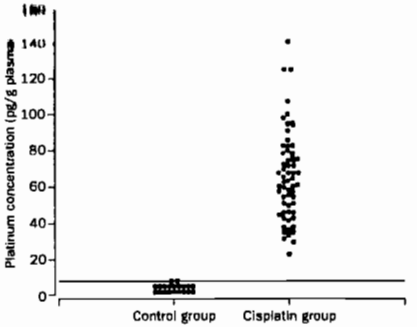


Figure 1: **Plasma platinum concentrations of 61 cured testicular cancer patients 10–20 years after cisplatin combination chemotherapy and 20 cured testicular cancer patients 10–20 years after orchidectomy**
Horizontal line=limit of quantification of platinum (6 pg/g plasma).

The platinum concentrations in the plasma of the 61 patients 10–20 years after cisplatin administration were significantly higher than those of the 20 control patients (cisplatin group: mean platinum concentration 64·9 pg/g plasma [SD 24·5] *vs* control group: 18 patients with platinum concentrations below the limit of detection and two patients with platinum concentrations at the limit of detection; Mann-Whitney *U* test, p<0·0001; figure 1). In all chemotherapy patients, the plasma platinum concentrations were above the limit of quantification, indicating that up to 20 years after

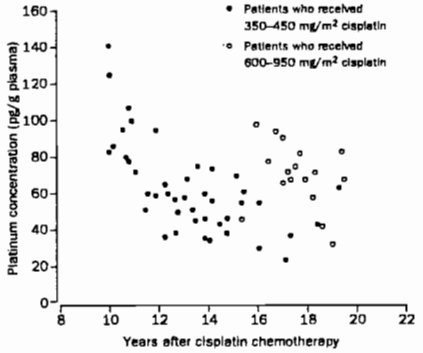


Figure 2: **Concentration-time plot of plasma platinum concentrations of 61 cured testicular-cancer patients 10–20 years after cisplatin combination chemotherapy**

(362) [4147]

TRANSDERMAL TESTOSTERONE TREATMENT IN WOMEN WITH IMPAIRED SEXUAL FUNCTION AFTER OOPHORECTOMY

JAN L. SHIFREN, M.D., GLENN D. BRAUNSTEIN, M.D., JAMES A. SIMON, M.D., PETER R. CASSON, M.D., JOHN E. BUSTER, M.D., GEOFFREY P. REDMOND, M.D., REGULA E. BURKI, M.D., ELIZABETH S. GINSBURG, M.D., RAYMOND C. ROSEN, PH.D., SANDRA R. LEIBLUM, PH.D., KIM E. CARAMELLI, M.S., AND NORMAN A. MAZER, M.D., PH.D.

ABSTRACT

Background The ovaries provide approximately half the circulating testosterone in premenopausal women. After bilateral oophorectomy, many women report impaired sexual functioning despite estrogen replacement. We evaluated the effects of transdermal testosterone in women who had impaired sexual function after surgically induced menopause.

Methods Seventy-five women, 31 to 56 years old, who had undergone oophorectomy and hysterectomy received conjugated equine estrogens (at least 0.625 mg per day orally) and, in random order, placebo, 150 μg of testosterone, and 300 μg of testosterone per day transdermally for 12 weeks each. Outcome measures included scores on the Brief Index of Sexual Functioning for Women, the Psychological General Well-Being Index, and a sexual-function diary completed over the telephone.

Results The mean (±SD) serum free testosterone concentration increased from 1.2±0.8 pg per milliliter (4.2±2.8 pmol per liter) during placebo treatment to 3.9±2.4 pg per milliliter (13.5±8.3 pmol per liter) and 5.9±4.8 pg per milliliter (20.5±16.6 pmol per liter) during treatment with 150 and 300 μg of testosterone per day, respectively (normal range, 1.3 to 6.8 pg per milliliter [4.5 to 23.6 pmol per liter]). Despite an appreciable placebo response, the higher testosterone dose resulted in further increases in scores for frequency of sexual activity and pleasure–orgasm in the Brief Index of Sexual Functioning for Women ($P=0.03$ for both comparisons with placebo). At the higher dose, the percentages of women who had sexual fantasies, masturbated, or engaged in sexual intercourse at least once a week increased two to three times from base line. The positive-well-being, depressed-mood, and composite scores of the Psychological General Well-Being Index also improved at the higher dose ($P=0.04$, $P=0.03$, and $P=0.04$, respectively, for the comparison with placebo), but the scores on the telephone-based diary did not increase significantly.

Conclusions In women who have undergone oophorectomy and hysterectomy, transdermal testosterone improves sexual function and psychological well-being. (N Engl J Med 2000;343:682-8.)

IN premenopausal women, the rate of production of testosterone is about 300 μg (1040 nmol) per day,[1] of which about half is derived from the ovaries and half from the adrenal glands.[2] In women who undergo bilateral oophorectomy before natural menopause, serum testosterone and estradiol concentrations decrease by approximately 50 and 80 percent, respectively.[3,4] These women are commonly treated with estrogen to prevent or ameliorate hot flashes, vaginal atrophy, osteoporosis, and heart disease.[5] Despite estrogen therapy, many surgically postmenopausal women have decreased sexual desire (libido), activity, and pleasure[6-8] and a decreased general sense of well-being.[9] These symptoms are believed to result from the lack of ovarian androgen production.

In a study of women in whom menopause had been induced by surgery, high doses of testosterone enanthate, given by intramuscular injection alone or in combination with estrogen, increased sexual desire, fantasies, and arousal more than placebo or estrogen alone.[6] In another study, therapy with testosterone and estradiol implants increased sexual activity, satisfaction, and pleasure and the frequency of orgasm more than estradiol alone.[10] However, the doses of testosterone enanthate were supraphysiologic, and in both studies the investigators knew which treatments the women received.[6,10,11] We undertook this study to determine the efficacy and safety of physiologic doses of testosterone, administered transdermally, in women who had impaired sexual function after surgically induced menopause.

METHODS

Study Subjects

We studied 75 healthy women, 31 to 56 years old, at nine clinical sites in the United States. All had undergone bilateral salpingo-oophorectomy and hysterectomy before natural menopause, at least 1 year but not more than 10 years earlier. All had serum testosterone concentrations of less than 30 ng per deciliter (1.0 nmol per liter) or serum free testosterone concentrations of less than 3.5 pg per milliliter (12.1 pmol per liter), which are below the median values for normal premenopausal women (Endocrine Sciences, Calabasas Hills, Calif.). All of the women had received conjugated equine estrogens at a daily dose of at least 0.625 mg orally for at least two months, had been in a stable, monogamous, heterosexual relationship for at least one year, and had a body-mass index (the weight in kilograms divided by the square of the height in meters) between 19.5 and 33.5. All were considered to have impaired sexual function on the basis of their affirmative answers to the following three questions: "At any time before surgery would you have characterized your sex life as active and satisfying?" "Since your surgery has your sex life become less active or less satisfying?" and "Would you prefer your sex life to be more active or more satisfying than it is now?"

Before enrollment the women completed the Brief Index of Sexual Functioning for Women,[12] a 22-item, multiple-choice questionnaire that provides scores pertaining to aspects of female sexuality (thoughts–desire, arousal, frequency of sexual activity, receptivity–initiation, pleasure–orgasm, relationship satisfaction, and problems affecting sexual function) and a composite score, ranging from −16 (poor function) to +75 (maximal function).[13] To qualify for the study, the women were required to have a composite score of less than 33.6, which is the mean value for normal women.[12] Women were excluded if they had received oral, topical, or vaginal androgen therapy in the previous three months or injectable or implantable androgen therapy in the previous six months; if they had more than 20 moderate or severe hot flashes per week, severe acne (grade 3 on the scale of Palatsi et al.[14]), moderate or severe hirsutism (score of 6 or more on the scale of Lorenzo[15]), hyperlipidemia, psychiatric illness, dyspareunia, or physical limitations that interfered with normal sexual functioning; or if they were taking glucocorticoids, selective serotonin-reuptake inhibitors, tricyclic antidepressants, antiandrogen agents, ginseng, yohimbine, phytoestrogens, dehydroepiandrosterone, or melatonin. The protocol was approved by the institutional review boards or ethics committees at all sites, and all the women gave written informed consent.

Study Design

After screening and a 4-week base-line period, the women began three consecutive 12-week treatment periods during which they received, in random order, the following regimens of transdermal patches applied twice weekly: two placebo patches (no active drug), one active and one placebo patch (nominal dose of testosterone, 150 μg per day), and two active patches (nominal dose of testosterone, 300 μg per day) (where the nominal dose is the amount of drug that will be absorbed by a person with average skin permeability during the application time). Neither the women nor the investigators knew the contents of the patches. Throughout the study, including the base-line period, the women received concomitant oral conjugated equine estrogens at their prestudy doses. The identical-appearing experimental patches (Watson Laboratories, Salt Lake City) were applied on the abdomen and were changed every three to four days.[16,17]

Serum Hormone Measurements

Serum free testosterone, bioavailable testosterone, total testosterone, dihydrotestosterone, and sex hormone–binding globulin were measured at base line and at weeks 4, 8, and 12 of each treatment period. Serum dehydroepiandrosterone sulfate, estradiol, estrone, luteinizing hormone, and follicle-stimulating hormone were measured at base line and at week 12 of each treatment period. Hormone assays were performed by Endocrine Sciences, Calabasas Hills, California.[17]

Evaluation of Sexual Function and Mood

Evaluation by means of the Brief Index of Sexual Functioning for Women was repeated at the end of the base-line period and at week 12 of each treatment period. The scores are expressed here as a percentage of the mean values derived from a previous study of 187 normal women between the ages of 20 and 55 years who had regular sexual partners.[13]

A telephone-based diary was also used to assess the frequencies of sexual thoughts, desires, and activities on a daily basis for 28 days during the base-line period and for the last 28 days of each treatment period. The women called a toll-free telephone number and responded to a series of recorded questions by using the telephone keypad. An overall frequency index was calculated as the sum of the frequencies of sexual thoughts, desires, and activities during the 28-day period.

Mood was assessed with the Psychological General Well-Being Index, a validated 22-item, multiple-choice questionnaire[18] that has been used in previous studies of postmenopausal women.[9,19] The Psychological General Well-Being Index provides scores for vitality, self-control, well-being, general health, depressed mood, and anxiety and a composite score that ranges from 0 (most negative affective experience) to 110 (most positive affective experience).[18]

Evaluation of Safety

Scores for hirsutism on the scale of Lorenzo (possible range, 0 to 20, with higher scores indicating greater hirsutism),[15] scores for acne on the scale of Palatsi et al. (possible range, 0 to 3, with higher scores indicating more acne),[14] facial-depilation rate (the number of times in the previous month that hair was removed from the chin or upper lip), serum lipid concentrations, fasting serum glucose concentrations, serum insulin concentrations, blood counts, indicators of liver function, and frequency of hot flashes were determined at base line and at the end of each treatment period. Tolerance of the skin to the transdermal systems and the occurrence of adverse events were recorded throughout the study.

Statistical Analysis

The primary efficacy end points were the composite score on the Brief Index of Sexual Functioning for Women and the overall frequency index from the telephone-based diary. Secondary end points included the scores for the various dimensions of the Brief Index of Sexual Functioning for Women and the composite and subscale scores from the Psychological General Well-Being Index. An intention-to-treat analysis was performed on data from all the women who completed the Brief Index of Sexual Functioning for Women at least once during treatment. Least-squares means corresponding to each treatment were estimated by a repeated-measures analysis of variance, with terms for period, sequence, and carryover effects (persistent effects from the preceding treatment period) included in the model.[20] Pairwise comparisons of values for each active dose with those for placebo (with base-line values subtracted) were performed with t-tests based on analysis of variance. On the basis of the categorical responses to question 7 of the Brief Index of Sexual Functioning for Women,[12,13] the percentages of women who reported having sexual fantasies, masturbating, or engaging in sexual intercourse at least once a week were estimated for descriptive purposes. In a post hoc analysis, composite and dimension scores on the Brief Index of Sexual Functioning for Women were analyzed for the subgroups of women who were less than 48 years old (the median age of the enrolled population) and those who were 48 or older. Serum hormone values during each treatment period were averaged and compared by analysis of variance

RESULTS

Seventy-five women were enrolled in the study and received at least one dose of transdermal study medication. The base-line characteristics of these 75 women are shown in Table 1. Eighteen women withdrew or were withdrawn from the study because of adverse events (three while receiving placebo, one while receiving 150 μg of testosterone per day, and two while receiving 300 μg of testosterone per day), poor compliance with the telephone diary (six women), or personal reasons (six women). Sixty-five women who had at least one evaluation for efficacy during treatment were included in the intention-to-treat analyses of scores on the Brief Index of Sexual Function-

From the Vincent Memorial Obstetrics and Gynecology Service, Department of Obstetrics and Gynecology, Massachusetts General Hospital, Boston (J.L.S.); the Department of Medicine, Cedars–Sinai Medical Center, Los Angeles (G.D.B.); Women's Health Research Center, Laurel, Md. (J.A.S.); the Department of Obstetrics and Gynecology, Baylor College of Medicine, Houston (P.R.C., J.E.B.); the Foundation for Developmental Endocrinology, Cleveland (G.P.R.); Salt Lake City (R.E.B.); the Center for Reproductive Medicine, Brigham and Women's Hospital, Boston (E.S.G.); the Department of Psychiatry, University of Medicine and Dentistry of New Jersey–Robert Wood Johnson Medical School, Piscataway (R.C.R., S.R.L.); and the Department of Clinical Research, Watson Laboratories–Utah, Salt Lake City (K.E.C., N.A.M.). Address reprint requests to Dr. Shifren at Vincent Memorial Obstetrics and Gynecology Service, VBK-113, Massachusetts General Hospital, Boston MA 02114, or at janshifren@hotmail.com.

Other authors were Kirtley P. Jones, M.D. (Department of Obstetrics and Gynecology, University of Utah Medical Center, Salt Lake City), and Claire A. Daugherty, M.S. (Anesta Corporation, Salt Lake City).

TABLE 1. BASE-LINE CHARACTERISTICS OF THE 75 WOMEN.

CHARACTERISTIC	VALUE
Age — yr	
Mean	47
Range	31–56
Body-mass index	
Mean	25.8
Range	19.5–33.5
Race or ethnic group — no. (%)	
White	62 (83)
Black	8 (11)
Hispanic	4 (5)
Asian	1 (1)
Years since oophorectomy and hysterectomy	
Mean	4.7
Range	1–10
Marital status — no. (%)	
Married	66 (88)
Cohabiting	7 (9)
Single, not cohabiting	2 (3)
Dose of conjugated equine estrogens — no. (%)	
0.625 mg/day	41 (55)
0.9 mg/day	12 (16)
1.25 mg/day	20 (27)
1.8 mg/day	1 (1)
2.5 mg/day	1 (1)
Previous androgen therapy — no. (%)	
Yes	23 (31)
No	52 (69)

ing for Women, the telephone-based diary, and the Psychological General Well-Being Index.

Serum Hormone Concentrations

The mean serum concentrations of free and bioavailable testosterone remained at low or low-normal values during placebo treatment, increased to mid-normal values during treatment with 150 μg of testosterone per day, and increased to high-normal values during treatment with 300 μg of testosterone per day (Table 2). The mean serum concentrations of total testosterone and dihydrotestosterone also increased and exceeded the respective normal ranges during treatment with 300 μg of testosterone per day. The mean serum concentration of sex hormone–binding globulin was high at base line because the women were taking oral conjugated equine estrogens and decreased slightly during testosterone treatment. The serum concentrations of dehydroepiandrosterone sulfate, estrone, estradiol, luteinizing hormone, and follicle-stimulating hormone did not change significantly during treatment.

Effects on Sexual Function and Mood

The mean (±SD) composite score on the Brief Index of Sexual Functioning for Women, expressed as a percentage of the mean value for normal women,[13] increased from 52±27 percent at base line to 72±38 percent during placebo treatment, 74±37 percent during treatment with 150 μg of testosterone per day, and 81±37 percent during treatment with 300 μg of testosterone per day (P=0.05 for the comparison with placebo) (Table 3). The scores for thoughts–desire, frequency of sexual activity, and pleasure–orgasm were lowest at base line and increased in a dose-dependent fashion. With the testosterone dose of 300 μg per day, the increases in scores for frequency of sexual activity and pleasure–orgasm were significantly greater than those with placebo (P=0.03 for both comparisons). The score for problems affecting sexual function was 116±48 percent of the normative mean at base line and decreased to 98±49 percent during treatment with 300 μg of testosterone per day (P=0.07 for the comparison with placebo).

To illustrate how the prevalence of particular types of sexual behavior varied during treatment, the following descriptive statistics were derived from the Brief Index of Sexual Functioning for Women. The percentage of women who reported having sexual fantasies at least once a week was 12 percent at base line, 10 percent during placebo treatment, 18 percent during treatment with 150 μg of testosterone per day, and 24 percent during treatment with 300 μg of testosterone per day. The percentage of women who reported masturbating at least once a week was 3 percent at base line, 5 percent during placebo treatment, and 10 percent during treatment with either 150 or 300 μg of testosterone per day. Finally, the percentage of women who engaged in sexual intercourse at least once a week was 23 percent at base line, 35 percent during treatment with either placebo or 150 μg of testosterone per day, and 41 percent during treatment with 300 μg of testosterone per day.

Compliance with the 28-day telephone-based diary was problematic and led to the withdrawal of six women. Missing calls averaged about four per month (14 percent of the data). The mean overall frequency index of sexual thoughts, desires, and activities was 13±12 events per month at base line and increased similarly by 5±12 events per month during treatment with placebo, 7±13 events per month during treatment with 150 μg of testosterone per day, and 6±13 events per month during treatment with 300 μg of testosterone per day. Although this index is less sensitive to treatment effects than the Brief Index of Sexual Functioning for Women, the changes in the overall frequency index of the telephone-based diary correlated significantly with the changes in the composite score on the Brief Index of Sexual Functioning for Women (r=0.54 for placebo; r=0.66 for 150 μg of testosterone per day; r=0.58 for 300 μg of testosterone per day; P<0.001 for all regimens). On the basis of the analysis-of-variance models used to analyze these primary outcome measures, there were no statistically significant effects of the treatment period, sequence, or carryover.

TABLE 2. MEAN (±SD) SERUM HORMONE CONCENTRATIONS IN THE WOMEN AT BASE LINE AND DURING EACH TREATMENT PERIOD.*

HORMONE†	BASE LINE	PLACEBO	150 μg OF TESTOSTERONE PER DAY	300 μg OF TESTOSTERONE PER DAY	NORMAL RANGE‡
Free testosterone (pg/ml)	1.1±0.7	1.2±0.8	3.9±2.4§	5.9±4.8§	1.3–6.8
Bioavailable testosterone (ng/dl)	2.0±1.4	2.2±1.3	7.1±4.1§	11.4±9.5§	1.6–12.7
Total testosterone (ng/dl)	21±10	22±12	64±25§	102±39§	14–54
Dihydrotestosterone (ng/dl)	7.6±4.0	8.4±4.8	17.8±7.0§	27.7±10.6§	4.4–20.4
Sex hormone–binding globulin (nmol/liter)	210±112	218±111	205±107¶	204±100§	36–185
Dehydroepiandrosterone sulfate (μg/dl)	61±34	60±34	58±33	57±34	60–255
Estrone (pg/ml)	148±114	130±133	143±110	146±88	32–159
Estradiol (pg/ml)	36±22	40±57	42±27	45±43	34–225
Luteinizing hormone (mIU/ml)	33.8±16.8	31.8±16.2	30.5±14.8	29.9±16.0	0.4–7.4
Follicle-stimulating hormone (mIU/ml)	42.7±22.8	41.4±21.1	39.0±20.4	38.2±20.9	1.7–7.2

*Seventy women with at least one hormone assessment during treatment were included in this analysis. All the women received conjugated equine estrogens throughout the study.

†To convert values for free testosterone to picomoles per liter, multiply by 3.467; to convert values for total and bioavailable testosterone to nanomoles per liter, multiply by 0.03467; to convert values for dihydrotestosterone to nanomoles per liter, multiply by 0.03444; to convert values for dehydroepiandrosterone sulfate to micromoles per liter, multiply by 0.02714; to convert values for estrone to picomoles per liter, multiply by 3.698; and to convert values for estradiol to picomoles per liter, multiply by 3.671.

‡The normal ranges in premenopausal women are from Endocrine Sciences, Calabasas Hills, California.

§P<0.001 for the comparison with placebo.

¶P=0.002 for the comparison with placebo.

TABLE 3. MEAN (±SD) SCORES ON THE BRIEF INDEX OF SEXUAL FUNCTIONING FOR WOMEN, EXPRESSED AS PERCENTAGES OF THE MEAN VALUES IN NORMAL WOMEN.*

DIMENSION	BASE LINE	PLACEBO	150 μg OF TESTOSTERONE PER DAY	300 μg OF TESTOSTERONE PER DAY
Composite score	52±27	72±38	74±37	81±37†
Thoughts–desire	48±31	67±40	72±40	77±40
Arousal	58±31	80±40	73±40	84±40
Frequency of sexual activity	41±31	53±41	58±40	64±40‡
Receptivity–initiation	68±33	89±39	86±39	92±39
Pleasure–orgasm	48±42	65±53	70±52	80±52‡
Relationship satisfaction	73±33	82±32	86±32	87±32
Problems affecting sexual function	116±48	108±49	97±49	98±49

*Sixty-five women from the intention-to-treat analysis were included in this analysis. Values are expressed as percentages of the mean values in normal women with partners, which were as follows: composite score, 33.6; thoughts–desire, 5.3; arousal, 6.2; frequency, 3.9; receptivity–initiation, 8.9; pleasure–orgasm, 4.9; relationship satisfaction, 8.9; and problems, 4.5.[13] All women received conjugated equine estrogens throughout the study. Least-squares mean values were estimated by analysis of variance, with terms for period, sequence, and carryover effects included in the model. P values are for the comparisons between the values during testosterone treatment (with base-line values subtracted) and the values during placebo treatment (with base-line values subtracted).

†P=0.05.

‡P=0.03.

The mean composite score on the Psychological General Well-Being Index was 78±15 at base line and increased by 1±14 during treatment with placebo, 2±14 during treatment with 150 μg of testosterone per day, and 5±14 during treatment with 300 μg of testosterone per day (P=0.04 for the comparison with placebo) (Table 4). There also were increases with testosterone treatment (indicative of improved mood) on the vitality, positive-well-being, depressed-mood, and anxiety subscales, which were significant at a dose of testosterone of 300 μg per day for positive well-being and depressed mood (P=0.04 and P=0.03 for the respective comparisons with placebo).

A post hoc analysis of the influence of age on the scores on the Brief Index of Sexual Functioning for Women was performed by comparing the subgroups of women under the median age of 48 years (31 women) with those 48 years of age or older (34 women). At base line, the mean composite score was 50±28 percent in the younger women and 53±27 percent in the older women. During placebo treatment, the composite score increased to 80±42 percent in the younger women, and there was no further improvement during testosterone treatment. In the older women, the composite score increased to 63±36 percent during placebo treatment, 76±37 percent during treatment with 150 μg of testosterone per day (P=0.03 for the comparison with placebo), and 81±38 percent during treatment with 300 μg of testosterone per day (P=0.003 for the comparison with placebo). The serum free testosterone concentrations were similar at base line and during treatment in both subgroups.

TABLE 4. MEAN (±SD) SCORES ON THE PSYCHOLOGICAL GENERAL WELL-BEING INDEX.*

INDEX SCALE†	BASE LINE	CHANGE WITH PLACEBO	CHANGE WITH 150 μg OF TESTOSTERONE PER DAY	CHANGE WITH 300 μg OF TESTOSTERONE PER DAY
Composite score	78±15	1±14	2±14	5±14‡
Anxiety	17±4	0±4	0±4	1±4
Depressed mood	12±2	0±2	0±2	1±2§
General health	12±2	0±2	0±2	0±2
Positive well-being	12±3	0±3	1±3	2±3‡
Self-control	12±3	0±2	1±2	0±2
Vitality	12±4	0±4	1±4	1±4

*Sixty-five women from the intention-to-treat analysis were included in this analysis. All the women received conjugated equine estrogens throughout the study. Least-squares mean changes from base line were estimated by analysis of variance, with terms for period, sequence, and carryover effects included in the model. P values are for the comparisons between changes from base line during testosterone treatment and changes from base line during placebo treatment.

†The range of possible scores is as follows: composite score, 0–110; anxiety, 0–25; depressed mood, 0–15; general health, 0–15; positive well-being, 0–20; self-control, 0–15; and vitality, 0–20.[18] For all scales, lower scores indicate a more negative affective experience and higher scores indicate a more positive affective experience.

‡P=0.04.

§P=0.03.

Safety

The hirsutism and acne scores did not change significantly during treatment (Table 5). The mean facial-depilation rate increased slightly during treatment with 300 μg of testosterone per day. The frequency of moderate or severe hot flashes averaged less than two per week at base line and was unaffected by testosterone treatment. The transdermal systems were well tolerated, with only one woman withdrawing because of a skin reaction caused by the placebo patches. Five serious adverse events occurred during the study. Four of these events (acute abdominal pain, angioplasty, bowel surgery, and a vasovagal episode) were considered to be unrelated to treatment; one (depression in a woman during placebo treatment) was considered to be possibly related. Treatment-related adverse events led four women to withdraw from the study (two who became anxious or agitated while receiving testosterone, one who had recurrence of a pink nipple discharge while receiving testosterone, and one with an application-site reaction). Transdermal testosterone treatment had no significant effects on the serum concentrations of total cholesterol, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, triglycerides, or fasting glucose or insulin; blood counts; or the results of liver-function tests (Table 5).

DISCUSSION

The women in this study had no ovarian androgen production and therefore had low serum concentrations of free and bioavailable testosterone at base line. These concentrations increased to the mid-normal and high-normal ranges, respectively, during transdermal treatment with 150 and 300 μg of testosterone per day. The supraphysiologic elevations in serum total testosterone and dihydrotestosterone at the higher dose were a consequence of the concomitant oral estrogen therapy, which raises serum concentrations of sex hormone–binding globulin[21] and reduces the clearance of androgens.[22,23] Serum estrogen concentrations did not change significantly during transdermal testosterone administration, indicating that aromatization of testosterone to estradiol[24] was minimal at the doses given.

Although they were receiving standard estrogen-replacement therapy, the base-line sexual function of the women was markedly impaired in comparison with that of normal women of similar age, as reflected by scores on the Brief Index of Sexual Functioning for Women.[13] The dimensions of thoughts–desire, arousal, frequency of sexual activity, and pleasure–orgasm were most affected. Although sexual function improved during placebo treatment, treatment with 300 μg of testosterone per day was associated with significantly greater improvement.

We can only speculate as to the origin of the strong placebo response in our study, why it was greater in the younger women, and why it tended to mask further effects of testosterone. As a condition of enrollment, all the women in our study wanted their sex lives to be more active or satisfying. Participating in the clinical trial may have facilitated communication within couples. In addition, the visible presence of the transdermal patches (active or placebo) might have been a stimulus to some women or their partners to increase sexual activity. Because the younger women had been in shorter relationships than the older women (13 vs. 18 years), they may have felt greater pressure to improve their sexual functioning. Finally, despite the use of statistical models that included terms for sequence and carryover effects, the crossover design (without washout periods) could have inflated the placebo response or caused "ceiling effects" in some couples who altered their patterns of sexual activity early in the study and then maintained the new patterns.

In contrast to the Brief Index of Sexual Functioning for Women, the 28-day telephone-based diary was less sensitive to treatment effects, and missed reporting days were a problem. These results are consistent with those of a recent methodologic study in which the required daily telephone reporting of sexual activity met with poor compliance.[25]

In regard to psychological status, testosterone replacement had a beneficial effect on well-being and depressed mood. The differences in the scores on the Psychological General Well-Being Index between the placebo and testosterone periods in our study are similar to the differences in scores between women who received estrogen alone after hysterectomy and bilateral oophorectomy and women who underwent hysterectomy without oophorectomy.[9] As expected, the serum free testosterone concentrations were higher in the women with intact ovaries.[9]

Finally, transdermal testosterone was not associated with clinically important changes in acne, hirsutism, or laboratory-test results, nor did it negate the beneficial effects of oral estrogen-replacement therapy on hot flashes and serum concentrations of high-density lipoprotein cholesterol.

In summary, treatment with transdermal testoster-

TABLE 5. MEAN (±SD) CLINICAL AND BIOCHEMICAL MEASURES AT BASE LINE AND DURING TREATMENT WITH TRANSDERMAL TESTOSTERONE PATCHES.*

MEASURE†	BASE LINE	PLACEBO	150 μg OF TESTOSTERONE PER DAY	300 μg OF TESTOSTERONE PER DAY	NORMAL RANGE‡
Androgenic skin effects					
Hirsutism score§	1.6±1.7	1.4±1.6	1.3±1.5	1.4±1.4	NA
Acne score¶	0.03±0.2	0.05±0.2	0.02±0.2	0.08±0.3	NA
Facial depilation‖	0.5±1.2	0.8±1.6	0.7±1.8	1.4±4.0**	NA
Serum lipids (mg/dl)					
Total cholesterol	214±31	221±36	220±39	221±34	134–253
HDL cholesterol	70±15	73±18	71±17	70±17	34–80
LDL cholesterol	117±26	121±36	122±39	123±33	72–164
Hematocrit (%)	39.5±2.6	39.4±2.4	39.4±2.5	39.7±2.4	34.9–44.5
Liver function					
Serum aspartate aminotransferase (U/liter)	20±5	19±6	19±5	19±5	12–31
Serum γ-glutamyltransferase (U/liter)	20±13	19±11	17±8	19±14	6–29
Serum albumin (g/dl)	4.3±0.3	4.2±0.3	4.2±0.3	4.2±0.4	3.5–5.0

*Sixty-seven women with at least one assessment during treatment were included in this analysis. All the women received conjugated equine estrogens throughout the study.

†HDL denotes high-density lipoprotein, and LDL low-density lipoprotein. To convert values for cholesterol to micromoles per liter, multiply by 0.026.

‡The normal ranges are from Mayo Medical Laboratories, Rochester, Minnesota. NA indicates that normal ranges were not available.

§Hirsutism was measured by the scale of Lorenzo (possible range, 0 to 20, with higher scores indicating greater hirsutism).[15]

¶Acne was measured by the scale of Palatsi et al. (possible range, 0 to 3, with higher scores indicating more acne).[14]

‖Facial depilation was measured on an index ranging from 0 to 28 times per month.

**P=0.04 for the comparison with base line.