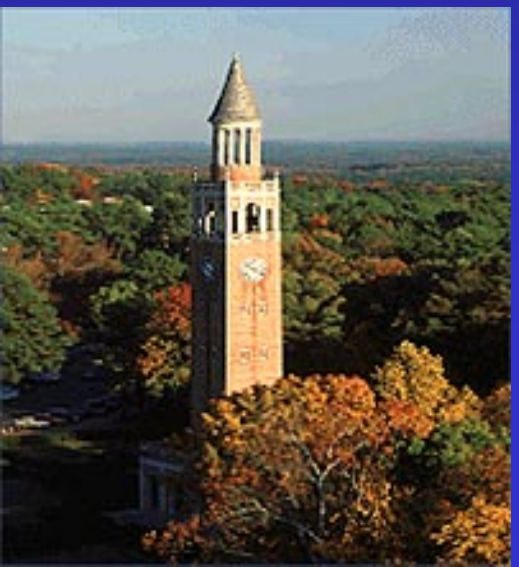
Validating Research Outcomes

Randomized Controlled Trial

vs.

Observational Data:

Differentiating Real-World from Perfect-World Data



William A. Marston, MD

Professor and Chief
Division of Vascular Surgery
University of North Carolina- Chapel Hill

Faculty Disclosure:
William Marston, MD

- Medical Advisory Boards
  - Advanced BioHealing
  - Baxter
  - Diversified Clinical Services

Goals and Objectives

Validating Research Outcomes

Objectives

- Define the GRADE system of clinical evidence and summarize the positive and negative applicability of these data sources in projecting clinical outcomes
- Review components of an RCT to outline the importance of RCT data concerning Advanced Therapeutics in directing clinical practice patterns
- Summarize published and unpublished ad-hoc data supporting the applicability of observational data to guiding physician practice
- Establishing a balance of information in developing recommendations and guidelines

WOUND SPRAY
TEA PRO
EQUINE
SilvaSorb
SILVER ANTIMICROBIAL DRESSING
SUSTAINED-RELEASE, SUPER ABSORBENT
Acticoat
Burn Dressing
BOUDREAUX'S
BUTT PASTE
DIAPER RASH OINTMENT (ZINC OXIDE)
• GOES ON AND CLEANS OFF EASILY
• TREATS AND PROTECTS
• PLEASANT SCENT
NET WT. 16 OZ. (453.6g)
Breakthrough in non-drug wound care
Natural Antibiotic-Free Healing
Helps Minimize Scarring
2nd Skin
Quick Heal
Advanced Natural Healing
proK
DuoDERM
Hydroactive Paste/Pasta/Pâte
mesalt
Dermafoam
Waterproof Foam Dressing
EPI-MAX
STERILE WOUND DRESSING
10 Dressings
4 in. x 4 in.
Dermagenics
lodosorb

What is the evidence that?

Boudreaux's Butt paste is any better than

Burt's bees wax

Outcomes in Wound Healing

- Clinically relevant outcomes
  - Wound closure
  - Reduced symptoms, pain
  - Function / Quality of Life
  - Limb preservation
- Surrogate outcomes
  - Short term wound size decrease
  - Control of bacteria
  - Control of drainage, odor
- Advantages of surrogate outcomes
  - Less costly
  - Less invasive
  - Shorter follow-up
  - Smaller sample size

Evaluative Instruments

- Objective clinical outcome measures
  - Venous clinical severity score
- Patient reported health-related quality of life
  - Generic measures
    - SF-36
    - Nottingham health profile
  - Disease specific measures
    - None specific to wounds

What is Evidence-Based Medicine?

Sackett DL; Spine 1998

Evidence-based medicine is the conscientious, explicit use of the current best evidence in making decisions about the care of individual patients. The practice of evidence-based medicine means integrating …"

- Individual clinical expertise
  - Clinical experience
  - Considered regard for patient's predicaments, rights, and preferences
- Best available external clinical evidence
  - Accuracy and precision of diagnostic tests
  - Power of prognostic markers
  - Efficacy & safety of therapeutic & preventative regimens

Levels of Evidence

Level of Evidence	Studies of therapy, prevention, etiology, harm
1a	Systematic review with homogeneity of RCTs
1b	Individual RCT with narrow confidence intervals
1c	All or none
2a	Systematic review with homogeneity of cohort studies
2b	Individual cohort study or low quality RCT
3a	Systematic review with homogeneity of case-control studies
3b	Individual case-control study
4	Case series
5	Expert opinion without explicit critical appraisal

GRADE

- Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) working group
  - Agency for Healthcare Research and Quality (US)
  - National Institute for Clinical Excellence (UK)
  - World Health Organization (WHO)

- Effort to unify competing systems of evidence evaluation

- 2 components of all treatment recommendations
  - Strength of recommendation
  - Strength of evidence supporting recommendation

Strength of Recommendation

- Reflects balance of benefits, risks, burdens, & cost
  - Includes patients values and preferences
  - Includes consideration of local resources and cost
- Grade 1 (strong) recommendations
  - Benefits >> risk
  - Informed patient would always choose therapy
  - “Just do it”
- Grade 2 (weak) recommendations
  - Benefits and risks more closely balanced
  - Different patients may make different choices
  - Requires knowledge of the evidence
  - Roles for decision aids and patient discussion

The Hierarchy of Clinical Evidence

Methodology	Quality	Bias	Estimate of Effect
Systematic Review of RCTs	High	Low	Precise
Randomized Clinical Trials			
Observational Studies			
Cohort Studies			
Case-Control Studies			
Case Series			Unknown
Expert Opinion	Low	High	Unknown

The 4 elements of Evidence Quality

GRADE Working Group; BMJ 2005

- Study Design
  - Randomized trials
  - Observational studies (Cohort, case-control, etc)
- Study Quality
  - Allocation concealment
  - Blinding
  - Loss to follow-up
- Consistency - Similar estimate of effect across studies
- *Directness - Similarity of subjects, interventions, and outcome measures to those of interest*
  - *Use of surrogate outcomes*
  - *Lack of direct comparisons of interventions*

Methodological Quality

- Reflects confidence in estimates of benefit / harm

Grade	Quality	Methodology
A	High	Homogenous RCTs without limitations Observational studies with overwhelming effects
B	Moderate	RCTs with limitations Exceptionally strong observational studies
C	Low	RCTs with multiple, serious limitations Observational studies Case reports Expert opinion

GRADE Recommendations

Guyatt et al, Chest 2006

Grade	Benefit vs Risk	Methodology	Implications
1A	Clear	High quality	Strong recommendation, Generalizeable
1B	Clear	Moderate	Strong Recommendation, Applies to most patients
1C	Clear	Low	Strong recommendation, but may change w better evidence
2A	Balanced	High quality	Weak recommendation, Action differs with patient/societal values
2B	Balanced	Moderate	Weak recommendation, Action differs with patient/societal values
2C	Uncertain	Low	Very weak recommendation, Alternatives equally reasonable

All RCTs are not created equally

- What is purpose of study?
  - FDA PMA approval
    - Subject to FDA audit
    - Closely monitored
    - Enrollment criteria strictly controlled
    - Protocol well defined, violations noted
  - Non FDA
    - May not be well monitored
    - Enrollment criteria less controlled
    - May have frequent protocol violations

Rigor of FDA PMA trial: example

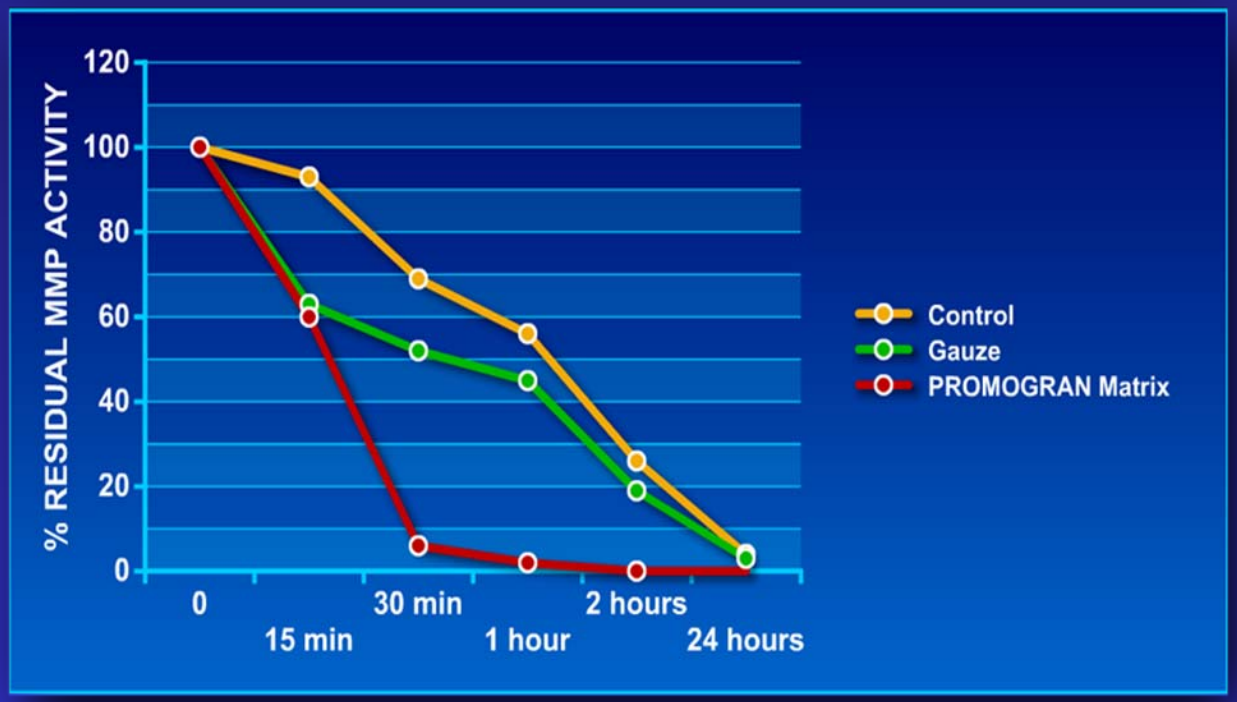
- DFU pivotal trial
- Several patients healed wound but failed to come for follow-up visit during time window for confirmation of healing visit
- Even if patient had healing confirmed at later visit, they were classified as “not healed” by study criteria
- Several patients classified as “healed” by investigator changed to “not healed” after monitor review of photographs

Non-PMA RCT

- Is design consistent across groups?
- Is study blinded?
- Is study monitored?
- Are results reported based on intent to treat?
- How are dropouts and protocol violations handled?
- Are investigators biased?

Do RCTs reflect “real world” outcomes? Promogran

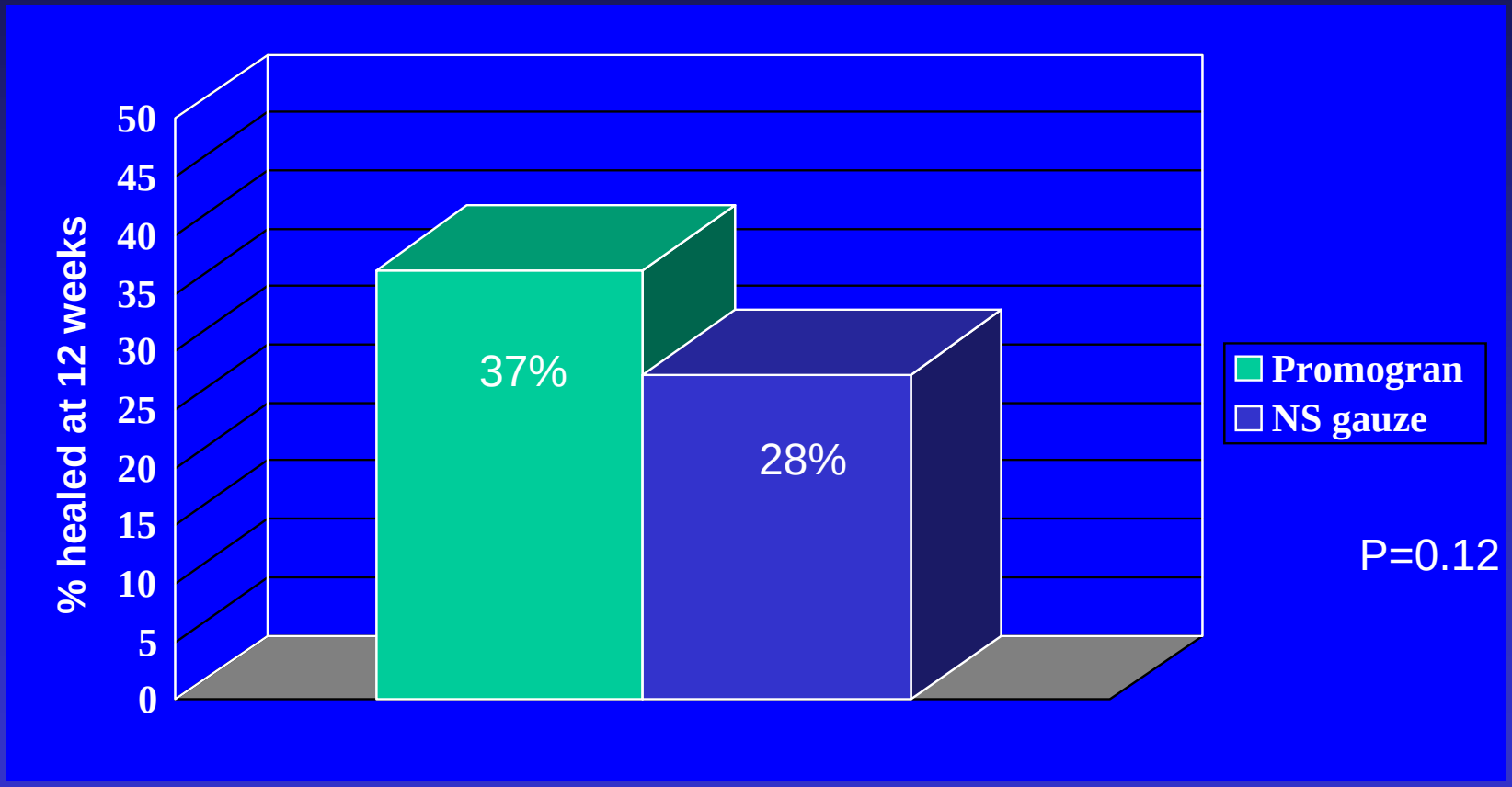
- Elevated MMP levels in non-healing DFUs
- ORC/Collagen binds MMPs

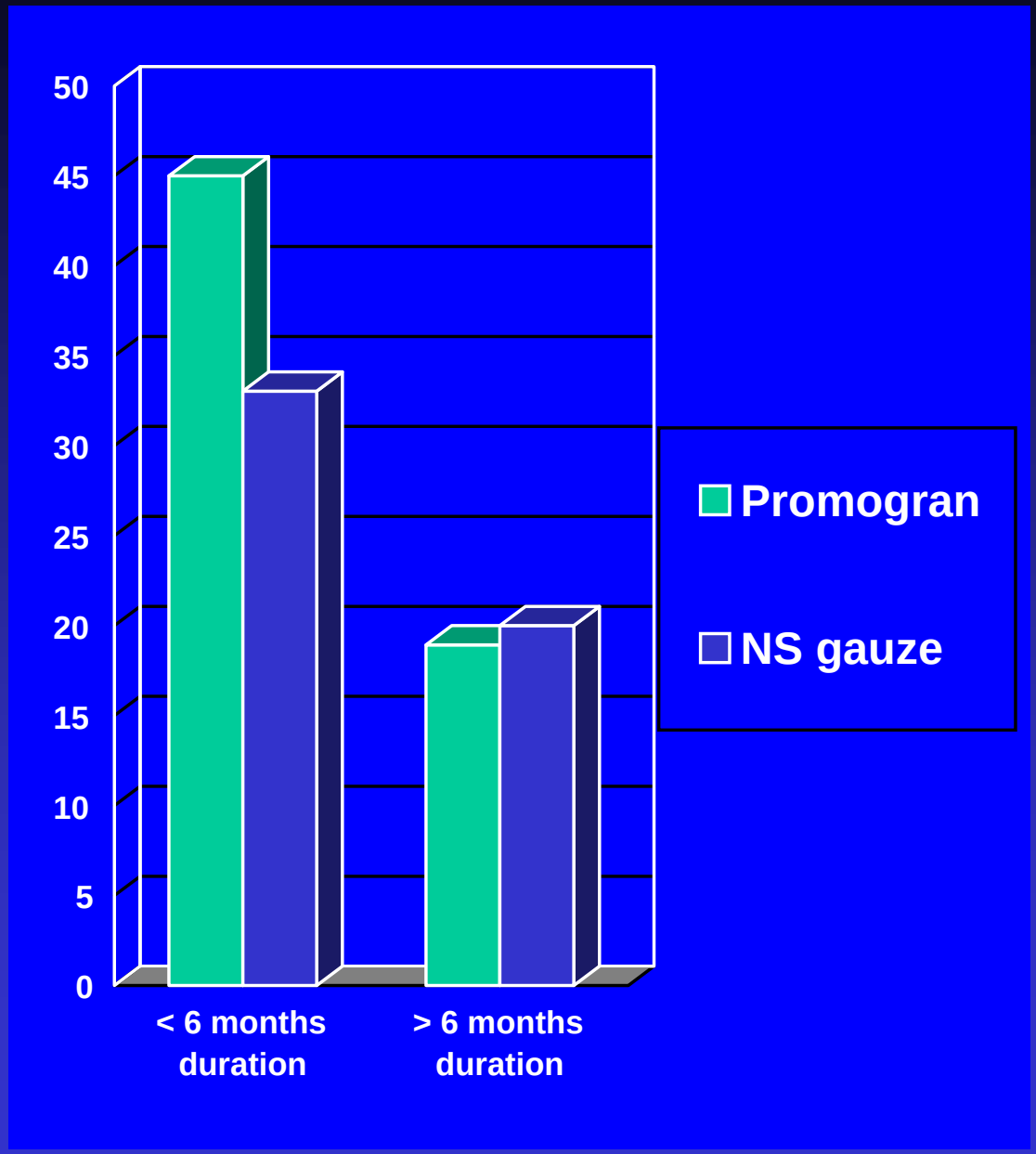


Promogran: Clinical studies

- Veves et al Arch Surg 2002;137:822-7
- Randomized controlled trial of Promogran in management of diabetic foot ulcers
- Goal: FDA indication for healing
- 276 patients with DFU enrolled
  - 138 Promogran plus standard care
  - 138 NS gauze plus standard care
- Endpoint: % healed at 12 weeks

Promogran DFU study: % healed at 12 weeks

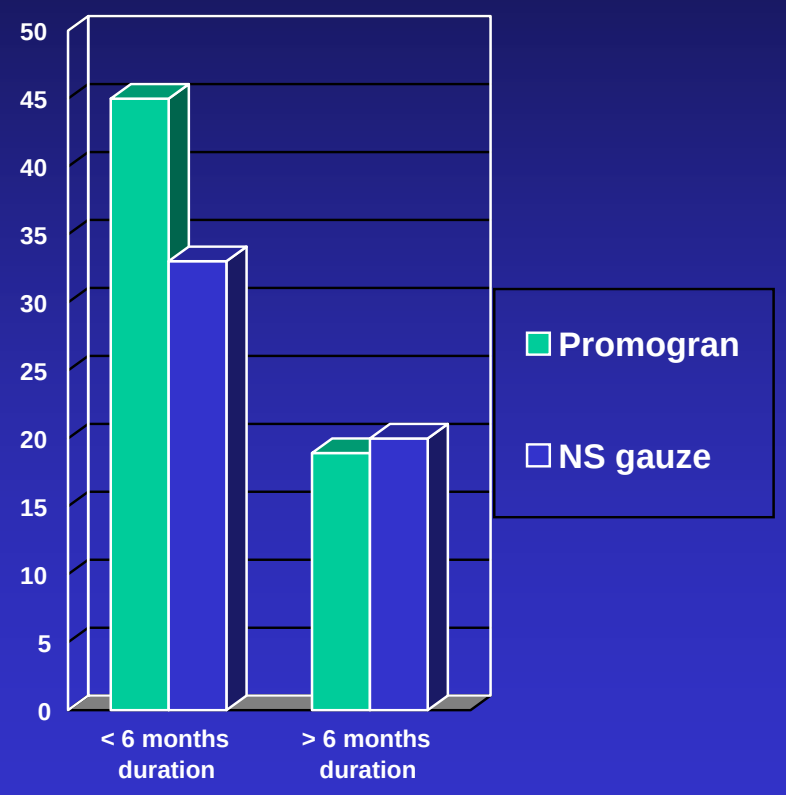



50
45
40
35
30
25
20
15
10
5
0
< 6 months duration
> 6 months duration
Promogran
NS gauze

Promogran study analysis

- Adequately powered, blinded, statistically valid, well designed and executed study
- Product not designed for use in rigid protocol from day 0 until healing
- Temporary use to reduce MMP levels then replace with another modality to accelerate tissue coverage
- MMP activity required for angiogenesis and epithelial migration

Promogran: < 6 month duration wounds



- Wounds of long duration have senescent burned out fibroblasts
- Little capacity to heal
- Low MMP production

- Study needed surrogate test for MMP levels

FDA Pivotal RCT

- Typically yield data that appears inferior to data from “real world” studies
- FDA is concerned with patient safety and potential of protocol to determine efficacy of novel therapy
  - Strict blinding
  - Strict inclusion/exclusion criteria
    - Many of easy-to-heal patients excluded
  - Data analysis by intent-to-treat
  - Close study monitoring with analysis of protocol violations
  - Standard of care may not really be SOC

Current management of VLU vs RCT

- Every visit wound is examined for
  - adequate edema control
  - Biofilm
  - Inflammation
  - Healing progress
- Treatment plan altered at each visit
  - Different mode of compression
  - Culture/antibacterial treatment
  - Anti-inflammatory strategy
  - LSS or other active therapy

DIAGNOSTICS!

Secondary study analysis

- Risk factors related to healing
  - Age
  - Gender
  - Wound size
  - Wound duration
  - Medical conditions
    - Renal failure
    - PAD
    - Poorly controlled diabetes

Secondary study analysis

- May be pre-planned or post-hoc
- Pre-planned provides designated list of secondary factors to be analyzed
- Post-hoc requires accounting of all factors studied to determine potential error of analysis

Secondary study analysis

- May yield interesting and important associations with healing leading to important treatment decisions
  - Example: Predictive potential of wound improvement at 4 weeks

BUT

- Analysis of a large number of secondary factors may lead to chance positive findings

Secondary study analysis

- If $p<.05$ is used for significance and 10 secondary factors are studied, the probability of finding a positive association where none actually exists is 40%

- If 40 secondary factors are studied, the probability is over 80%

Secondary study analysis

- Statisticians suggest that secondary analysis should utilize higher standards for significance
  - P = 0.05/number of secondary factors studied
  - If 10 studied, p< .005 necessary

- Or, factors identified with standard analysis should be confirmed by another study designed to test this specific association
- Some journals refuse to publish post-hoc analyses

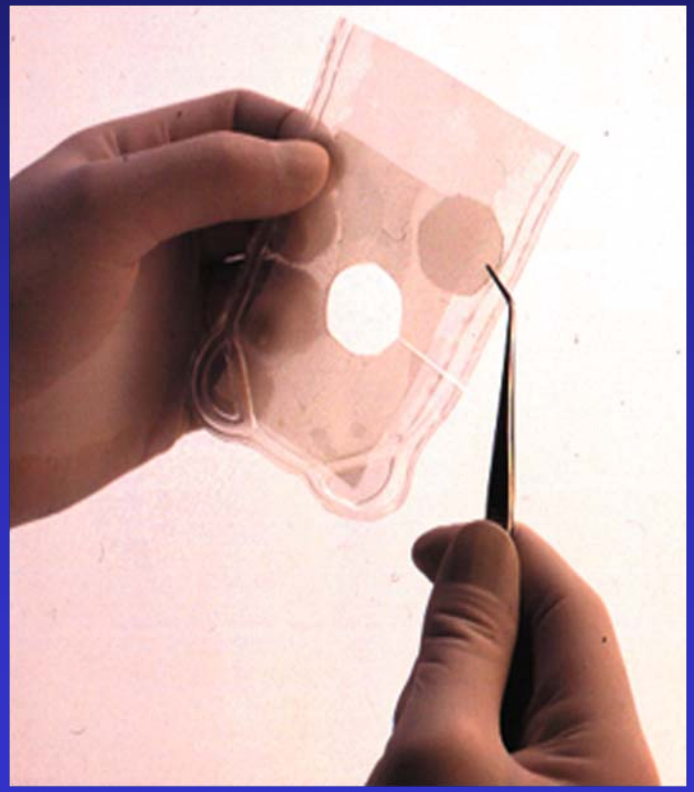
RCT quality issues

- Sample size
- Proper blinding and randomization
- Dropouts
- Handling of dropouts or others not completing study
  - Intent-to treat analysis
- Multi-center vs single center

Quality of evidence in treatment of chronic wounds

Examples

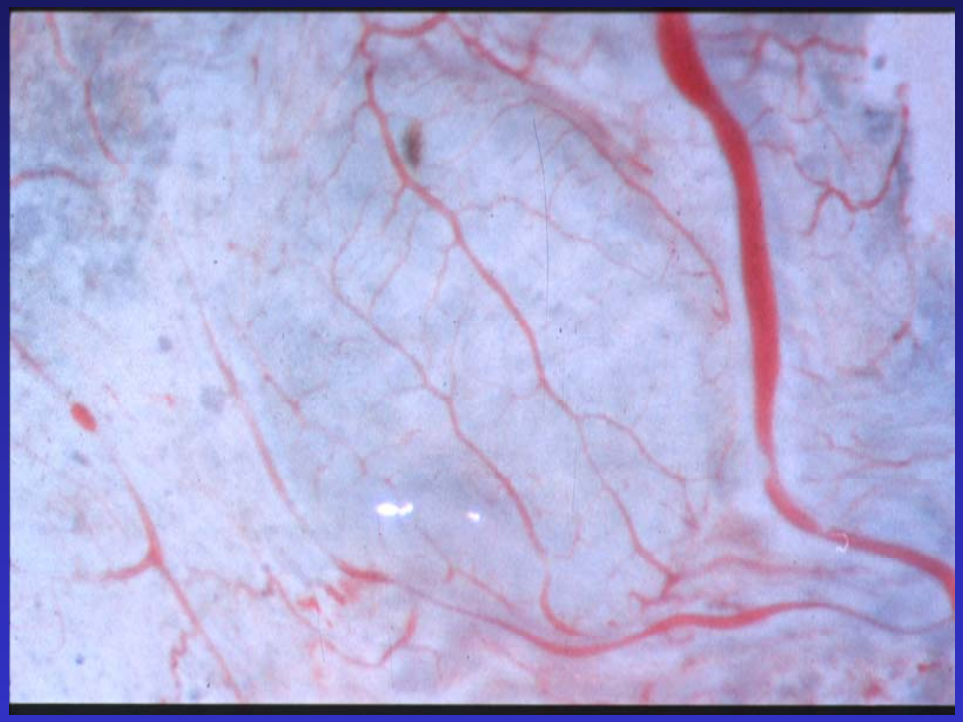
Dermagraft: Dermal fibroblasts w/out epidermal layer



Growth factors produced

- PDGF A chain
- IGF
- HGF
- TGF beta
- VEGF

- Angiogenesis into graft tissue supported



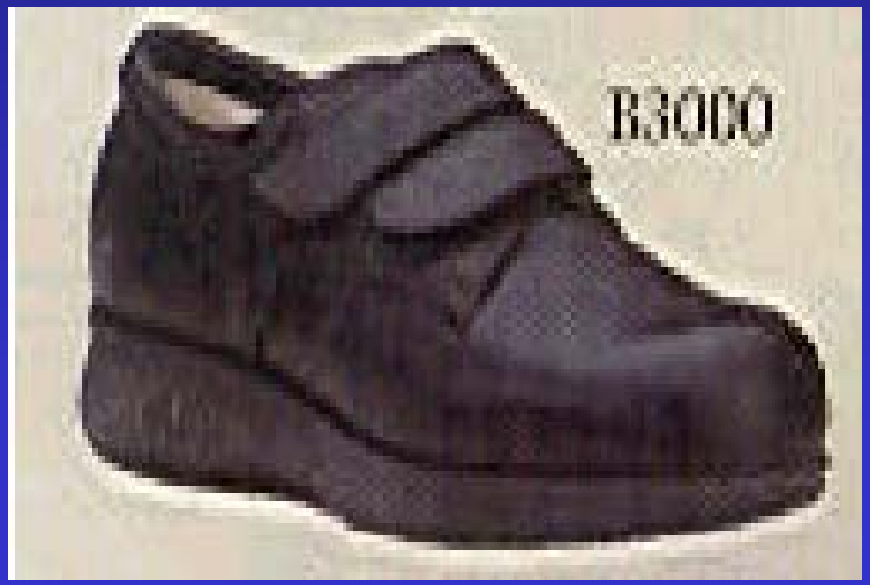
DG pivotal trial for DFU

- Multicenter prospective randomized clinical trial of DG for DFU undertaken with metabolic activity in optimal range
- Patient accrual 1998-2000
- 314 patients enrolled at 35 centers
- Ulcer size 1 cm^2 to 20 cm^2
- Ulcer extends through dermis
- Patient has normal arterial circulation
  - Ankle/brachial index > 0.7

Ambulatory Protocol

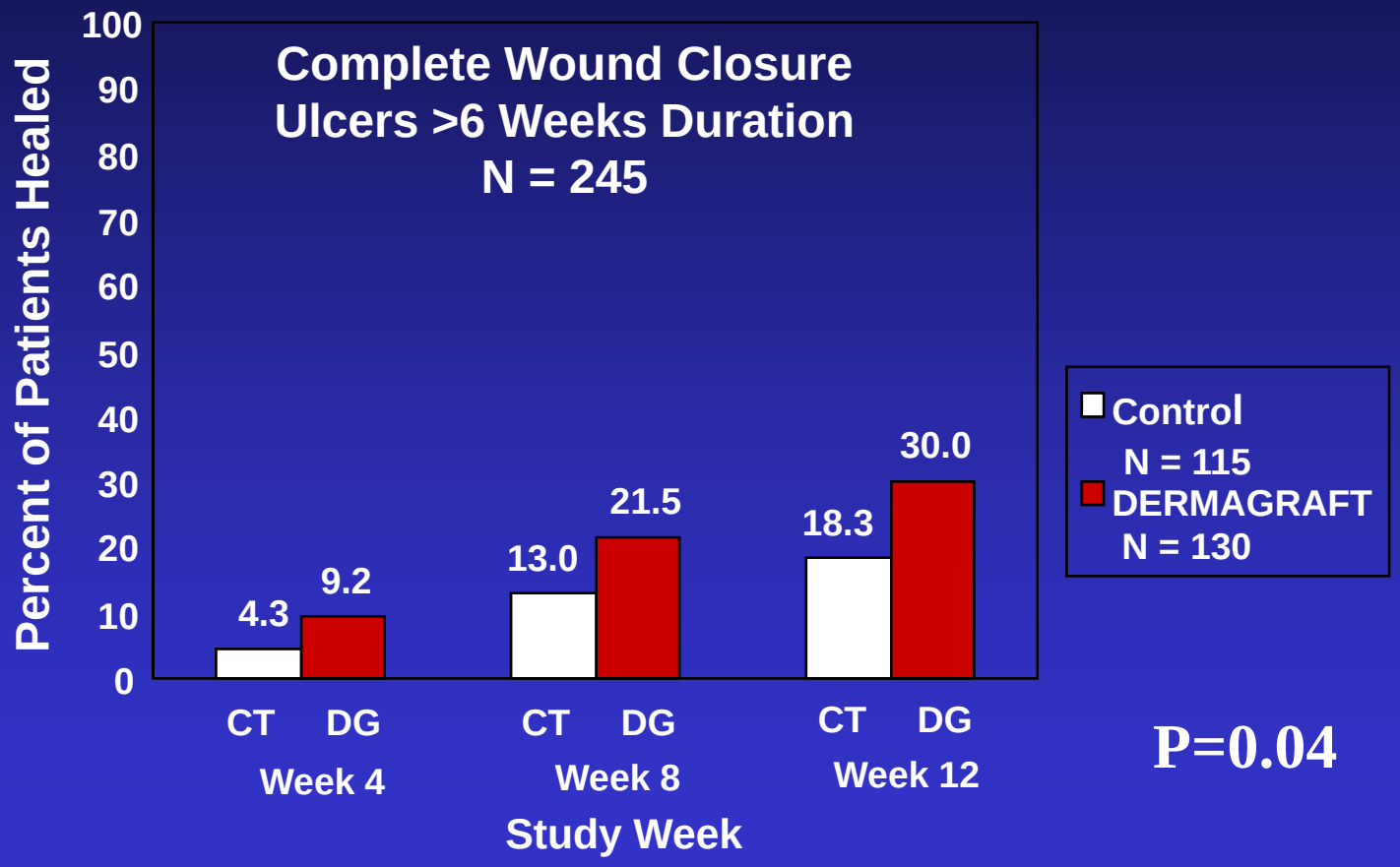
- Orthotics with custom molded inserts
- Limited ambulation recommended

Biomechanical Ambulator



Healing Sandal

Clinical Study Results



P=0.04

Adverse events

- Number of ulcer related AEs was significantly less in the DG group ($p<0.05$)
  - 19% in DG group
  - 33% in control group

  - Ulcer related AEs defined as local wound infection, cellulitis, osteomyelitis

RCT quality analysis

- Study adequately powered
- Not blinded; investigators knew which patients received DG
- Dropout rate acceptable
- Statistical analysis based on intent-to-treat
- Healing required confirmation visit
- Rigorous monitoring

Reaction to study

Wound clinicians

- Control healing rate unacceptably low
- Improvement in treatment group not very impressive
- Not a very good protocol

Statisticians and efficacy reviewers

- Protocol well designed to determine product safety and efficacy
- Offloading must be identical across groups
- SOC wet to dry

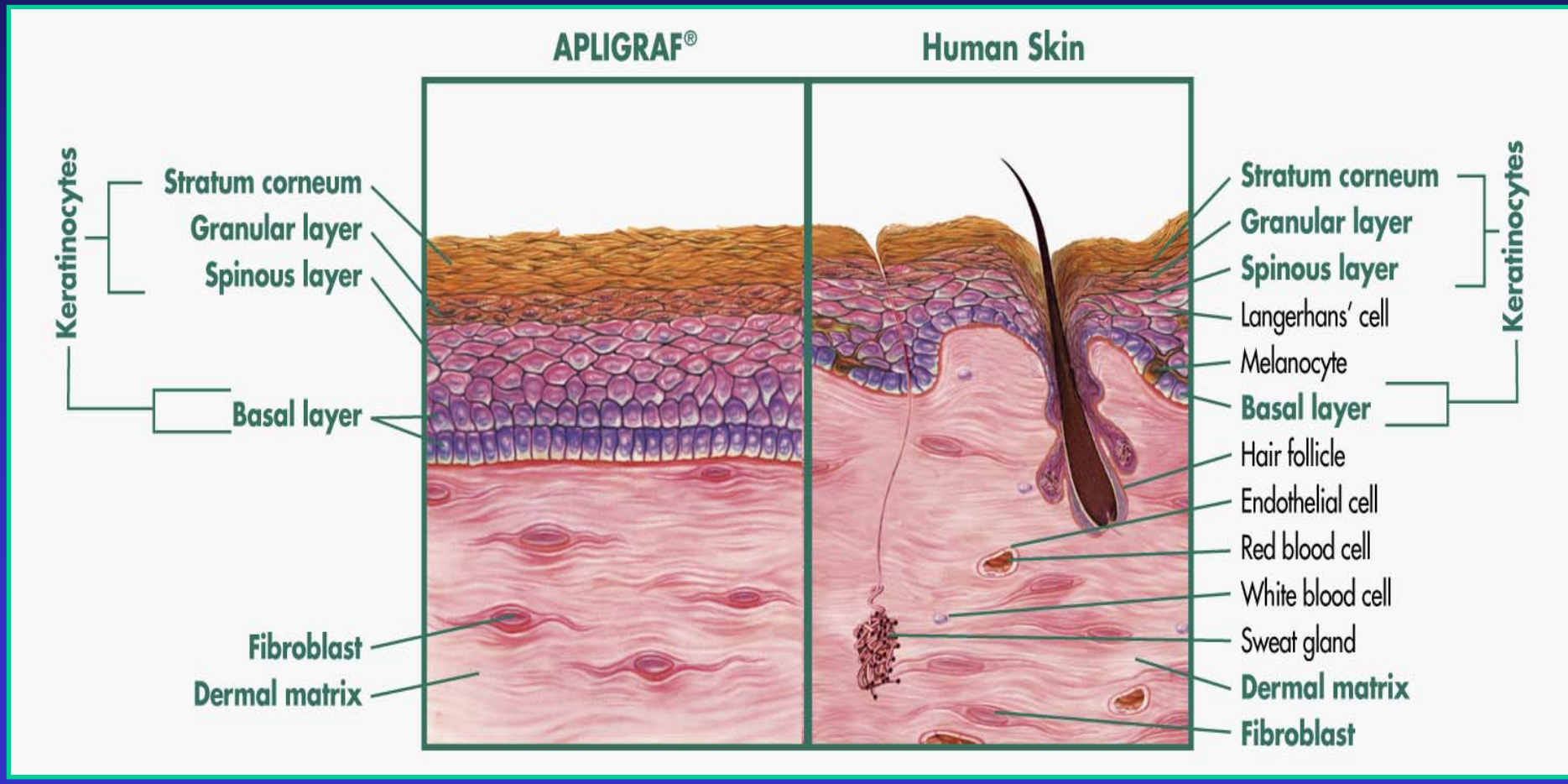
Clinical Study Results

- Percent of wound area closure by Week 12:
  - DERMAGRAFT group 91%
  - Control group 78%

$P<0.05$

- Wounds healing rapidly during screening run-in excluded
- Over 15% of DG treatment group did not receive full course of therapy - maintained in DG group for intent to treat analysis
- Patients reported being ambulatory an average of 8 hours per day.

Apligraf

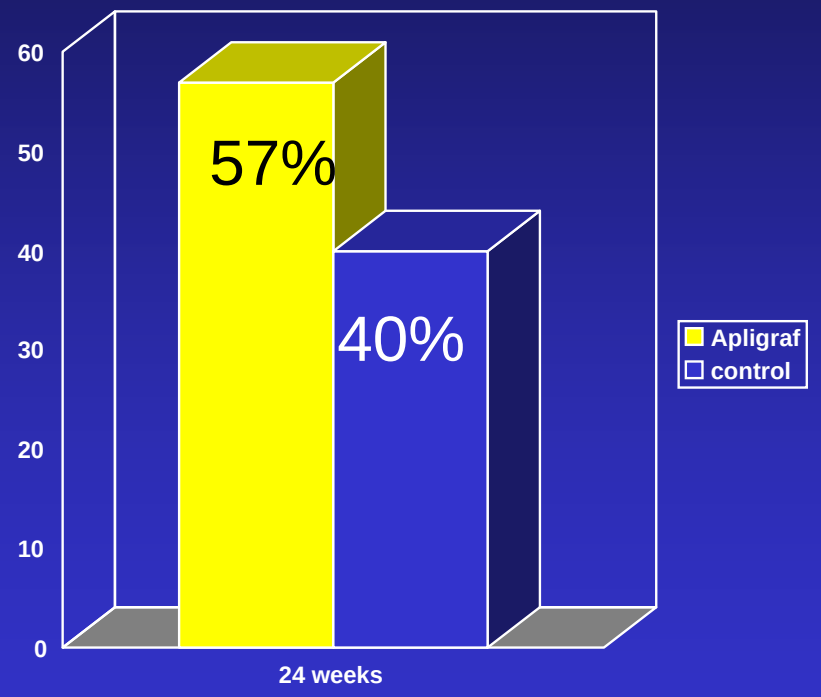

APLIGRAF®
Human Skin
Keratinocytes
Stratum corneum
Granular layer
Spinous layer
Basal layer
Fibroblast
Dermal matrix
Stratum corneum
Granular layer
Spinous layer
Langerhans' cell
Melanocyte
Basal layer
Hair follicle
Endothelial cell
Red blood cell
White blood cell
Sweat gland
Dermal matrix
Fibroblast
Keratinocytes

Apligraf: Method of application

- Mesh graft or cut slits
  - 1.5 to 1
- Cut to fit ulcer outline
- Suturing not necessary
- Primary nonadherent dressing
  - Mepitel
  - May use silver, cadexomer iodine

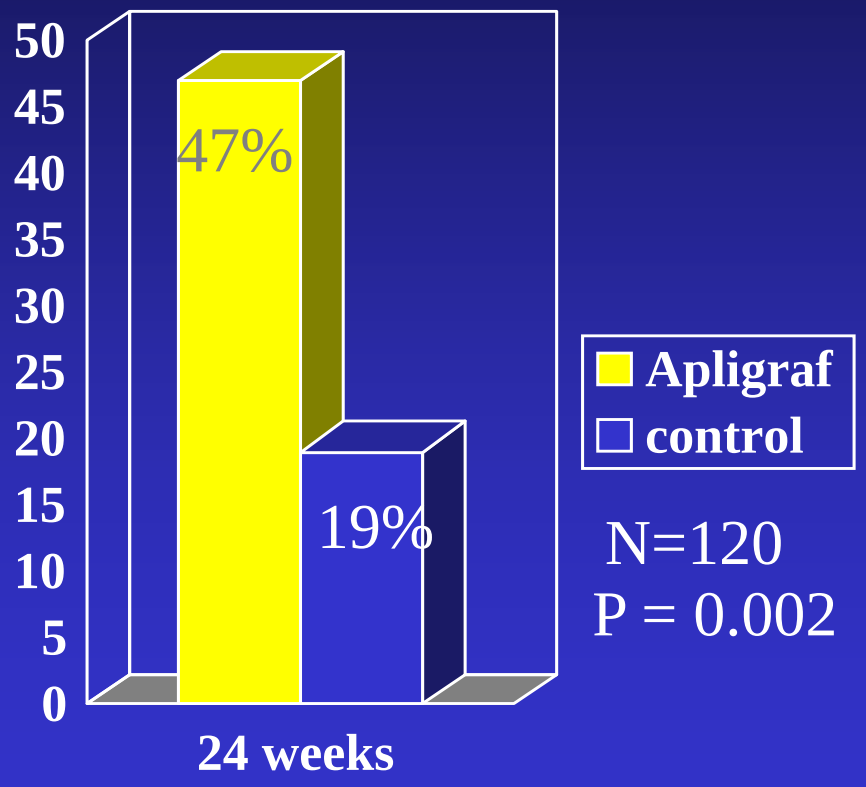
Apligraf: venous leg ulcer pivotal trial

- 240 patients
- Ulcers present > 1 month
- Average 3.3 applications
- Applied with compression
- Percent closed at 24 weeks compared to compression alone

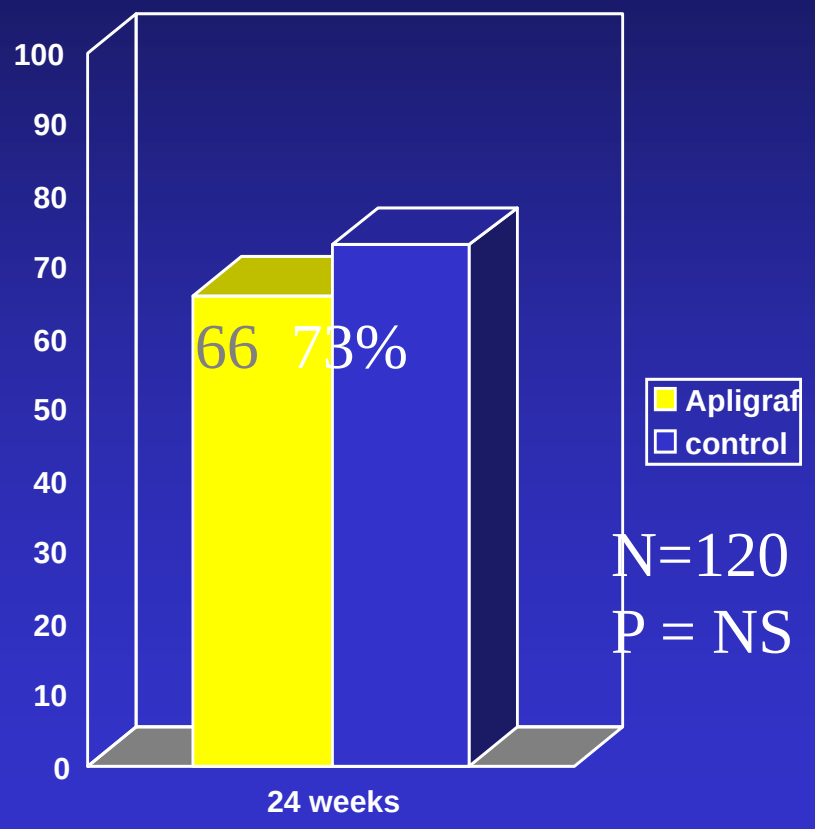


P=0.02

Apligraf: venous leg ulcer pivotal trial



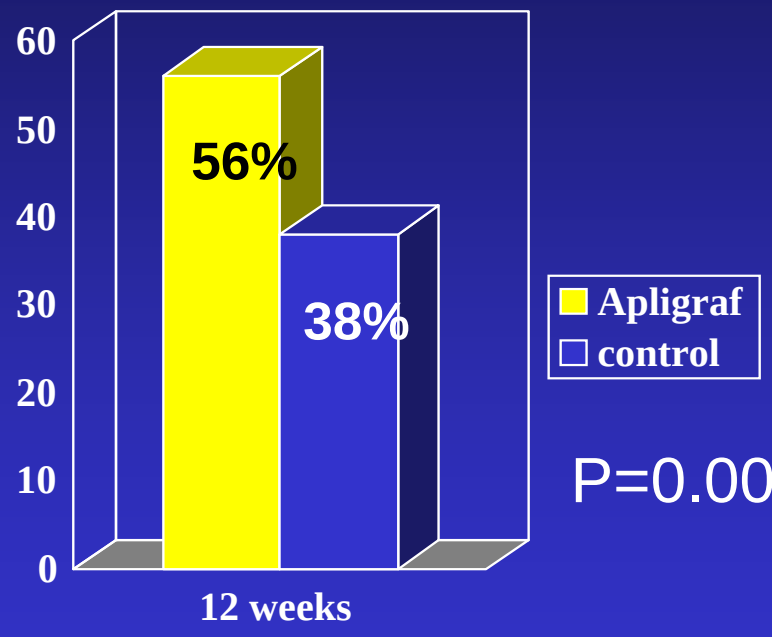
Ulcer duration > 1 year



Ulcer duration < 1 year

Apligraf in diabetic foot ulcers

- 208 patients randomized prospectively
- 24 wound care centers
- All received standard wound care, etc
- followed 12 weeks



P=0.003

RCT quality analysis

- Study adequately powered
- Not blinded; investigators knew which patients received AG
- Dropout rate acceptable
- Statistical analysis based on intent-to-treat
- Healing required confirmation visit
- Rigorous monitoring

Comparison of results of multiple RCTs: Meta-analysis

Heterogeneity of studies

- How similar are inclusion/exclusion criteria?
- Are treatment protocols similar for SOC?
- Are outcome measures similar?

Dermagraft vs Apligraf DFU

Dermagraft

- Offloading protocol via ambulatory diabetic shoe or sandal w custom insert
- Enrollment and run-in criteria

- 18% closure rate in SOC at 12 weeks
- 60% relative increase in % healed at 12 wks

Apligraf

- Offloading not standardized; most used DH walker or full length cast boots

- 38% closure rate in SOC at 12 weeks
- 68% relative increase in % healed at 12 wks

Apligraf or Dermagraft for DFU

Grade	Benefit vs Risk	Methodology	Implications
1A	Clear	High quality	Strong recommendation, Generalizeable
1B	Clear	Moderate	Strong Recommendation, Applies to most patients
1C	Clear	Low	Strong recommendation, but may change w better evidence
2A	Balanced	High quality	Weak recommendation, Action differs w patient/societal values
2B	Balanced	Moderate	Weak recommendation, Action differs with patient/societal values
2C	Uncertain	Low	Very weak recommendation, Alternatives equally reasonable

Observational studies

- Pros
  - If large may provide level of information that smaller RCTs cannot on outcomes in general population
  - More reflective of real world results
- Cons
  - Not well standardized
  - Ripe for investigator bias
  - Difficult to use for comparative study of different treatments due to patient selection factors and practice pattern variation

Curative Health Services: circa 1995

- “80% of wounds in our advanced wound healing centers close completely within 8 weeks”
  - Multiple publications, most not peer-reviewed

Margolis et al, DFUs: Predicting which ones won't heal

- Studied 27,630 DFU cases in Curative Health Services system database
- 47% healed at 20 weeks
- Factors associated with poor healing
  - Wound size > 2 cm^2
  - Wound duration > 2 months
  - Wagner grade $\geq$ 3
  - Presence of multiple foot wounds

Am J Med 2003;115:627-631

Small observational studies

Relationship of pressure relief to healing with Dermagraft

- 14 patients treated with Bledsoe conformer boots and strict ambulation limitation
- DG applied to diabetic plantar ulcers, average 2.7 pieces per patient
- At 12 weeks 64% healed
- 5 of remaining 6 > 50% closed
- One AE, cellulitis in study ulcer

Observational Studies

- Size
- Outcomes measures
  - Objectivity
  - Clinically relevant
- Potential for investigator bias
- Corroboration from multiple sites

- Do studies provide “overwhelming evidence”?

SAWC poster session: (Rough numbers)

- Case series section (about 150 posters)
- < 10% of posters described outcomes in more than 10 patients treated
- Only 3 found that described outcomes in > 50 treated patients

Examples of use of GRADE criteria for modalities used in the treatment of chronic wounds

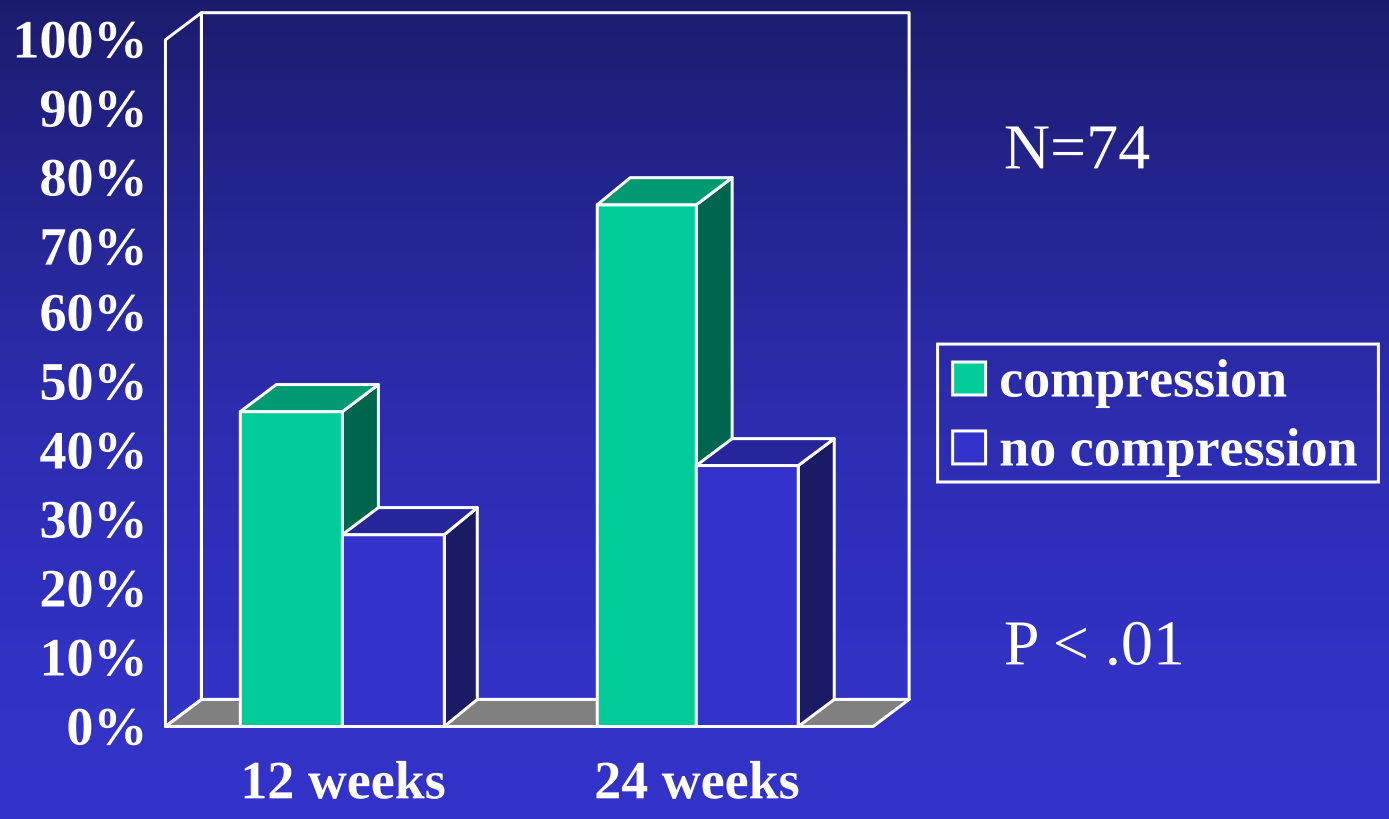
Example:1A recommendation

- Compression for venous leg ulcers

Cochrane Library meta-analysis 2000

- 8 prospective randomized studies found evaluating use of compression for treatment of VLU
- Appears clear that compression is better than no compression
- High strength is better than low strength

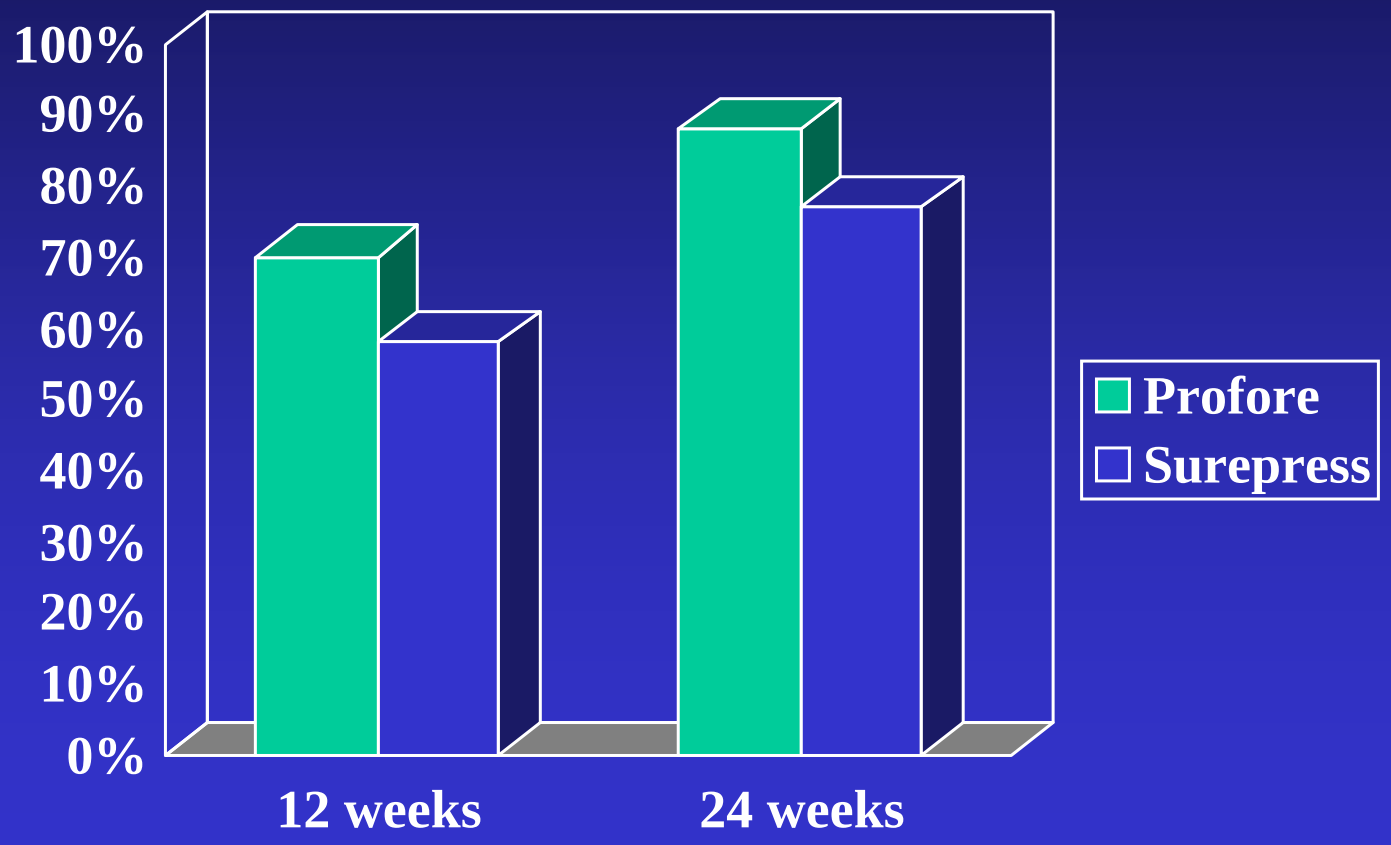
Healing of VLU with/wo compression



N=74

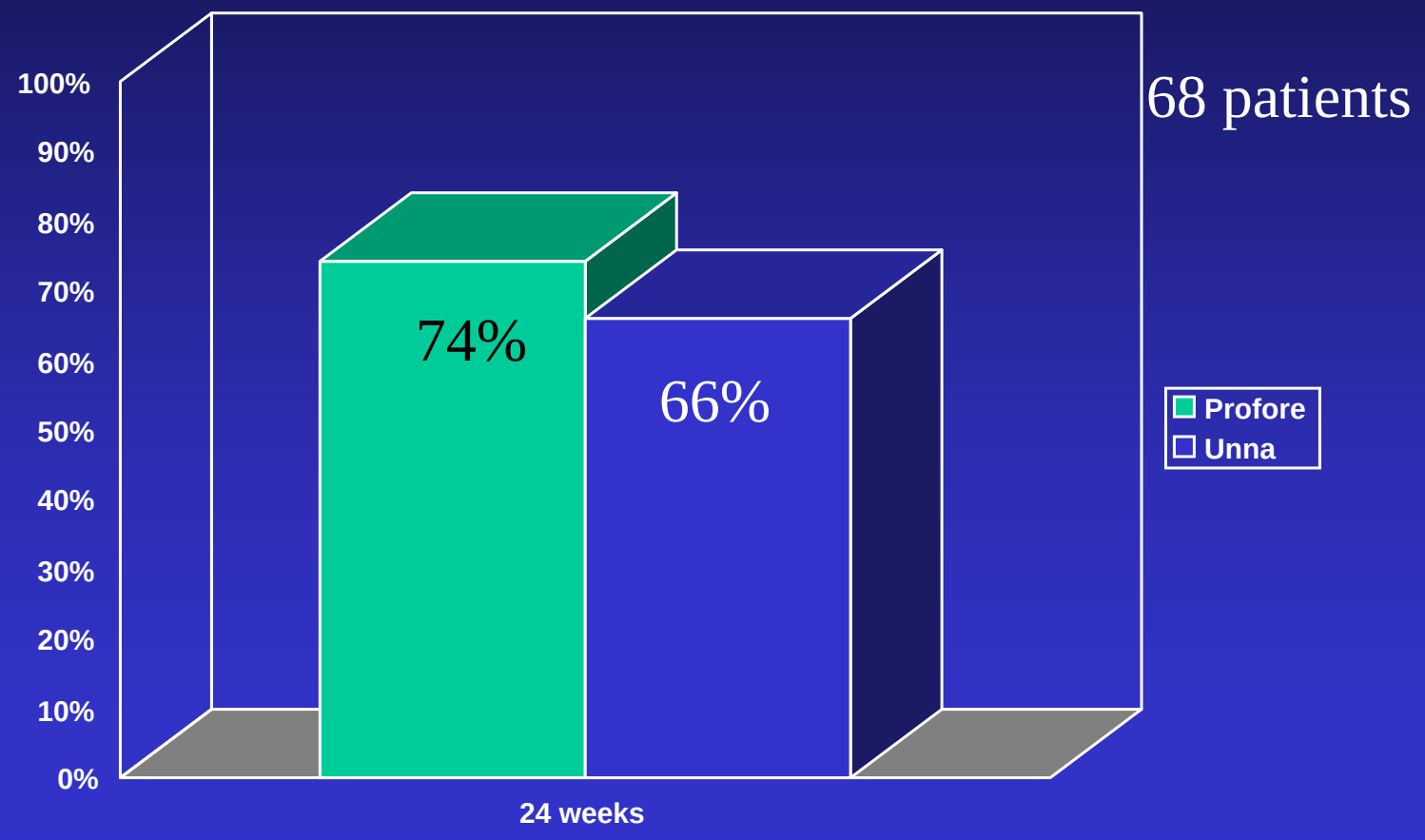
P < .01

Randomized trial of 4-layer elastic compression (Profore) versus 2-layer elastic compression (Surepress)



Moffatt et al, Wound Rep Regen 2003

Randomized trial of 4-layer elastic compression (Profore) versus Unna's paste boot



68 patients

Polignano et al, J Wound Care 2004;13:21-4

High strength compression for VLU

Just do it

Grade	Benefit vs Risk	Methodology	Implications
1A	Clear	High quality	Strong recommendation, Generalizeable
1B	Clear	Moderate	Strong Recommendation, Applies to most patients
1C	Clear	Low	Strong recommendation, but may change w better evidence
2A	Balanced	High quality	Weak recommendation, Action differs w patient/societal values
2B	Balanced	Moderate	Weak recommendation, Action differs with patient/societal values
2C	Uncertain	Low	Very weak recommendation, Alternatives equally reasonable

1C Example: Antibiotics for wound infection

- No studies randomizing patients to antibiotics vs no antibiotics
- Numerous trials comparing different antibiotics

Antibiotics for wound infection

Grade	Benefit vs Risk	Methodology	Implications
1A	Clear	High quality	Strong recommendation, Generalizeable
1B	Clear	Moderate	Strong Recommendation, Applies to most patients
1C	Clear	Low	Strong recommendation, but may change w better evidence
2A	Balanced	High quality	Weak recommendation, Action differs w patient/societal values
2B	Balanced	Moderate	Weak recommendation, Action differs with patient/societal values
2C	Uncertain	Low	Very weak recommendation, Alternatives equally reasonable

Silver dressings for diabetic (or venous) ulcers

- No quality evidence that silver dressings result in improvement in clinically relevant outcome
- May reduce bacterial counts in vivo
- No randomized studies with improved healing or reduced rates of infection
  - Aquacel AG study

Further evidence may alter treatment recommendations

Silver dressings for diabetic foot ulcers

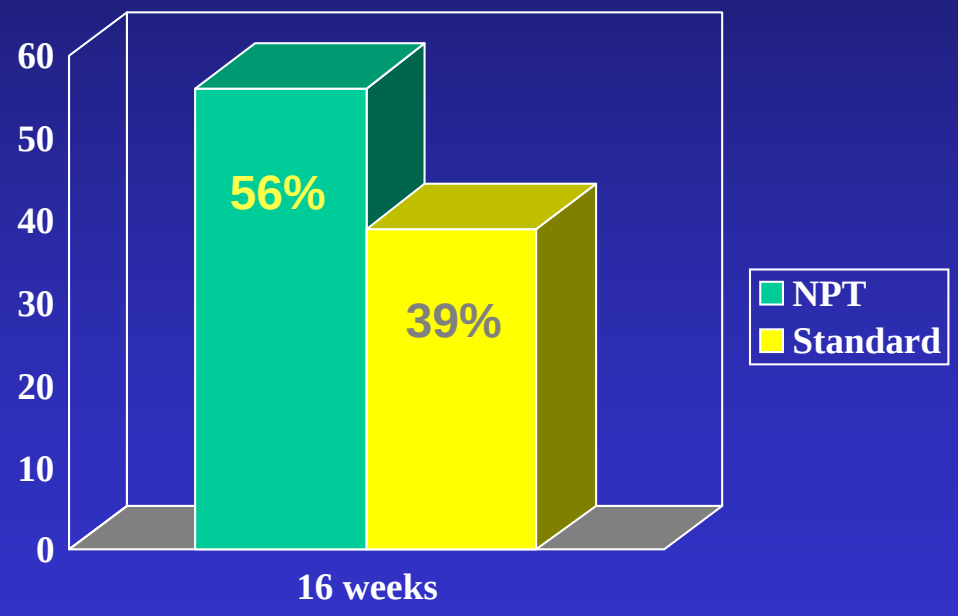
Grade	Benefit vs Risk	Methodology	Implications
1A	Clear	High quality	Strong recommendation, Generalizeable
1B	Clear	Moderate	Strong Recommendation, Applies to most patients
1C	Clear	Low	Strong recommendation, but may change w better evidence
2A	Balanced	High quality	Weak recommendation, Action differs w patient/societal values
2B	Balanced	Moderate	Weak recommendation, Action differs with patient/societal values
2C	Uncertain	Low	Very weak recommendation, Alternatives equally reasonable

NPWT: grade of evidence?

- Prior to 2005 literature search yielded 3 small randomized trials of NPT for chronic wounds
- Total 56 patients

Negative pressure wound therapy after partial diabetic foot amputation: a multicentre, randomized controlled trial.

- Armstrong DG, Lavery LA; Diabetic Foot Study Consortium.
- Lancet. 2005 Nov 12;366(9498):1704-10.
- 162 patients with partial foot amp randomized to NPT or standard care for 16 weeks



$P < 0.05$

NPWT for adjuvant treatment of DFU

Grade	Benefit vs Risk	Methodology	Implications
1A	Clear	High quality	Strong recommendation, Generalizeable
1B	Clear	Moderate	Strong Recommendation, Applies to most patients
1C	Clear	Low	Strong recommendation, but may change w better evidence
2A	Balanced	High quality	Weak recommendation, Action differs w patient/societal values
2B	Balanced	Moderate	Weak recommendation, Action differs with patient/societal values
2C	Uncertain	Low	Very weak recommendation, Alternatives equally reasonable

Studies on HBO for DFU

Study	N	Outcome Variables	Results (Treated vs. Control Group)	Comments
Doctor et al (1992)	30	AR; II	**AR: 2 vs 7** ($p < .05$); II: 19 to 3 vs 16 to 12 ($p < .05$)	**4 treatments over 2 weeks**
Faglia et al (1996)	68	AR; TcPO2	**AR: 3 vs. 11** ($p = .016$); PtcO2: increase 14 vs 5 mmHg ($p = .0002$)	38 ± 8 treatments
Lin et al (2001)	29	PtcO2	PtcO2: 58 vs 36 mmHg ($p < .01$)	30 treatments; only abstract published
Abidia et al (2003)	16	CE; WS	CE: **5 vs 0** (+1 y; $p = .026$); WS: **100% decrease (median) vs 52%** (+6 weeks; $p = .027$)	**30 treatments over 6 weeks**
Kessler et al (2003)	28	WS	WS: **42% vs 22% (2 weeks**; $p = .037$). No difference at 4+ weeks (2 closed HBOT, 0 cntrl).	**20 treatments over 2 weeks**

Listing of RCTs. AR = amputation rate; CE = complete epithelialization (healing); II = infection incidence; TcPO2: transcutaneous oximetry of foot; WS = wound size;

HBO for adjuvant treatment of DFU

Grade	Benefit vs Risk	Methodology	Implications
1A	Clear	High quality	Strong recommendation, Generalizeable
1B	Clear	Moderate	Strong Recommendation, Applies to most patients
1C	Clear	Low	Strong recommendation, but may change w better evidence
2A	Balanced	High quality	Weak recommendation, Action differs w patient/societal values
2B	Balanced	Moderate	Weak recommendation, Action differs with patient/societal values
2C	Uncertain	Low	Very weak recommendation, Alternatives equally reasonable

Impact of RCTs on clinical market

Why don’t wound clinicians respond to products supported by quality data?

Previously reported reasons

- Lack of marketing by wound companies
- Difficulty finding clinicians caring for wounds at advanced centers
- Apathetic or uneducated end-users
  - “I’m waiting for something that really makes a difference”
- Need for visible results from treatment plan
- Inadequate cost-efficacy data

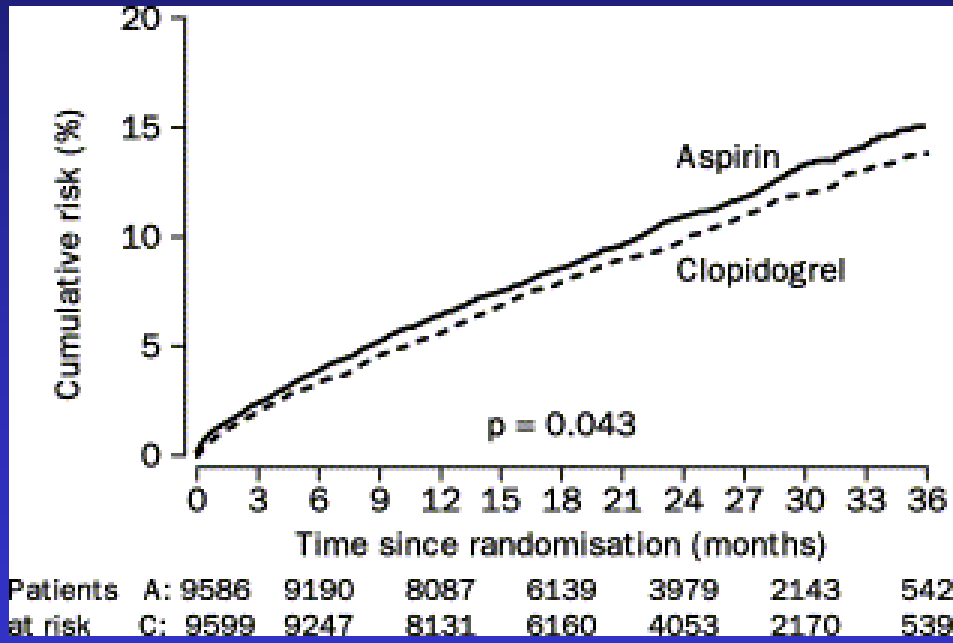
The cardiology model

- Large clinical trials of pharmacologic therapies
- Small absolute but significant benefit
- Express improvement as “relative risk reduction”
- Prescribe to everyone

Caprie trial: Clopidogrel vs Aspirin for patients at risk for ischemic events

Ischemic event

- Asymptomatic MI
- Symptomatic non-fatal MI
- Fatal MI
- Stroke



Caprie trial specifics: Use of RRR

- Nearly 20,000 patients enrolled
- Incidence of ischemic event at 2 yrs
  - Aspirin 9.6%
  - Clopidogrel 7.9%

- Absolute reduction in event incidence 1.7%
- **Relative risk reduction** **18%**
- Cost
  - Aspirin $2-3/month
  - Clopidogrel $60-80/month

Apply RRR to wound healing

- Dermagraft DFU pivotal trial
  - 18% healed SOC
  - 30% healed with DG
  - **60% relative increase in healed wounds**

- Apligraf DFU pivotal trial
  - 38% healed SOC
  - 56% healed with AG
  - **68% relative increase in healed wounds**

Number needed to treat

- To prevent an ischemic event
  - Need to treat 59 patients with clopidogrel instead of ASA
  - Cost for each prevented event $85,000

- To heal one additional wound
  - Need to treat 8 patients with DG or 6 patients with AG
  - Cost for each additional healed wound
  - $15,000 - 25,000 depending on # used per pt

What is the difference between treating CAD and healing a chronic wound?

Summary: Evaluation of evidence

- Grade criteria not levels of evidence
- Incorporate patient condition, values, treatment desires in therapy selections
- All RCTs are not created equally
- Beware the investigator with a big stake in the game