RESTORATION LONGEVITY and CLINICAL EVALUATION

S.C. BAYNE

UM, School of Dentistry
Ann Arbor, MI 48109

The real test of treatment effectiveness is the long term success of the procedure or restoration. **[CLICK]** In this module, we will introduce you to a way of describing restoration longevity and then look at how clinical evaluation is used to track longevity.

EVIDENCE-BASED DENTISTRY

What is the source of the truth? How much is there?

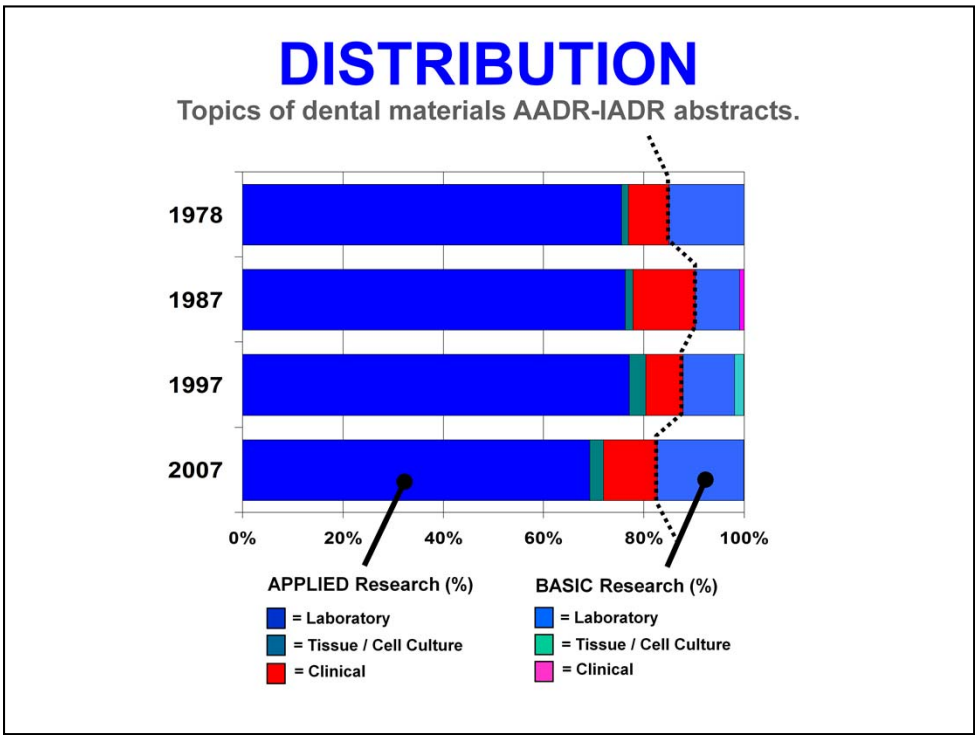
The good news is that 10% of everything in practice may be evidence-based.

The bad news is that we "don't know" which 10% it is.

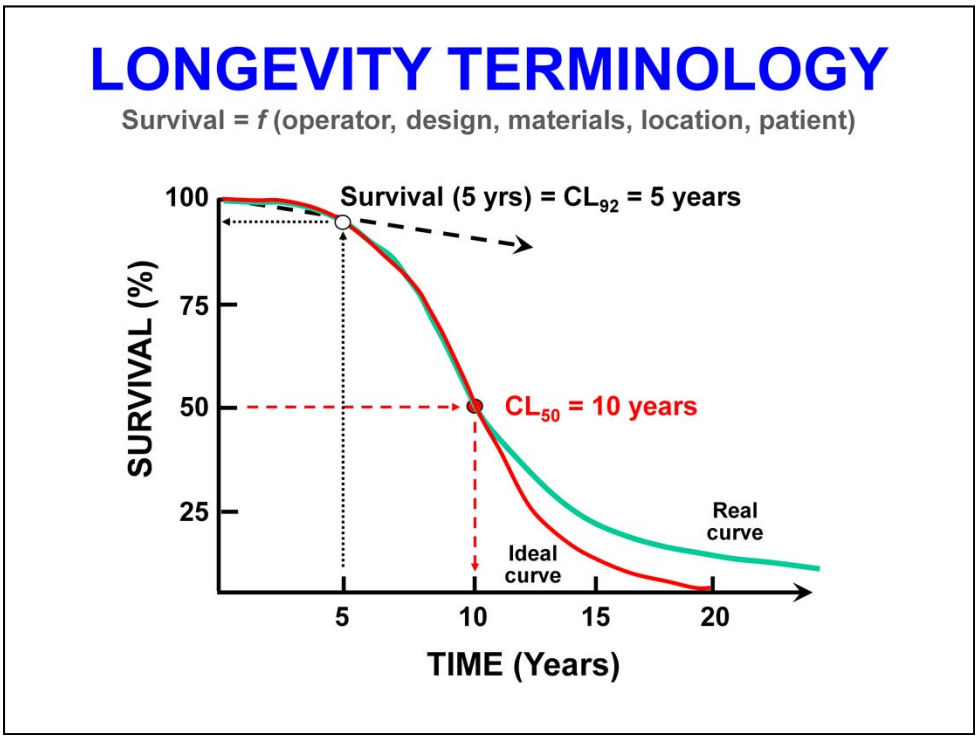
Clinical research and evidence-based dentistry go hand-in-hand. **[CLICK]** What is the source of your truth for evidence-based dentistry? Why TV, of course, …… **[CLICK]** and CNN is the best of all. Ha!

How much evidence is there for the average things that are performed in general dental practice each day? **[CLICK]** The good news is that about 10% of everything in practice may be evidence-based. **[CLICK]** The bad news is that we "don't know" which 10% it is.

[CLICK] Now let's look at how much evidence we really want. It would be wonderful to have 100% evidence, **[CLICK]** and at the moment we only have 10%. **[CLICK]** In reality it takes 10 or more years to do good clinical trials to determine good evidence. **[CLICK]** Therefore, we are always working with newly emerging information for the other part.



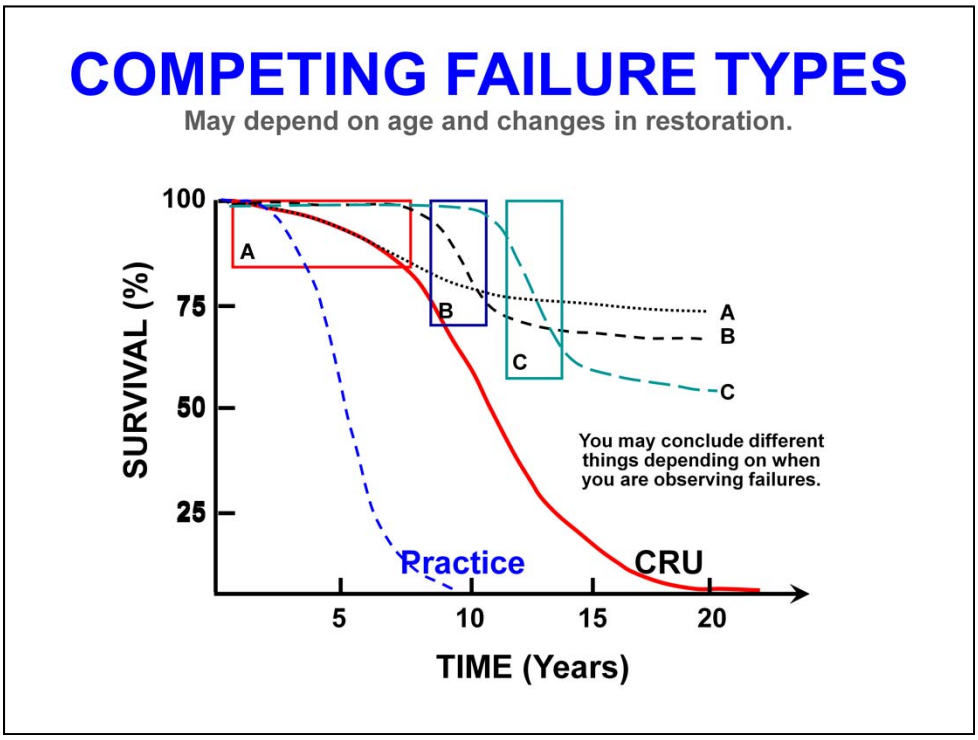
There is further proof of this problem of missing clinical evidence from the DMG abstracts at research meetings at the International Association for Dental Research for the years 1978, 1987, 1997, and 2007. **[CLICK]** The bars show applied research to the left **[CLICK]** and basic research to the right. Clinical components of these are shown as RED or PINK. Only around 10% of the overall DMG research activities involved clinical research.



Clinical performance generally is represented in terms of survival of a pool of restorations versus time. Imagine plotting a curve that starts at 100% and decreases with time. **[CLICK]** This is called a survival or success curve. If it starts at 0% and increases with time it is a failure curve. Both show the same events. The red curve is a reverse s-shaped curve. **[CLICK]** An actual curve is typically stretched out at the end like the green curve.

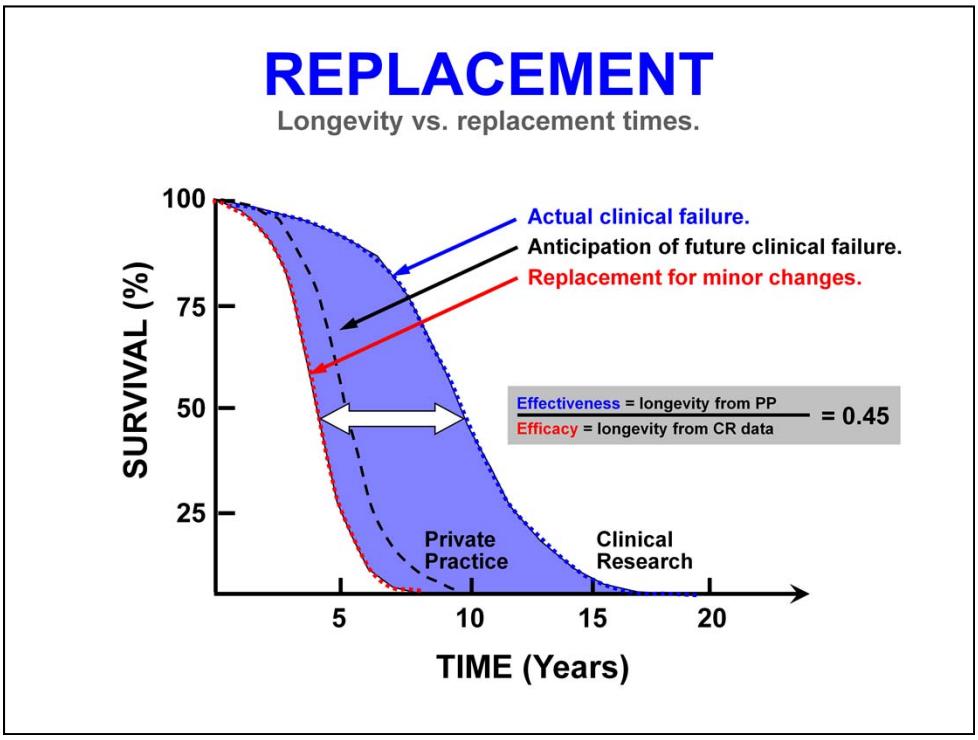
Since it is often inconvenient to report the entire curve, one may choose a key point on the curve to report instead. **[CLICK]** The most important point is the CL50 or half-life for the pool of restorations. Another way of stating this is that it is the time at which half of the restorations are still surviving or at which half of the restorations have failed. Since most clinical trials are not conducted for very long times, very little of the actual longevity curve is known – and the CL50 is rarely known. **[CLICK]** Therefore, some earlier point that is more convenient is reported. For a material with a survival rate of 92% at 5 years, one would report CL92=5 years.

[CLICK] One of the errors that many investigators make -- is to linearly extrapolate downfield events from a small amount of clinical data. Be careful.



A longevity curve represents the sum of many contributions to clinical failure and these may occur at different rates at different times. **[CLICK]** In the case shown above, mechanism A dominates at first. **[CLICK]** Mechanism B then begins to operate. **[CLICK]** In later stages of failure, mechanism C is the main one. These mechanisms might involve secondary caries, wear, and bulk fracture. The RED curve is the sum of the three dotted curves shown in the figure.

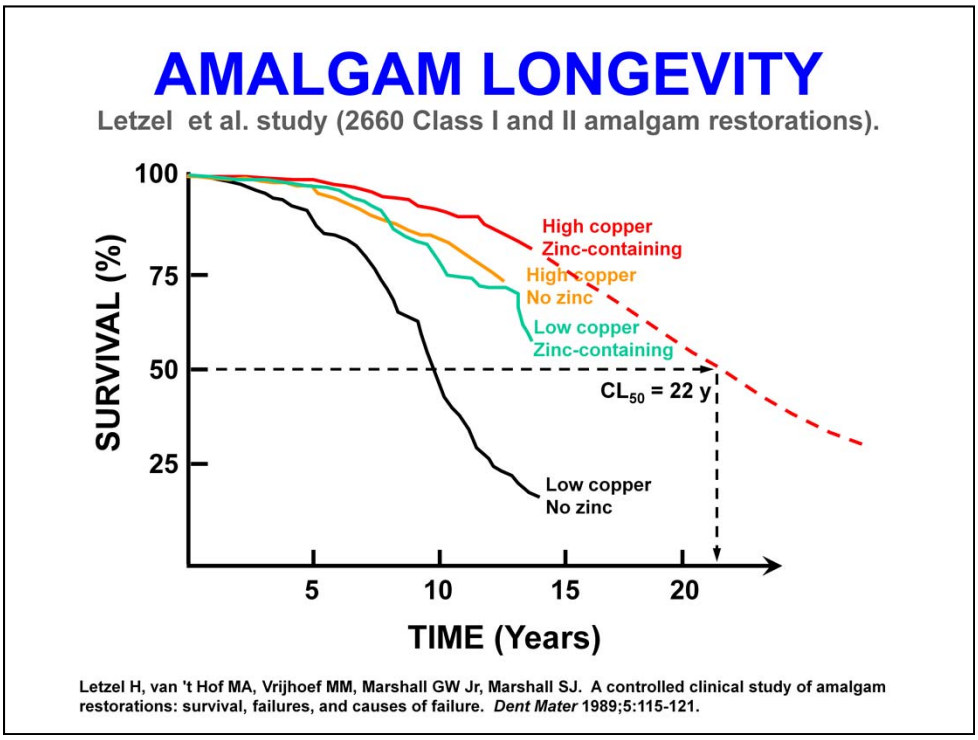
The actual curve for clinical practice may occur much sooner that this. **[CLICK]** Is this related to other types of failures? No. Let's take a closer look at what types of decisions are involved.



While we can calculate an average longevity curve for all private practitioners, each individual has their own curve. The curve for a specific dentist may be shifted toward shorter or longer times.

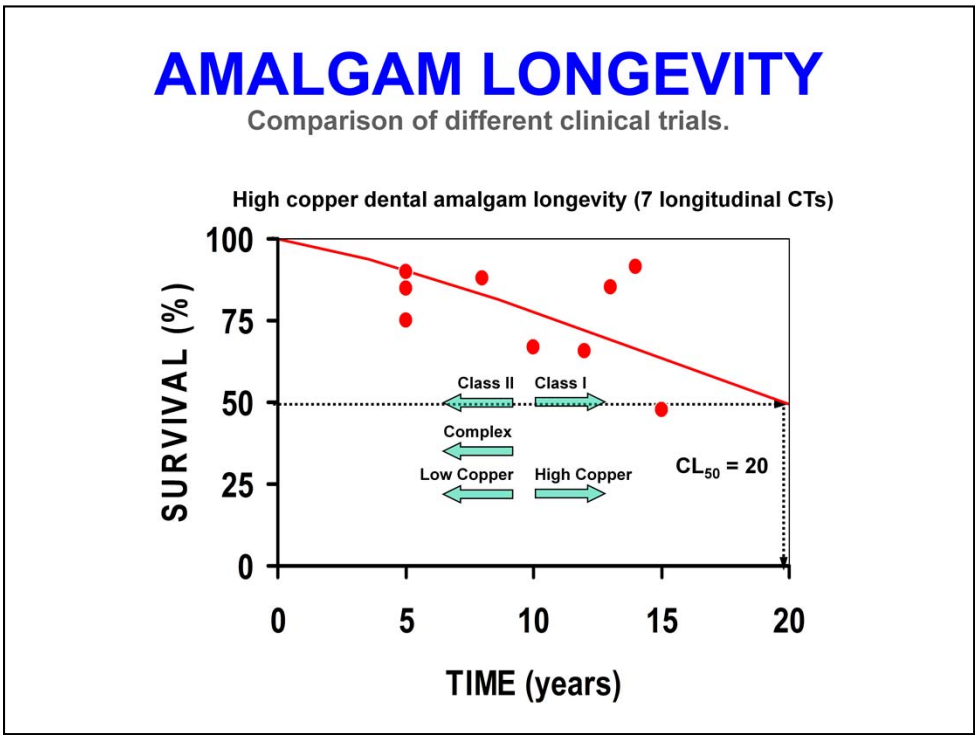
Look at the **[CLICK]** range of situations occurring in a private practice. **[CLICK]** The LD50 for a dentist who only places restorations when true clinical failure occurs is much greater than the average in clinical practice. **[CLICK]** Another dentist might "replace" restorations by trying to anticipate their failure. **[CLICK]** Even another dentist may replace restorations for minor changes in quality. This last situation is more typical in countries with national health insurance and minimal requirements to qualify for compensation.

[CLICK] The actual replacement time is called the "effectiveness." The true longevity value is often called the "efficacy." The ratio of effectiveness to efficacy is about a one-half. If a restoration lasts 18 years in a clinical trial, it will be replaced typically at 9 years in clinical practice.



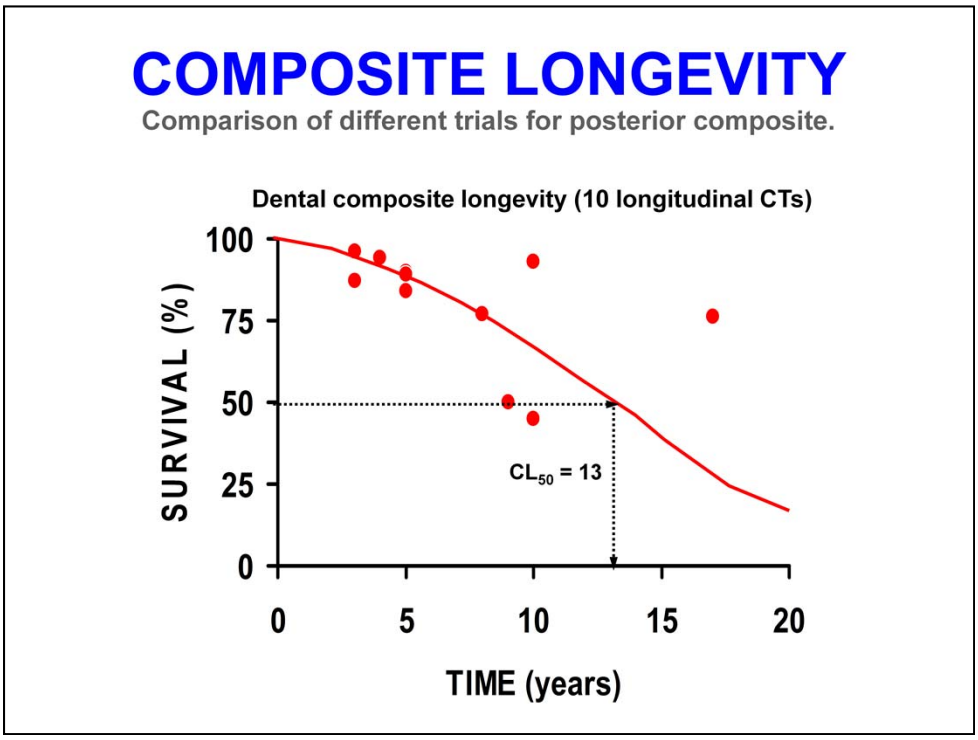
Shown above are the longevity curves for dental amalgam for a study done by Letzel et al. looking at the differences between low and high copper alloys. This study involved 2660 Class I and II dental amalgam restorations over 13-14 years. **[CLICK]** High-copper zinc-containing compositions performed the best. **[CLICK]** Low-copper amalgams were much worse. Despite the long term recalls involved, only part of the middle portion of the s-curve is beginning to appear. Zinc appears to improve the performance. **[CLICK]** High copper amalgam with no zinc appeared to perform slightly worse than its counterpart. **[CLICK]** Low copper amalgam with zinc was substantially better than its counterpart.

Taking some latitude, **[CLICK]** I have extrapolated the curve for high copper zinc-containing amalgam to the point that **[CLICK]** we can determine an CL50 value of about 22 years. If this is the efficacy, then the effectiveness for clinical practice would be about 45% as much or around 10 years which seems to be true.



This same study is now shown as part of a larger collection of 7 amalgam clinical trials representing survival rates from 7-15 years. If the total information is been treated as a homogeneous pool, a survival curve can be estimated for the pool. **[CLICK]** If all these assumptions were valid, then for dental amalgam, the CL50 would be about 20 years. **[CLICK]** Of course, there are many other factors that are not being carefully considered here.

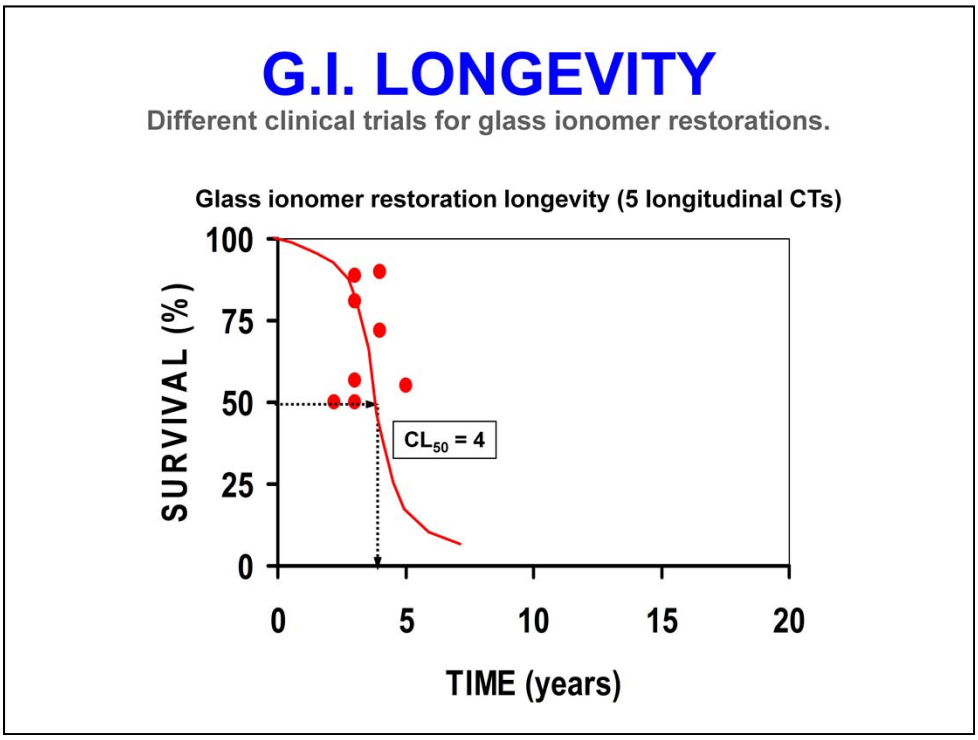
[CLICK] As shown by the arrows, amalgams that are Class II, complex, and/or low-copper types tend to drag the curve to the left and decrease the CL50. Class I restorations, made from high-copper amalgam, tend to push the curve to the right. Most studies have a complicated mixture of these variables in their designs.



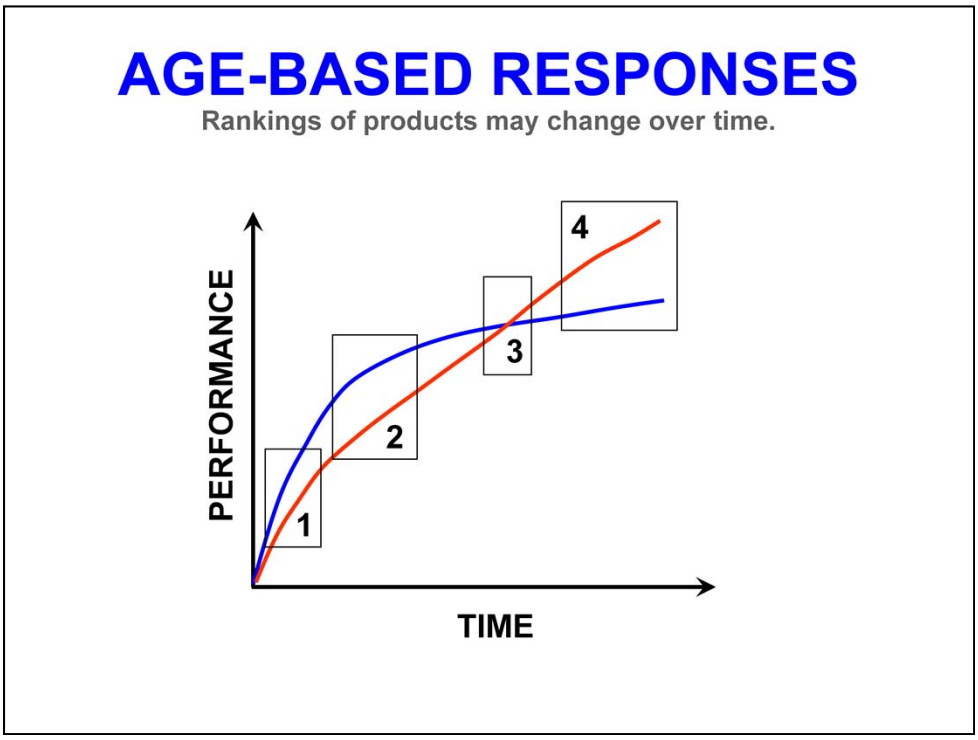
Intuitively, you would expect dental amalgam to last longer than dental composite. But dental composite may be better than you think.

Here are some discrete examples of 10 longevity studies reported in the literature for dental composite. The first thing that you will notice is that there is substantial scatter of the survival rates. It is very important to remember that the points toward the right represent the oldest versions of composites. The 17-year point is for UV-cured composites as reported by Wilder. Newer versions of composites have only short-term clinical trials and are collected to the left-hand side of the graph. **[CLICK]** The survival curve represents data collected over a broad range of composite compositions, curing strategies, associated bonding systems, and surface finishing techniques. **[CLICK]** Even with all of these assumptions, the estimated survival curve demonstrates a CL50 = 13 years. For this same collection of materials, one might guess that the effectiveness in private practice could be about 6 years.

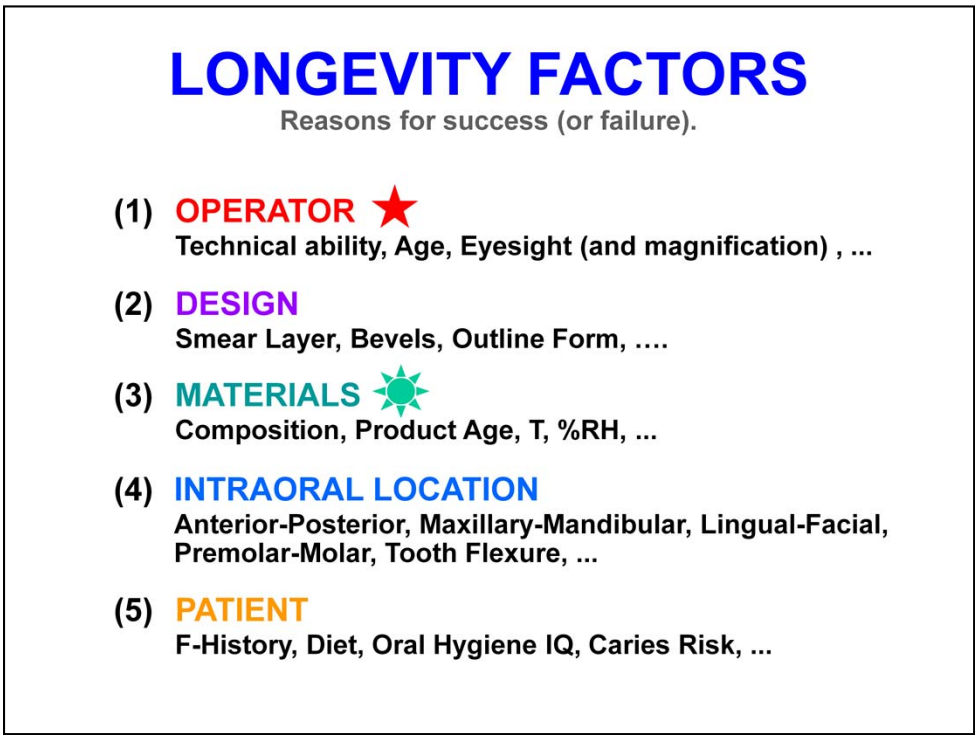
Of course, if we only considered the composites of the last decade or so, then the CL50 would be even much better.



Glass ionomers as Class I and II restorations have not survived as well as amalgams or composites. Shown above is a collection of data from 5 longitudinal studies. **[CLICK]** An estimated longevity curve is included **[CLICK]** with a projected CL50 of about 4 years. Most of these restorations are replaced for reasons of secondary caries. Despite the poor average performance, glass ionomers have done well in the hands of a few dedicated operators that have perfected the techniques of using them.



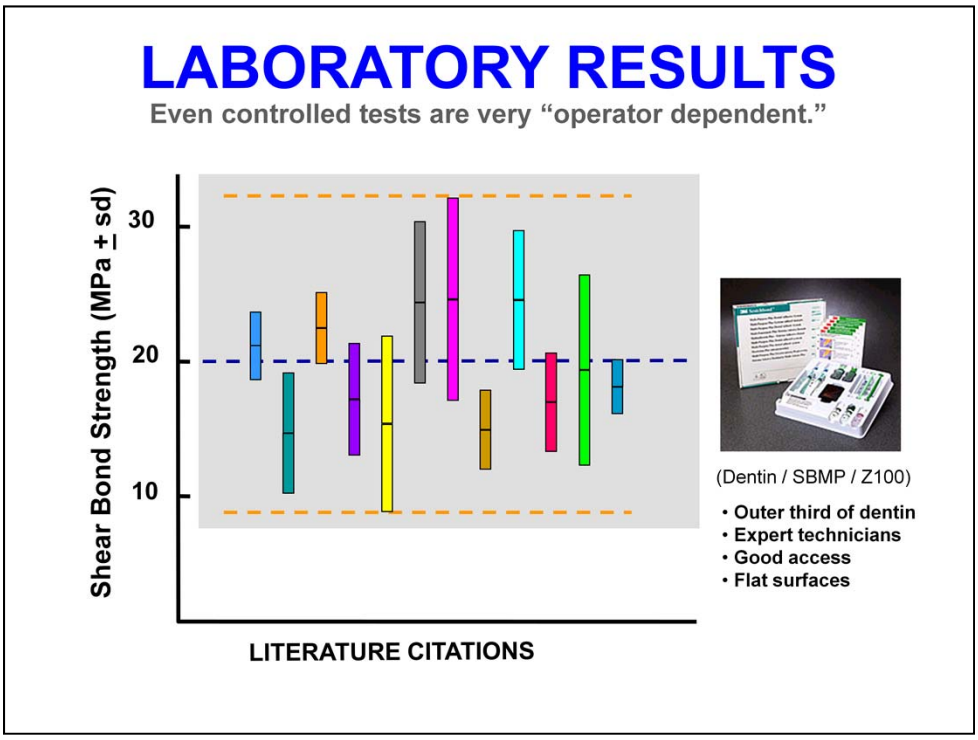
Different things do not necessarily demonstrate the same patterns of behavior over time in clinical trials. Consider the following two curves, **[CLICK]** Blue and **[CLICK]** Red. **[CLICK]** At first the performance of the blue curve is better. **[CLICK]** At a later time, blue prevails. **[CLICK]** At an even later time, both are equal. **[CLICK]** Finally, at a very late time, red begins to out perform blue. If you need to make a judgment about which treatment is better, beware that choices made for one recall time may not predict future recall times.



Longevity depends on 5 major classes of factors.

(1) **[CLICK]** Operator factors include things such as the individual's technical ability, age, and eyesight limitations.
(2) **[CLICK]** Cavity design factors depend on the outline form and surfaces of the preparation.
(3) **[CLICK]** Materials factors include the procedural details and inherent material's properties.
(4) **[CLICK]** Intraoral location factors are important to understanding clinical risk. Restorations that are large and in first molars are at much more risk than small anterior ones.
(5) **[CLICK]** Finally, patient factors are associated with F history, diet, oral hygiene IQ, and other risks for secondary caries.

[CLICK] Of all of these, "operator factors" are by far-and-away the most important. Generally, we estimate that they affect 50% of the longevity. **[CLICK]** The least important are the "material's factors."



Even in well controlled laboratory situations, operator effects are large. Here is a good example. **[CLICK]** The published bond strengths measured in several dental research laboratories under standardized and controlled conditions are reported on the figure above for 3M's Scotchbond MultiPurpose used to bond Z100 composite to dentin.

[CLICK] A typical value for shear bond strength of bonding systems is 20 MPa. Each bar reports the mean value for a different reported study plus or minus the standard deviation. **[CLICK]** If you look at the combined scatter (see the orange lines), you will notice that the highest value is 33 MPa and the lowest is 9 MPa. **[CLICK]** This represents tremendous variation for ideal conditions in a laboratory testing environment. If there is this much variation in the laboratory, imagine how much variation occurs under more poorly controlled clinical conditions.

LONGEVITY RISK FACTORS

Hypothetical calculation of risk factors.

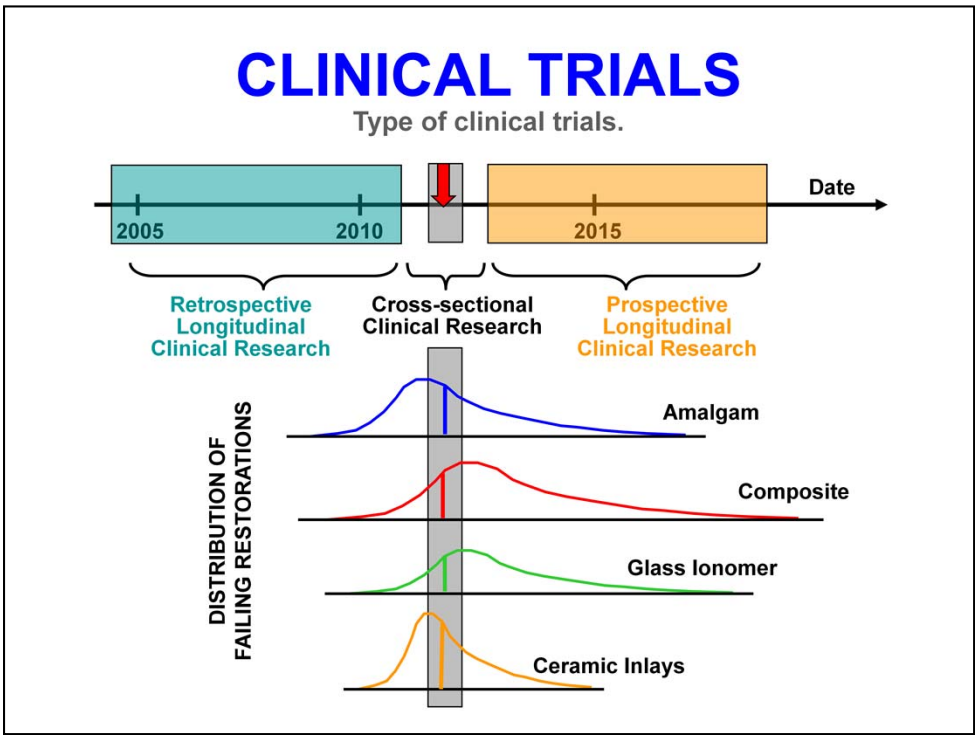
Clinical Factors	Relative Risk	Practitioner Risk-Aversion	Estimated Practice Longevity
Operator	50%	0.50	0.25
Design	10%	0.60	0.06
Material	10%	0.50	0.05
Site	25%	0.30	0.08
Patient	5%	0.20	0.01
	100%		
Effectiveness =			*0.45*

Every practitioner will have a different personal “relative risks” and different levels of personal “risk-aversion’ while doing restorative dentistry.

[CLICK] As an academic exercise, average relative risks were assigned to each of the 5 categories of clinical factors. The operator is clearly the most important category.

[CLICK] Within an actual practice, a dentist can not select which patients will be treated, but he or she may decide on how much average risk to undertake. A value of 1.0 is total risk-aversion and represents treating only the most conservative and safest conditions. 0.00 represents the highest risk.

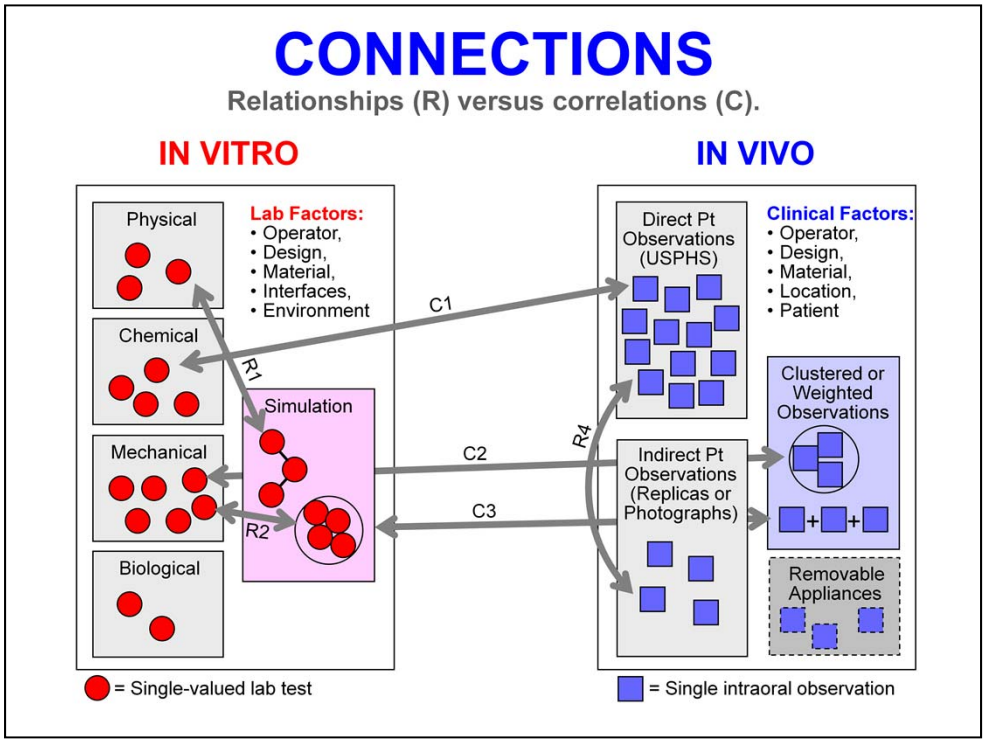
[CLICK] Multiplying the relative risk by the risk-aversion lets us estimate contributions to average longevity in practice which is called “effectiveness”. **[CLICK]** This is generally just less than half of what is reported in longitudinal clinical trials.



Longitudinal clinical trials follow a controlled population of patients and restorations over time. **[CLICK]** These can be conducted prospectively **[CLICK]** or retrospectively. Prospective trials are generally much better and guarantee that the appropriate controls and standardization are in place.

[CLICK] Other clinical trials are conducted as cross-sectional ones in which all patients and their restorations are examined in a short period of time as evidence of success or failure. These trials are relatively easy to conduct but are generally missing most of the key information about the patients and the variables associated with placement of the original restorations.

[CLICK] The results of cross-sectional trials are different than for longitudinal trials. **[CLICK]** Cross-sectional trials survey patients and restorations in different stages of their distributions of failure. Very rarely do these effects average out properly. Thus, cross-sectional trials tend to report failure as faster by a factor of 2-to-3 times as much.



One of the great frustrations for restorative materials research is that there is almost no correlation of in vitro (or laboratory, on the LEFT) and in vivo (or clinical, on the RIGHT) research outcomes. Both **[CLICK]** are affected by factors surrounding the creation of the restorations.

Properties or measurements can be considered singly or grouped together in simulations. **[CLICK] [CLICK]** One can try to establish relationships **[CLICK] [CLICK]** between different single properties and simulations, or **[CLICK]** between different clusters of observations. Ultimately what one would like to identify is correlations **[CLICK] [CLICK] [CLICK]** between in vitro and in vivo situations.

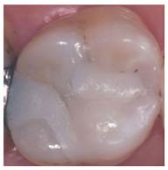
Standards-Based Controlled Trials

United States Public Health Service (USPHS) categories.

USPHS Categories:	Alternative Designations:
• Color match (c.m.)	Color stability
• Marginal discoloration (m.d.)	Marginal staining; Interfacial staining
• Secondary caries (s.c.)	Resistance to secondary caries
• Anatomic form (a.f.)	Occlusal wear; wear; resistance to wear
• Marginal adaptation (m.a.)	Marginal integrity; marginal ditching
• Surface texture (s.t.)	Surface roughness

Criteria (Rating Scale):

A	= alfa	= clinically "ideal"
B	= bravo	= clinically "acceptable"
C	= charlie	= clinically "unacceptable"



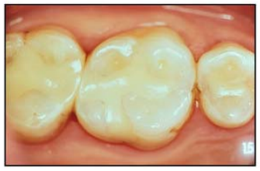
Cat.	A	B	C
c.m.	80	20	0
m.d.	82	18	0
s.c.	98	---	2
a.f.	93	7	0
m.a.	91	9	0
s.t.	91	9	0

Clinical trials of restorative dental materials rely on comparison of test materials to a standard scale for several categories of interest. The original process was developed by Cvar and Ryge in the 1960s from their work at the United States Public Health Service (USPHS). **[CLICK]** The original categories (and alternative descriptions) are shown. **[CLICK]** These have been expanded considerably in recent years. **[CLICK]** For each category the scale was A = alfa = clinically ideal; **[CLICK] [CLICK]** B = bravo = clinically acceptable; and **[CLICK] [CLICK]** C = charlie = clinical failure.

[CLICK] The number of A, B, and C ratings are then tabulated for the different categories for any single recall.

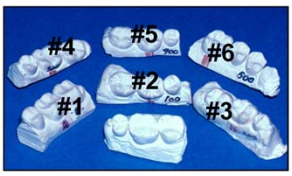
CLINICAL WEAR EVALUATION

Leinfelder indirect wear evaluation.



OCCLUSAL WEAR DIFFICULT TO SEE DIRECTLY.

OCCLUSAL WEAR APPARENT IN CLINICAL CASTS.



Conversion Scale for CAST Ratings

CAST #1	Category 1.0	=	0 μm
	Category 1.5	=	46 μm
CAST #2	Category 2.0	=	92 μm
	Category 2.5	=	156 μm
CAST #3	Category 3.0	=	221 μm
	Category 3.5	=	272 μm
CAST #4	Category 4.0	=	322 μm
	Category 4.5	=	352 μm
CAST #5	Category 5.0	=	382 μm
	Category 5.5	=	438 μm
CAST #6	Category 6.0	=	493 μm

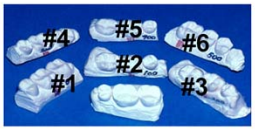
One of the most popular and useful indirect evaluation methods is the Leinfelder Method for estimating occlusal wear. Since it is extremely difficult to visually assess wear intra-orally, an impression is taken to produce a cast **[CLICK]** where the wear is more obvious.

The cast is compared to a series of 6 standard casts **[CLICK]** that represent approximately 100 μm steps. **[CLICK]** Each cast is compared to the standards and then rated as equal to a standard or as falling between two of the adjacent standard casts. **[CLICK]** Then the rating is converted into an estimate of the wear in microns as shown here.

CLINICAL WEAR EVALUATION

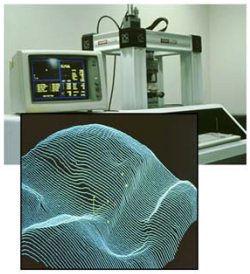
Comparison of accuracy of different methods.

Leinfelder Method

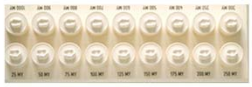


M-L Method

Laser Scanner



Vivadent Method

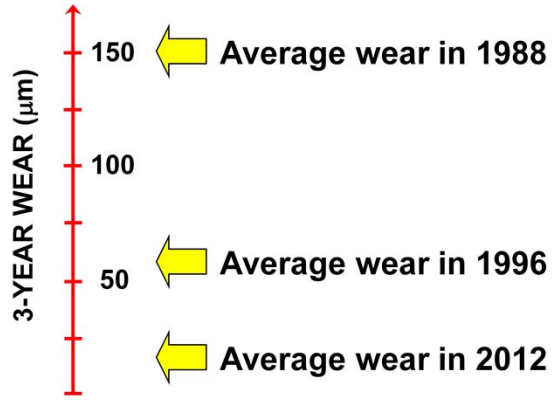


Name:	Scale:	Steps:	Application:	Precision:	Accuracy
Leinfelder Method	0-500 μm	11	Visual	± 100 μm	± 50 μm
Vivadent Scale	0-1000 μm	18	Visual	± 25 μm	± 50 μm
M-L Scale	0-500 μm	18	Visual	± 25 μm	± 50 μm
Laser Scanner	0-500 μm	- -	Automatic	± 1 μm	± 50 μm

There are several other methods of indirect wear measurement that have been examined as well. **[CLICK]** The M-L standards use small round cylinders with depressions of standard depths as a substitute for the Leinfelder casts. However, the M-L ratings generally underestimate wear. **[CLICK]** The Vivadent method uses a series of standard Class I restorations in a molar that are increasingly deeper and deeper. However, these are not quite as reliable as the Leinfelder rating method either. **[CLICK]** Finally, more elegant methods exist such as laser scanning -- to digitize the surface and to rate wear. However, digital images at high magnification demonstrate that most surfaces do not have sharp transitions at margins or edges. Therefore, computer interpretation of the images is complicated if not impossible.

Despite differences in precision among these methods, all methods have about the same accuracy. The laser scanning method is extremely expensive and therefore not commonly used.

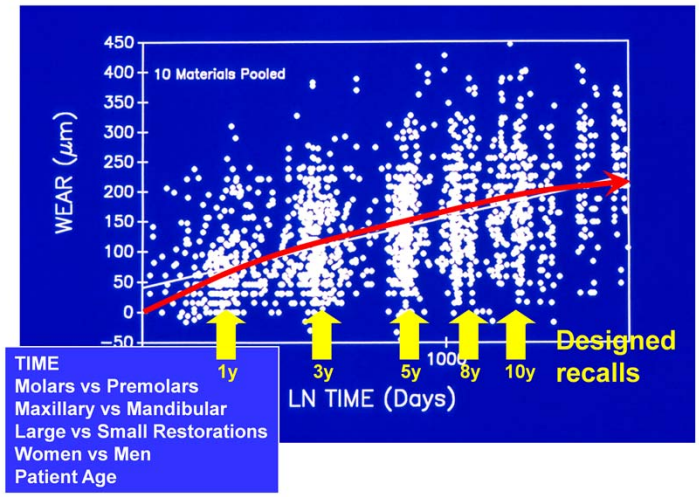
Name:	Scale:	Steps:	Application:	Precision:	Accuracy
Leinfelder Method	0-500 μm	11	Visual	± 100 μm	± 50 μm
Vivadent Scale	0-1000 μm	18	Visual	± 25 μm	± 50 μm
M-L Scale	0-500 μm	18	Visual	± 25 μm	± 50 μm
Laser Scanner	0-500 μm	- -	Automatic	± 1 μm	± 50 μm



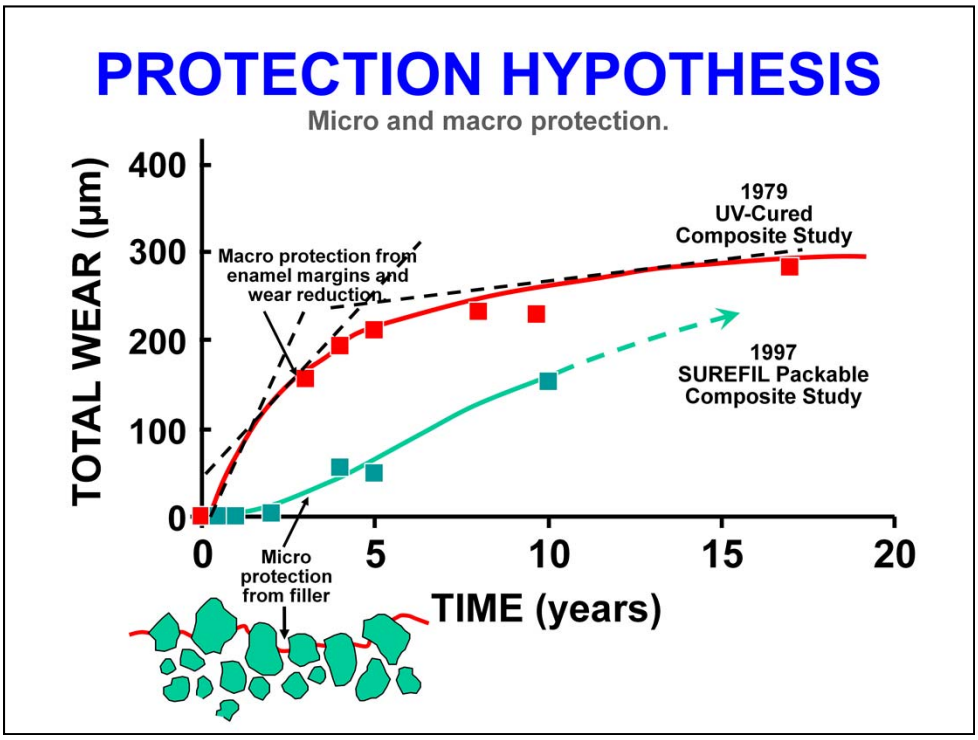
The variation in recorded measurements (or accuracy) shows the limit of these scales. Until recently, most composites underwent wear in clinical trials that was typically in the range of 50-150 μm over 3 years. **[CLICK]** Therefore, scales such as the Leinfelder Method could provide reasonable estimates of wear. However, at the present time the wear of newer materials is substantially lower. The variability associated with any measurements is so great that it is almost impossible to compare different materials at this time.

LONG-TERM CLINICAL TRIALS

Advantages.



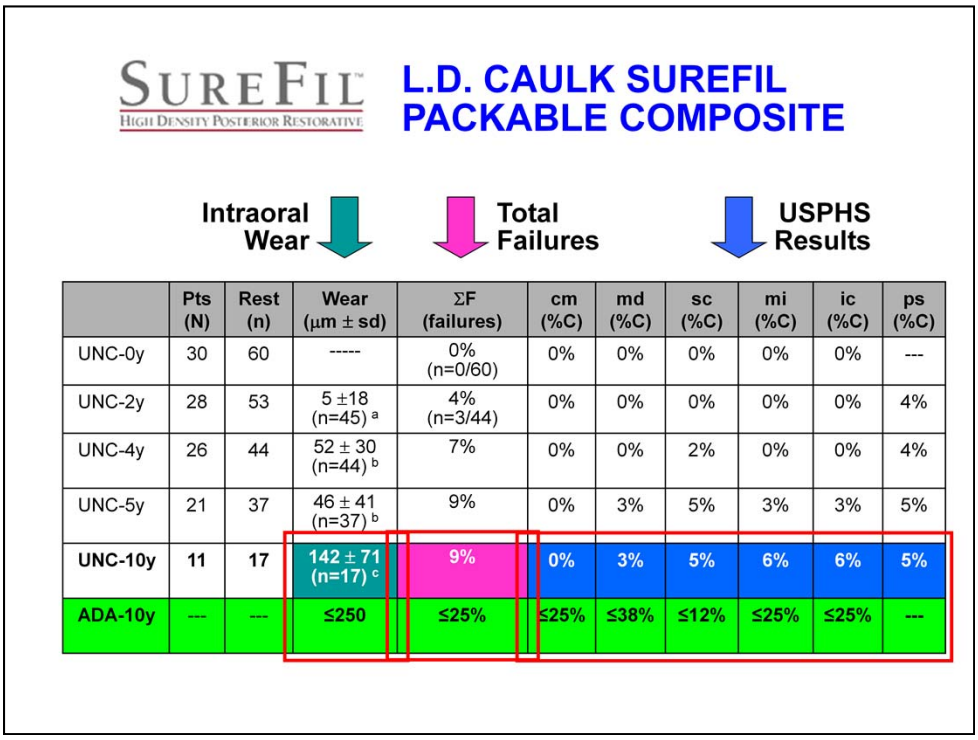
One of the great advantages of a longitudinal trial is that the effect of time can be considered on all the possible variables. **[CLICK]** In the figure above, Leinfelder wear values for a pool of 600 restorations involving multiple recalls is shown. **[CLICK]** Wear rates decreased over time. Time was the most important variable. **[CLICK]** Restorations in molar teeth showed the most wear. Maxillary teeth generally wore more than mandibular teeth.



For conservative composite restorations (<1/3rd intercuspal distance and no cusp capping), the pattern of wear over the long term (for UV-cured composites) has been established by Dr. Al Wilder as decreasing with time. **[CLICK] [CLICK] [CLICK]** Recently, the wear behavior of a packable low-wear composite has been reported up to 10-years by the same team.

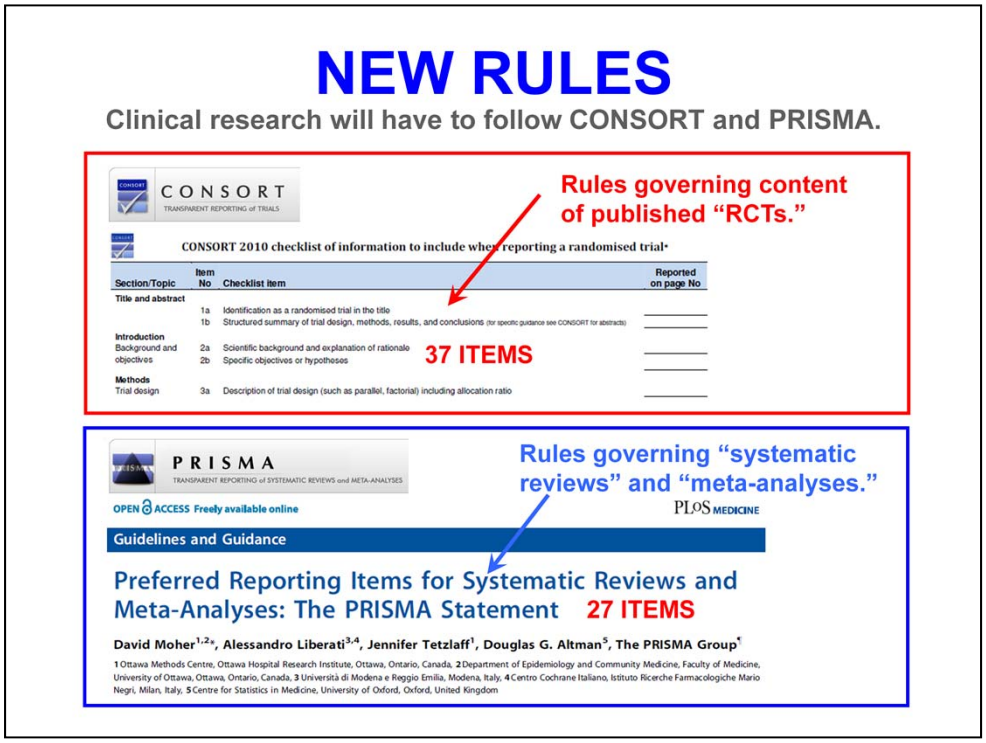
[CLICK] Micro-protection occurs as the filler particles come closer and closer together in systems with more filler, smaller particles, and more well-packed ones. It involves just the wear of the soft resin matrix between the hard filler particles. Newer composites with much narrower spaces between particles have much lower wear rates. **[CLICK]** Macro-protection occurs, as restorations wear a little bit, and the lowered surface becomes sheltered by the remaining cavosurface margins of the preparation making further abrasive contact of lower force and frequency. Therefore, with time, the wear rate decreases to a point that total sheltering occurs. For the restorations followed by Dr. Wilder, this level is about 250 µm after 3-5 years. Posterior composite restorations do not wear out and do not expose dentin.

The foundation of the Protection Hypothesis was originally suggested by Dr. Jorgensen (Sweden) and later demonstrated by Dr. Bayne with clinical trials. Wear is typically measured with reference to the exposed cavity preparation margins. There are laboratory methods to simulate clinical wear that have some correlation with observed clinical values. These are frequently used as an alternative to clinical trials, but are not very good for accurately comparing newer versions of composites that wear very slowly.



	Pts (N)	Rest (n)	Wear (μm ± sd)	ΣF (failures)	cm (%C)	md (%C)	sc (%C)	mi (%C)	ic (%C)	ps (%C)
UNC-0y	30	60	-----	0% (n=0/60)	0%	0%	0%	0%	0%	---
UNC-2y	28	53	5 ±18 (n=45) [a]	4% (n=3/44)	0%	0%	0%	0%	0%	4%
UNC-4y	26	44	52 ± 30 (n=44) [b]	7%	0%	0%	2%	0%	0%	4%
UNC-5y	21	37	46 ± 41 (n=37) [b]	9%	0%	3%	5%	3%	3%	5%
UNC-10y	**11**	**17**	**142 ± 71 (n=17) [c]**	**9%**	**0%**	**3%**	**5%**	**6%**	**6%**	**5%**
ADA-10y	**---**	**---**	**≤250**	**≤25%**	**≤25%**	**≤38%**	**≤12%**	**≤25%**	**≤25%**	**---**

The posterior composite clinical trial results at UNC are now available for the 10 years. Shown above is the clinical performance compared to the extrapolated ADA standard for an amalgam-substitute material. Note that Surefil passes in all the categories indicated in green **[CLICK]**, pink **[CLICK]**, and blue **[CLICK]**. The average wear at 5 and 10 years for large cusp-capping restorations was only 46 and 142 microns respectively.



We want more clinical research. However the differences in methodologies and documentation has always made combining their information difficult. Therefore, there are now 2 relatively new rules for clinical research to be published. They must follow the CONSORT and PRISMA rules. What are these two things?

CONSORT = Consolidated Standards of Reporting Trials

These rules require that randomized controlled trials (RCTs) follow very strict requirements for reporting all of their parameters. **[CLICK]** There are actually 37 different steps involved. **[CLICK]** Journals now are requiring that these rules be following for consideration for publication.

[CLICK] PRISMA = Preferred reporting items for systematic reviews and meta-analyses

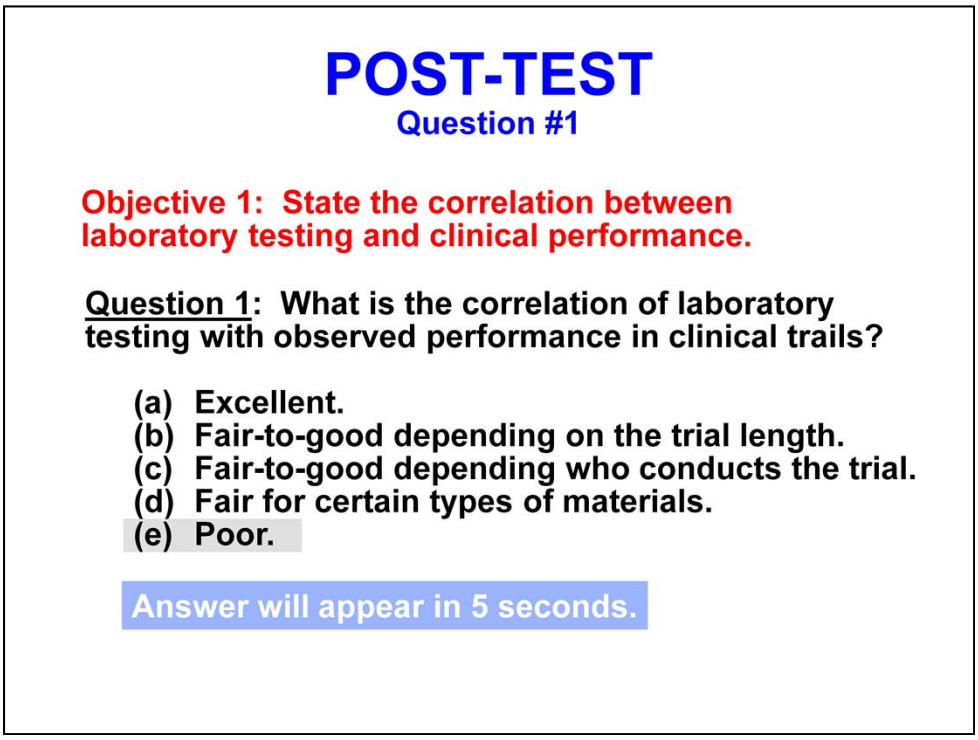
These new rules try to provide a framework for combining information in reviews and involve 27 specific steps.

While they may ultimately help the process of collecting and reporting clinical research, at the moment they seem to be having the opposite effect.

SUMMARY

- Longevity curves are ideal to describe clinical performance.
- Most clinical trials are extremely short (1-5 years).
- Very little clinical research has been done to date.
- Carefully controlled clinical trials can be related to dental practice in terms of "effectiveness."
- Operator factors dominate the 5 categories of factors.
- Clinical research utilizes "direct" (USPHS) and "indirect" evaluations to collect information about restorations.
- Longitudinal trials are better than cross-sectional ones.
- New composites show very low wear rates.

Here is a quick summary of the entire presentation. Longevity curves provide a useful tool to describe clinical performance. Most clinical trials are short, and rarely exceed 5 years. Very little clinical research has been accomplished thus far. Recently, more practice-based research is being accomplished, but the effectiveness is less than half that of more controlled trials. The principal risk for outcomes is due to operator factors. Restorative materials are usually evaluated using direct and indirect observations – and in most cases via longitudinal trials. From examination of many clinical trials of wear of posterior composites over the last 30 years, it can be shown that new composites wear at very low rates.



Question 1: What is the correlation between laboratory testing and observed performance in clinical trials?

POST-TEST

Question #2

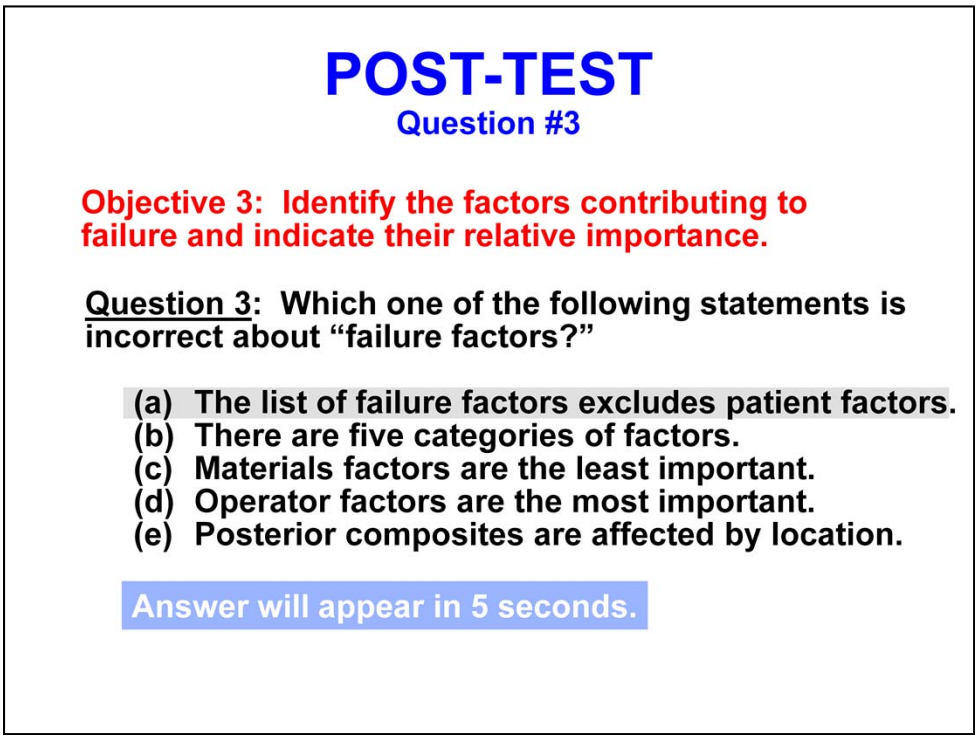
Objective 2: Draw the shape of a failure curve and identify its various characteristics.

Question 2: Which one of the following statements is incorrect about a "failure curve?"

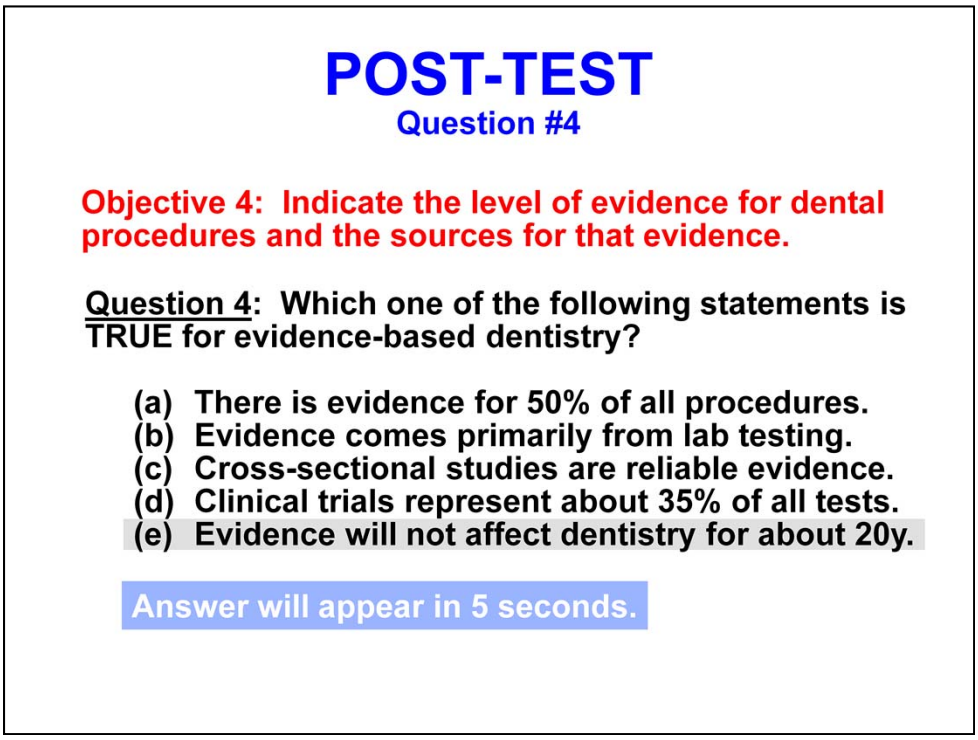
(a) It generally has a reverse "s-shape."
(b) Most curves for real materials are not known.
(c) Curves for trials and practices are the same.
(d) Curves are typically described by the CL50.
(e) Curves represent a combination of failure types.

Answer will appear in 5 seconds.

Question 2: Which one of the following statements is incorrect about a "failure curve?"



Question 3: Which one of the following statements is incorrect about "failure factors?"



Question 4: Which one of the following statements is TRUE for evidence-based dentistry?

POST-TEST

Question #5

Objective 5: Indicate the appropriateness of using posterior composites as permanent restorations.

Question 5: Which one of the following statements is UNIMPORTANT for posterior composites?

(a) Restoration finish predominantly controls wear.
(b) Wear rates tend to decrease with time.
(c) Large molar restorations still perform well at 5y.
(d) Interparticle spacing controls wear rates.
(e) Clinical performance has been tracked 15-20 y.

Answer will appear in 5 seconds.

Question 5: Which one of the followingn statements is unimportant for posterior composites?

THANK YOU.