Session V: Special Topics—Anthony R. Volpe, Chairman

A Critical Evaluation of Some Problems Associated with Clinical Caries Trials by Computer Simulation

K.H. LU and L. VAN WINKLE

Oregon Health Sciences University, Portland, Oregon 97201

J Dent Res 63(Spec Iss):796-804, May, 1984

A natural question one may ask is that if we could observe the caries process in action and conduct live clinical trials, what could a simulated clinical trial tell us? The answer is that we can observe neither the caries process nor a clinical trial very precisely. This is so because all clinical data are obtained through the effort and instrumentation of an examiner. Furthermore, the caries process, as a dynamic process, follows Heisenberg's principle of uncertainty. During the course of a clinical trial, not only the caries process changes, but the examiner changes also, and they influence each other. Unfortunately, a perfect examiner does not exist. However, through time and experience he may become more experienced, although not necessarily better. What we obtained as data was actually a blurred version of the true process through the lens of instrumental limitations and certain human errors. In view of the above uncertainties, a simulated trial where a "true" and a blurred set can be created at will, according to acceptable principles and distributions, will give us a tool to look into the nature of the process and probable solutions to some of the very gnawing problems. There are many interesting aspects; however, due to the constraints of time and space, we will concentrate on the following three problems:

(1) *Measurement.* Is there a more appropriate measure for caries experience than DMFS?

(2) *The effects of professional intervention.* What happens to the caries experience of a population if there is professional intervention? The benefits notwithstanding, from the clinical trial point of view, such intervention produces some fillings, not because the surfaces were carious, but because of the restorative procedures. To what degree does this loss of innocent surfaces further blur the measure of caries experience?

(3) *The effects of diagnostic errors.* What happens to the caries experience of a population if the examination is subject to, say, x-percent of error? The question is innocent enough. The answers, however, are quite complex, because they depend on the kind of error that has been committed. Was it unbiased error, *i.e.,* the error could go either way? Was it biased error, *i.e.,* the error is deliberately committed for some purpose? The consequences of these errors could be quite different.

A brief description of the simulation of a caries process and clinical trial.

Simulation has proved to be quite a humbling task. According to some beliefs, the world was created in six days, and on the seventh day the Creator rested. However, if we view the record of man's own millions of years of evolution, there is ample evidence suggesting that trial and error were at play. What we see today is what was left. But nature does not always bury its mistakes: The evolutionary process is succinctly, but honestly, recapitulated through various stages of embryonic development. By the same token, in order to simulate the current caries process, it is imperative that we simulate the target population from infancy to adulthood and specify the underlying distributions that govern all the changes and turns of this phenomenon, however big or small. In the passages to follow, we shall give a brief description of this procedure.

The procedure of simulation. –

(1) *The selection of an underlying distribution which governs the survival time of dental surfaces.* What we are seeking is a mathematical representation of the distribution of survival time with respect to length of exposure and caries lesions. To put it more explicitly, we are particularly interested in parametric changes associated with the following situations:

(a) the values of parameters which render the survival time independent of length of exposure, and

(b) the values of parameters which make the survival time a function of length of exposure such that a surface that has been exposed to the cariogenic environment longer has less chance of surviving a caries attack than one which has been exposed for a shorter period. Since, from past experience, the survival time could be described very well by the exponential distribution (Carlos and Gittelsohn, 1965), it seems reasonable to use a more general distribution which contains the exponential as a special case. We know that the gamma distribution has this property, and this is the reason for adopting the gamma distribution as a descriptive model of survival time.

The Gamma Distribution. – If random variable x is gamma-distributed with probability density function,

$$f(x, a, r) = \frac{a}{(r-1)!}(ax)^{r-1}e^{-ax} \qquad 0 \leq x,$$

$$= 0 \text{ otherwise}$$

where r is a finite positive integer. Then

(i) if $r = 1$
$P\{x \geq t_i + t_n \mid x \geq t_i\} = P\{x \geq t_j + t_n \mid x \geq t_j\} = Pr\{x \geq t_n\}$

(ii) if $2 \leq r$, $t_j > t_i$
$P\{x \geq t_i + t_n \mid x \geq t_i\} < Pr\{x \geq t_n\}$

(iii) if $t_i, t_j, t_n > 0$, $2 \leq r < \infty$, $t_j > t_i$, then
$P\{x \geq t_i + t_n \mid x \geq t_i\} > P\{x \geq t_j + t_n \mid x \geq t_j\}$

We observe that, for $r = 1$, the gamma distribution reduces to the exponential distribution with probability density function

$$f(x, a, 1) = ae^{-ax}.$$

According to property (i), the survival time is independent of exposure times t_i and t_j ($t_i \neq t_j$), because it is identically equal to that of new surface, $P\{x \geq t_n\}$.

If $r \geq 2$, according to property (ii),

$$P\{x \geq t_i + t_n \mid x \geq t_i\} < P\{x \geq t_n\}$$

The surface which has been exposed for a period of t_i is always less likely to survive a caries attack than a new surface.

Furthermore, according to property (iii), we know that

$$P\{x \geq t_i + t_n \mid x \geq t_i\} > P\{x \geq t_j + t_n \mid x \geq t_j\}$$

It suggests that, as long as the survival time follows a non-exponential gamma distribution, the surface with a longer exposure is less likely to survive caries attack than one with a shorter exposure. We have thus arrived at a theoretical interpretation of the difference in survival time behavior, *i.e.*, there has been a change in the value of r=1 to r≥2. It would be of heuristical importance to locate those factors responsible for such a change.

It can be shown that the associated distribution of caries lesions is the negative binomial distribution.

(2) *The distribution of eruptions and exfoliation of the deciduous dentition.* The times of eruption and exfoliation of each of the deciduous teeth were assumed to be normally distributed with various means and S.D. (Logan, *et al.*, 1933).

(3) *The time lapse between deciduous exfoliation and permanent eruption.* The time lapses between deciduous exfoliation and permanent eruption are assumed to be normally distributed, with means = four months and S.D. = six months.

(4) *The distribution which determines the sequence and time of eruption for the permanent dentition.*

(5) *The onset of caries attack rates specified according to vulnerability of tooth surfaces from existing experimental data.* The caries process stress was applied to the dentition at two years of age.

(6) *The changing nature of survival time due to changes in the immediate neighborhood.* It was assumed that a qualitative change in a surface would also affect the survival time of the adjacent surface of a neighboring tooth. These changes were brought about as follows: When a surface is restored by filling, its neighbor's survival time gains the advantage of being multiplied by a factor with distribution N (2, 1.5); and if a surface is absent from the neighborhood, its neighbor's survival gets an advantage factor with distribution N (4, 1.5). On the other hand, if a surface acquired a new neighbor by eruption, its survival time reverted immediately to its natural expectancy.

From a data sample gathered from 1109 adults, we classified the surfaces according to the respective decay risks as shown in Table 1a. (For convenience, it is converted to categories as in Table 1b.) By manipulating the parameters a and r, we were able to assign the "observed" risks as follows:

DECIDUOUS					
RISK	VH	H	M	L	VL
R	2.000	2.000	3.000	2.000	4.000
ALPHA	1.500	1.000	1.000	0.500	0.800
PERMANENT					
RISK	VH	H	M	L	VL
R	3.000	3.000	3.000	5.000	5.000
ALPHA	1.000	0.500	0.200	0.200	0.100

where VH = very high, H = high, M = medium, L = low, and VL = very low.

(7) *The distribution of examiner's error.* We reiterate that the perfect examiner does not exist. All examiners are subject to a certain error percentage We are assuming that we have a reasonable examiner whose error rate is no more than 10%. The examiner is unbiased in that his errors can go in either direction. Note, however, that in the state SI (sound surface developing an incipient lesion), the error can go only in one direction. Similarly, $D_2 \rightarrow D_1$ can also go only in one direction. But there are many more S surfaces than D_2's; therefore, the error rates may be equal, but there will be more $S \rightarrow I$ errors than $D_2 \rightarrow D_1$ errors. The following is the diagnostic error matrix of our examiner: The effects of a biased examiner will be dealt with later.

TABLE 1A
RISKS OF DECAY OF DENTAL SURFACES IN ADULT POPULATION (N=1109)

		Posterior				Anterior	
		B_1	B_2	M_1	M_2	Inc.	Can.
Occlusal	Upper	0.55	0.64	0.88	0.83	–	–
	Lower	0.33	0.60	0.81	0.84	–	–
Lingual	Upper	0.05	0.07	0.51	0.27	0.13	0.06
	Lower	0.02	0.07	0.24	0.19	0.01	0.01
Buccal	Upper	0.06	0.05	0.19	0.25	0.06	0.07
	Lower	0.08	0.10	0.52	0.37	0.01	0.01
Mesial	Upper	0.25	0.49	0.66	0.43	0.30	0.20
	Lower	0.10	0.30	0.60	0.53	0.04	0.05
Distal	Upper	0.52	0.59	0.52	0.25	0.29	0.24
	Lower	0.30	0.56	0.59	0.28	0.05	0.07

TABLE 1B
PROBABILISTIC WEIGHTING – VERSION I

		M_2	M_1	B_2	B_1	C	I
Occlusal	Upper	VH	VH	H	H	–	–
	Lower	VH	VH	H	M	–	–
Lingual	Upper	M	H	VL	VL	VL	L
	Lower	L	M	VL	VL	VL	VL
Buccal	Upper	M	L	VL	VL	VL	VL
	Lower	M	H	L	VL	VL	VL
Mesial	Upper	M	H	M	M	M	M
	Lower	H	H	M	L	VL	VL
Distal	Upper	M	H	H	H	M	M
	Lower	M	H	H	M	VL	VL

TABLE 1C
DIAGNOSTIC ERROR MATRIX OF OUR EXAMINER

	States				
True State	S	I_1^*	I_2	D_1	D_2
S=0	0.9000,	0.1000,	0.0000,	0.0000,	0.0000
I_1	0.0500,	0.9000,	0.0500,	0.0000,	0.0000
I_2	0.0000,	0.0500,	0.9000,	0.0500,	0.0000
D_1	0.0000,	0.0000,	0.0500,	0.9000,	0.0500
D_2	0.0000,	0.0000,	0.0000,	0.1000,	0.9000

*I_1 and I_2 are successive stages of incipiency; D_1 and D_2 are successive stages of actual decay. These were subsequently collapsed in the model to I and D, respectively.

(8) *The specification of professional intervention.* There are four aspects we need to determine:

(a) The distribution of the time lag between decayed and filled. The initiation of restoration is always based on

"clinical" judgment and is assumed to be normally distributed, with mean = 12 months, variance = six months.

(b) The equations of protection: (i) the protection offered by professional intervention, *e.g.*, the surface life of a filling for this time is assumed to be permanent, *i.e.*, safe until age 30; and (ii) protection by a preventive program is assumed in the linear form $y = A + Bx$, where y = the enhanced survival time, x = the true survival time at time t, A = a constant, and B = the coefficient of protection. In our simulation, the equations are as follows:

	A	B
Control	0	1
Treatment	1	1.5

(c) The distribution of sound surfaces lost due to restorative procedures.

By consulting life records in our files, the surface involvements by filling are assumed as follows:

	Probability of Type of Restoration			
	0	MO	MD	MOD
Occlusal Decay	0.40	0.24	0.24	0.12
	M,D	ML,DL	MB,DB	MLB,DLB
Mesial & Distal Decay	0.10	0.36	0.36	0.18

(9) *The transition distributions between states are assumed to be normal,* with mean and S.D. as listed below:

	Mean (yr)	S.D. (yr)
SI_1	*	
I_1I_2	0.50	0.50
I_1D_1	0.60	0.60
D_1D_2	0.75	0.50
D_2M	2.00	0.50
D_2F	1.00	0.50

*Incipience appearance controlled by gamma distribution, as listed earlier.

The Harris 300 computer was used to keep track of all surface development at all ages, between infancy and thirty years. In compliance with the above specifications, two populations were created. All subjects were identified and raised from infancy, their deciduous teeth erupted and exfoliated, and their permanent teeth erupted according to the distribution as specified. In order to simulate a clinical trial, two samples of 400 subjects each were selected at random. Each sample consisted of five age groups (8, 9, 10, 11, and 12), each having 80 subjects. To one sample was assigned a preventive program, and the other was used as a control. The trial was run until all subjects were thirty years old. However, for our purposes, we shall look only at the first three years of this experiment. Because the search for an appropriate measure of the caries experience has a direct bearing on how we analyze the data, we now need to examine the rationale before we examine the simulation results.

The rationale of an appropriate measure. — As a rule, all cariostatic agents tested in clinical trials are preventive rather than restorative. Once a DMF has occurred, the agent in question has very little effect on it. On the other hand, if an agent were to be effective, it must work in a manner that either prevents the sound surface from developing a lesion or prolongs the time from incipience to frank caries. Logically, the measure of caries experience should record the change from sound to incipience plus that from incipience to frank caries. Since we examine the subjects only at fixed time intervals, it is possible that a surface might move from incipience to frank caries or filled lesion within the interval; a similar development could occur from sound as well. The measure to be formulated, therefore, must consider these situations. Furthermore, there are intrinsic differences in resistance to caries attack among various types of surfaces on different teeth. This factor should also be incorporated into the simulation process. In the next section, these concepts will be treated with more heuristic rigor.

The construction of a new measure of caries experience. — Heuristically speaking, the systematic search for a new measure of caries experience requires a careful, detailed examination of the transitional behavior of the caries process over time, say, between points in time, t_i and t_{i+k}, where $i = 0,1,2,...r-1$; $K = 1,2,...r-1$, and $i+k \leq r$ and r = the total number of time periods the subjects were examined.

For the sake of clarity, it is obvious that, at any time, t, a surface can be in only one of the following five states:

(1) The sound state (S): A surface is said to be in the sound state if it has no observable signs of any caries activity;

(2) The incipient state (I): A surface is said to be in the incipient state if it has signs of caries activity but with no observable cavitation;

(3) The decayed state (D): A surface is said to be in the decayed state if it has an observable cavitation;

(4) The filled state (F): A surface is said to be in a filled state if it has a filling involvement;

(5) The missing state (M): A surface is said to be in the missing state if, after eruption, it has become absent from the oral cavity.

$$\text{Let } I = I_1 + I_2 \text{ and } U = D + F + M.$$

We may succinctly say that a surface at any time t must be in one of the three states, namely, S, I, or U.

Since the purpose of clinical trials is to detect the differences in the changes of states of the groups, the most serious defect of the current DMFS measure is that it only measures the end results, but pays no attention to the transition. We shall now delineate all possible transitions that could take place in a time interval.

Consider two consecutive times — say, t_i and t_{i+k}. It is also obvious, for any time interval (t_i, t_{i+k}), that a surface at t_{i+k} can be in only one of the following nine transition states:

Transition	Description and remarks of transition states
1. SS	The surface remains sound for the interval; its condition did not become worse — a favorable holding transition.
2. II	A surface remains in incipiency for the interval; also a favorable holding transition.
3. UU	Surface remains a DMF for the interval. Its condition could not become worse, from the preventive point of view. A moot holding transition.

4. SI	Surface changed from sound to incipiency; its condition took a turn for the worse – a natural deteriorative transition.
5. IU	Surface changed from incipiency to DMF; condition took a turn for the worse – a natural deteriorative transition.
6. SU	Surface changed from sound directly to DMF, a drastic change for the worse. Most transitions of this kind, however, were due to the involvement of innocent surfaces by restorative procedures. Only in rampant caries would an SU occur naturally. An artificial deteriorative transition.
7. IS	An incipient surface observed as sound – a diagnostic reversal.
8. UI	A DMF was observed as an incipiency – a diagnostic reversal.
9. US	A DMF was observed as sound – an unlikely diagnostic reversal.

The geneological flow of the states and transactions is sketched in Fig. 1.

From Fig. 1 it is apparent that the DMFS at any time t comprises IU + SU + UU. Since the cariostatic agent in question cannot act upon the UU, its inclusion in the measure serves no useful purpose.

The transition UU is mostly past history and has little bearing on the current new caries increment. Since SU is from sound to DMFS, it consists mostly of innocent surfaces lost due to restorative procedures. The only part pertaining to caries activities in DMFS is IU. The usefulness of DMFS as a measure of caries activity is obviously very inadequate. It appears that SI and IU are the only two transitions dealing with current caries activities; the others end up in U, but they got there due to some other irrelevant circumstances. We shall now examine the simulated data and see how all these transitions fared in the course of the clinical trials.

Findings from the simulation results.

A. *Test of transition means.* – (1) The transition means of the treatment and the control groups were compared by *t* tests, and the results are summarized in Tables 2a and 2b. We see that the results in general point in a similar direction and magnitude, indicating that, although our examiner is subject to 10% diagnostic error, his records nonetheless have shown the same tendencies with respect to the true condition. On the other hand, if he has biased diagnostic errors, the result will be quite different, which will be dealt with later. We observe that the favorable holding state transitions did not come into significant agreement until the third year. This is quite disappointing, mainly because the variances of these two transitions are quite large.

(2) The two natural deteriorative transitions, SI and IU, came into significant difference after one year. The artificial deteriorative transition, SU, and the existing holding state, UU, also came into significance in the second year.

(3) There were significantly more reversals, IS, in the treatment group than in the control group, as expected.

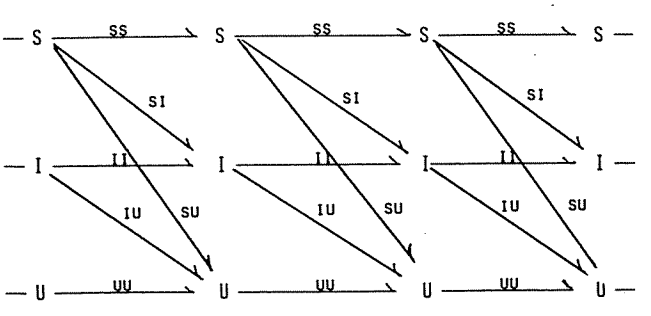

Fig. 1

TABLE 2A
COMPARISONS OF TRANSITION MEANS BETWEEN TREATMENT GROUPS BY DIAGNOSTIC STATUS AND YEARS

	True	Treatment	Control	*t* Values
Year 1	SS	63.76	61.96	1.20
	II	1.37	0.68	9.99
	UU	3.32	3.36	–0.17
	SI	0.00	1.59	–26.14
	IU	0.46	1.10	–9.97
	SU	0.82	0.93	–1.07
Year 2	SS	78.80	75.68	1.73
	II	0.82	0.87	–0.78
	UU	4.59	5.38	–2.62
	SI	1.04	1.79	–8.50
	IU	0.55	1.42	–12.40
	SU	0.87	1.35	–4.13
Year 3	SS	92.19	87.29	3.03
	II	1.41	0.99	5.35
	UU	6.00	8.15	–6.18
	SI	1.18	2.14	–9.75
	IU	0.45	1.68	–17.24
	SU	0.68	1.50	–7.64

TABLE 2B
COMPARISONS OF TRANSITION MEANS BETWEEN TREATMENT GROUPS BY DIAGNOSTIC STATUS AND YEARS

	Diagnosed	Treatment	Control	*t* Values
Year 1	SS	57.32	56.13	0.84
	II	1.45	0.89	7.61
	UU	3.27	3.34	–0.28
	SI	3.14	4.31	–9.33
	IU	0.53	1.15	–9.33
	SU	0.80	0.91	–1.07
	IS	3.18	2.87	2.16
Year 2	SS	71.14	68.51	1.60
	II	1.02	1.12	–1.37
	UU	4.60	5.38	–2.58
	SI	4.67	5.18	–2.89
	IU	0.63	1.43	–11.04
	SU	0.81	1.33	–5.08
	IS	3.80	3.52	1.82
Year 3	SS	83.28	78.98	2.91
	II	1.56	1.29	3.04
	UU	6.01	8.12	–6.08
	SI	5.47	6.02	–2.80
	IU	0.50	1.76	–16.80
	SU	0.65	1.50	–8.03
	IS	4.25	4.09	2.12

$t \geq 1.65$ sig. $P \leq 0.05$ for one-tailed test.
$t \geq 2.00$ sig. $P \leq 0.01$ for two-tailed test.

TABLE 3
THE *t* TESTS OF DIFFERENCES BETWEEN TREATMENT AND CONTROL MEANS OF THE THREE MEASURES (DMFS, TCI, AND NCI) BY DIAGNOSTIC STATES AND BY YEAR

	True	Treatment	Control	*t* Values
Year 1	DMFS (IU+SU+UU)	4.39	5.38	−2.62
	TCI (SI+IU+SU)	1.27	3.61	−17.57
	NCI (SI+IU)	0.46	2.50	−50.90
Year 2	DMFS	6.00	8.15	−6.18
	TCI	2.45	4.55	−13.72
	NCI	1.58	3.20	−14.38
Year 3	DMFS	7.13	11.34	−10.92
	TCI	2.30	5.33	−17.71
	NCI	1.62	3.83	−20.21
	Diagnosis	**Treatment**	**Control**	***t* Values**
Year 1	DMFS	4.59	5.38	−2.62
	TCI	4.47	6.37	−9.31
	NCI	3.66	5.41	−10.63
Year 2	DMFS	6.03	8.14	−6.07
	TCI	6.10	7.95	−8.06
	NCI	5.29	6.61	−6.07
Year 3	DMFS	7.16	11.37	−10.89
	TCI	6.61	9.27	−10.57
	NCI	5.97	7.78	−8.36

TABLE 4A
MEANS OF DMFS, TCI, AND NCI AND THEIR RESPECTIVE COMPONENTS ACCORDING TO TREATMENT GROUPS, DIAGNOSTIC STATUS, AND YEARS

Control True Status	SI	IU	SU	UU	
Year 1					
DMFS	–	1.10	0.92	3.36	5.38
TCI	1.59	1.10	0.92	–	3.61
NCI	1.59	1.10	–	–	2.69
Year 2					
DMFS	–	1.42	1.35	5.38	8.15
TCI	1.79	1.42	1.35	–	4.56
NCI	1.79	1.42	–	–	3.21
Year 3					
DMFS	–	1.69	1.50	8.15	11.34
TCI	2.14	1.69	1.50	–	5.33
NCI	2.14	1.69	–	–	3.83
Control Diagnosed Status	**SI**	**IU**	**SU**	**UU**	
Year 1					
DMFS	–	1.14	0.91	3.34	5.39
TCI	4.32	1.14	0.91	–	6.37
NCI	4.32	1.14	–	–	5.46
Year 2					
DMFS	–	1.43	1.33	5.38	8.14
TCI	5.19	1.43	1.33	–	7.95
NCI	5.19	1.43	–	–	6.62
Year 3					
DMFS	–	1.76	1.50	8.12	11.37
TCI	6.02	1.76	1.50	–	9.27
NCI	6.02	1.76	–	–	7.78

B. *Analytical comparisons of the three measures.* –
(1) The DMFS = IU + SU + UU.
(2) The total caries increment (TCI).
TCI = SI + IU + SU.
(3) The net caries increment (NCI). NCI = SI + IU.

I. *The tests of significance between the treatment and control means of the measurements* – namely, DMFS, total caries increment, and net caries increment. The detailed *t* test results are given in Table 3. It is quite evident that all three measures are significantly different in their means and treatment of the control groups. However, significant differences in this case are not enough: We need to know how the significance comes about. More precisely, how much of the variation in each of the measures came from pertinent components? – namely, from transitions SI and IU.

II. *Analysis of the partition of the means.*

(a) Partition analysis gives the component contribution toward the mean of each measure. The observed values of DMF, TCI, and NCI are tabulated in Tables 3, 4a, and 4b, and 5a and 5b, with respect to their component contributions. We note that the formation of the group means of DMFS, although it contains the IU transition, can only account for as low as 6% and at most 38% of its magnitude. Had we used the DMFS as a measure of caries activity, and supposing that we did obtain a significant difference, we cannot be sure from our results that the significant difference was indeed due to the transition IU. The total caries increment (TCI) fared much better; however, its component SU was mostly reflecting conversions of sound surfaces into fillings. It does not quite measure the incre-

TABLE 4B
MEANS OF DMF, TCI, AND NCI BY THEIR RESPECTIVE COMPONENTS BY TREATMENT GROUPS AND YEARS

True Treatment Group	SI	IU	SU	UU	
Year 1					
DMFS	–	0.45	0.82	3.32	4.59
TCI	0.00	0.45	0.82	–	1.27
NCI	0.00	0.45	–	–	0.45
Year 2					
DMFS	–	0.54	0.87	4.59	6.00
TCI	1.04	0.54	0.87	–	2.45
NCI	1.04	0.54	–	–	1.58
Year 3					
DMFS	–	0.44	0.68	6.00	7.12
TCI	1.18	0.44	0.68	–	2.30
NCI	1.18	0.44	–	–	1.62
Diagnosed Treatment Group	**SI**	**IU**	**SU**	**UU**	
Year 1					
DMFS	–	0.53	0.80	3.27	4.60
TCI	3.14	0.53	0.80	–	4.46
NCI	3.14	0.53	–	–	3.60
Year 2					
DMFS	–	0.63	0.80	4.60	6.03
TCI	4.67	0.63	0.80	–	6.10
NCI	4.67	0.63	–	–	5.30
Year 3					
DMFS	–	0.50	0.65	6.01	7.16
TCI	5.47	0.50	0.65	–	6.61
NCI	5.47	0.50	–	–	5.97

TABLE 5A (CONTROL GROUP)
COMPONENT CONTRIBUTION TOWARD THE MEANS OF DMFS, TCI, AND NCI BY TREATMENT GROUPS, DIAGNOSTIC STATUS, AND YEARS

Control True		Analysis of Means				
		SI	IU	SU	UU	% of Info. Due to SI + IU
Year 1	DMFS	0.00	0.20	0.17	0.63	20
	TCI	0.44	0.30	0.26		74
	NCI	0.59	0.41			100
Year 2	DMFS		0.17	0.17	0.66	34
	TCI	0.39	0.31	0.30	0	70
	NCI	0.56	0.44	0	0	100
Year 3	DMFS	0	0.15	0.13	0.72	28
	TCI	0.40	0.32	0.28	0	72
	NCI	0.56	0.44	0	0	100
Diagnosed Control		SI	IU	SU	UU	
Year 1	DMFS		0.21	0.17	0.62	38
	TCI	0.68	0.18	0.14	0	86
	NCI	0.79	0.21	0	0	100
Year 2	DMFS	0	0.17	0.16	0.66	18
	TCI	0.64	0.18	0.18	0	82
	NCI	0.91	0.09	0	0	100
Year 3	DMFS	0	0.15	0.13	0.72	15
	TCI	0.65	0.19	0.16	0	84
	NCI	0.77	0.23	0	0	100

TABLE 5B (TREATMENT GROUP)
COMPONENT CONTRIBUTION TOWARD THE MEANS OF DMFS, TCI, AND NCI BY TREATMENT GROUPS, DIAGNOSTIC STATUS, AND YEARS

Treatment True		Analysis of Means				
		SI	IU	SU	UU	% of Info. Due to SI + IU
Year 1	DMFS	0	0.10	0.18	0.72	10
	TCI	0	0.35	0.65	0	35
	NCI	0	1.00	0	0	100
Year 2	DMFS		0.09	0.15	0.76	9
	TCI		0.41	0.22	0.36	63
	NCI		0.66	0.34	0	100
Year 3	DMFS	0.00	0.06	0.10	0.84	6
	TCI	0.51	0.19	0.30	0.00	70
	NCI	0.72	0.28	0.28	0.00	100
Diagnosed						
Year 1	DMFS	0	0.12	0.17	0.71	12
	TCI	0.70	0.12	0.18	0	82
	NCI	0.87	0.13	0	0	100
Year 2	DMFS		0.10	0.13	0.76	10
	TCI	0.77	0.10	0.13	0	87
	NCI	0.88	0.12	0	0	100
Year 3	DMFS	0	0.07	0.09	0.84	7
	TCI	0.84	0.08	0.08	0	92
	NCI	0.92	0.08	0	0	100

TABLE 6A
ANALYSIS OF CORRELATION R^2

		SI	IU	SU	UU	% of Info. Due to SI + SU
Control True						
Year 1	DMFS	0.00	0.08	0.16	0.76	8
	TCI	0.31	0.28	0.41	–	59
	NCI	0.56	0.44			100
Year 2	DMFS	–	0.08	0.12	0.80	8
	TCI	0.34	0.31	0.35	–	65
	NCI	0.70	0.30	–	–	100
Year 3	DMFS	–	0.06	0.14	0.80	6
	TCI	0.37	0.24	0.39	–	61
	NCI	0.82	0.18	–	–	100
Diagnosed Control						
Year 1	DMFS	–	0.11	0.24	0.75	11
	TCI	0.62	0.16	0.22	–	78
	NCI	0.80	0.20	–	–	100
Year 2	DMFS	–	0.08	0.12	0.80	8
	TCI	0.65	0.15	0.20	–	80
	NCI	0.84	0.10	–	–	100
Year 3	DMFS	–	0.06	0.14	0.80	6
	TCI	0.62	0.14	0.24	–	76
	NCI	0.82	0.18	–	–	100

TABLE 6B
ANALYSIS OF CORRELATION R^2

		SI	IU	SU	UU	% of Info. Due to SI + IU
Treatment True						
Year 1	DMFS	0.00	0.03	0.23	0.74	3
	TCI	0.00	0.17	0.83	0.00	17
	NCI	0.00	1.00			100
Year 2	DMFS	0.00	0.03	0.15	0.82	3
	TCI	0.22	0.17	0.61	–	39
	NCI	0.67	0.33	–	–	100
Year 3	DMFS	–	0.02	0.08	0.90	2
	TCI	0.35	0.16	0.49	–	51
	NCI	0.73	0.27	–	–	100
Diagnosed True						
Year 1	DMFS	–	0.04	0.22	0.74	4
	TCI	0.56	0.10	0.34	–	68
	NCI	0.85	0.15			100
Year 2	DMFS	–	0.04	0.14	0.82	4
	TCI	0.64	0.12	0.24	–	76
	NCI	0.87	0.13			100
Year 3	DMFS	–	0.02	0.08	0.90	2
	TCI	0.64	0.12	0.24	–	76
	NCI	0.87	0.13			100

ment due to caries activity alone. The net caries increment (NCI), although much smaller, is 100% relevant information. From the above analysis of the contribution of the means, it is obvious that NCI is far superior to either DMFS or TCI.

(b) Analysis of the partitioning of R^2. Since each measure is a linear sum of two or three components, the multi-

ple correlations (R^2's) between the measures and their respective components are by definition equal to unity. By applying path analysis to this set of data, we can break down the correlation R^2 into the contribution from each of the contributing components. The results of the above analysis are given in Tables 6a and 6b. We see that in the DMFS the SI and IU can account for no more than 11% of its R^2. Any further reliance on this measure of caries experience would appear unwise indeed. In view of the above analyses, it appeared that both the TCI and the NCI are superior to the DMFS. In view of the percent of relevant information, the net caries increment (NCI) should be used.

C. *The effect of biased examiner's error.* –

(1) The Biased High-Low Errors: In a demonstration program, a cross-sectional examination of different age groups was obtained at the baseline, and the same schools were examined again in subsequent years. As the participants grew older, the age groups were compared with the respective age group at baseline. The design of this program is sketched in Fig. 2. There is a potential danger if there is examiner bias, unconscious or conscious. To demonstrate this danger, the control group of the current clinical trial was subjected to a 15% bias examination toward the higher end at the beginning, and then 15% bias in the opposite direction in subsequent years. In other words, the examiner calls one step higher 15% of the time at the baseline and one step lower, 15%, in subsequent years. The results are given in Table 7. We see from Table 7 that each subsequent year's comparison against the baseline was significant; in fact, at the fourth year, even the increment level was increased, but we know for a fact that, in this group, there was no treatment effect. All the differences were obtained by the examiner's deliberate errors. If we did not know this fact, it could pass as a significant result of a demonstration project.

(2) The Biased Low-High Errors: For some reason, some examiners considered diagnostic reversals a source of "embarrassment". The very simple gamesmanship required to avoid such "embarrassment" is deliberately to call the subjects with a lower status at the baseline examination and call them one status higher in subsequent years. To demonstrate this point, the same populations were subjected to 15% lower bias errors in the first year, and the 15% bias was reversed in the other direction in subsequent years. The results are presented in Tables 8a, 8b, and 8c, according to the changes of the various transitions and caries measures. As we examine the results of Tables 8a, 8b, and 8c, we see that, in general, the high-low game lowers the t value, although in the present case it doesn't matter, because most of them were significant anyway. However, where the

SCHEMATIC DIAGRAM OF A DEMONSTRATION PROGRAM

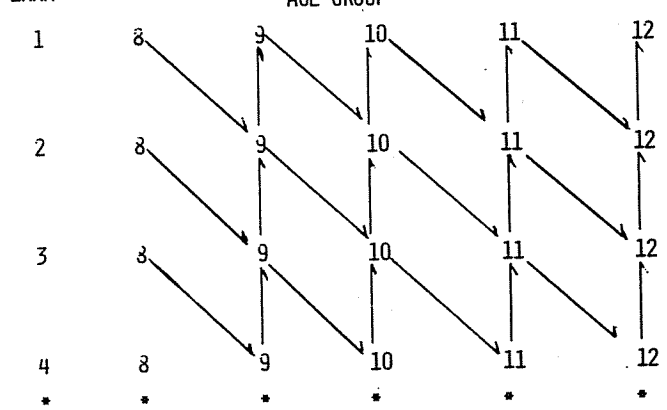

Fig. 2

TABLE 7
SAME POPULATION SUBJECT TO BIASED EXAMINER'S ERROR. 15% BIAS TOWARD H AT YEAR 1 AND 15% BIAS IN OPPOSITE DIRECTION IN SUBSEQUENT YEARS

Age	Baseline DMFS	Subsequent DMFS	Differences
9	0.98	0.68	−1.88*
10	2.61	2.23	−1.86*
11	5.16	4.60	−1.89*
12	8.48	7.63	−2.30**

TABLE 8A (FIRST-YEAR DATA)
SAME SAMPLE UNDER LOW-HIGH ERROR BIAS, 15% TOWARD LOW AT YEAR 1 AND 15% BIAS IN OPPOSITE DIRECTION IN SUBSEQUENT YEARS CLASSIFIED BY TREATMENT GROUPS, TRANSITION STATUS, AND CARIES EXPERIENCE MEASURES

Year 1

		Treatment	Control	t
SI	True	0.00	1.57	−26.80
	Diag.	9.68	10.79	−3.63
IU	True	0.48	1.23	−11.05
	Diag.	0.69	1.41	−9.47
SU	True	0.72	0.98	−2.64
	Diag.	0.72	1.13	−4.00
UU	True	3.31	3.15	0.64
	Diag.	3.16	3.01	0.66
DMFS	True	4.50	5.36	−2.90
	Diag.	4.57	5.55	−3.28
TCI	True	1.20	3.77	−19.26
	Diag.	11.09	13.33	−6.30
NCI	True	0.48	2.80	−26.42
	Diag.	10.37	12.20	−5.58

TABLE 8B (SECOND-YEAR DATA)
SAME SAMPLE UNDER LOW-HIGH ERROR BIAS, 15% TOWARD LOW AT YEAR 1 AND 15% BIAS IN OPPOSITE DIRECTION IN SUBSEQUENT YEARS CLASSIFIED BY TREATMENT GROUPS, TRANSITION STATUS, AND CARIES EXPERIENCE MEASURES

Year 2

		Treatment	Control	t
SI	True	1.04	1.77	−8.55
	Diag.	10.79	10.81	−0.06
IU	True	0.51	1.39	−13.27
	Diag.	0.70	1.62	−12.45
SU	True	0.99	1.56	−4.59
	Diag.	0.84	1.34	−4.57
UU	True	4.50	5.36	−2.91
	Diag.	4.54	5.51	−3.29
DMFS	True	6.00	8.32	−6.52
	Diag.	6.08	8.48	−6.73
TCI	True	2.54	4.73	−13.24
	Diag.	12.33	13.77	−3.79
NCI	True	1.55	3.17	−15.13
	Diag.	11.49	12.43	−2.75

TABLE 8C (THIRD-YEAR DATA)
SAME SAMPLE UNDER LOW-HIGH ERROR BIAS, 15% TOWARD LOW AT YEAR 1 AND 15% BIAS IN OPPOSITE DIRECTION IN SUBSEQUENT YEARS CLASSIFIED BY TREATMENT GROUPS, TRANSITION STATUS, AND CARIES EXPERIENCE MEASURES

Year 3

		Treatment	Control	t
SI	True	1.18	2.12	−9.45
	Diag.	12.80	13.00	−0.59
IU	True	0.49	1.67	−15.24
	Diag.	0.65	1.87	−14.41
SU	True	0.73	1.62	−7.88
	Diag.	0.63	1.49	−8.63
UU	True	6.00	8.31	−6.52
	Diag.	6.05	8.46	−6.78
DFMS	True	7.22	11.60	−11.19
	Diag.	7.33	11.83	−11.42
TCI	True	2.40	5.41	−17.28
	Diag.	14.08	16.36	−6.09
NCI	True	1.67	3.80	−16.06
	Diag.	13.45	14.87	−4.05

differences are not large, it may occasionally render a highly significant difference into insignificance in some very important component transitions. Furthermore, the rather unusual high rise in caries activities in both groups in the first year can be equally embarrassing.

We are now in a position to answer the first two questions:

(1) There is a better measure of caries experience than DMFS. It is the NIC; and

(2) Professional intervention comprised from 8% to 83% of the information, as seen in Tables 8a and 8b, depending on the measure. In general, it had a higher contribution to TCI than to DMFS. It had no effect on NCI.

Discussion.

(1) *The problem of examiner error.* In the simulated populations, we had the unusual luxury of knowing what the true status was; therefore, we would examine the effect of known examiner errors. However, in real life, the true status is unknown, and no single examiner is entirely accurate. There have been various measures proposed for examiner accuracy — *e.g.*, reliability, reproducibility, repeatability, etc. — but none of these is satisfactory.

In the final analysis, the accuracy of examination depends above all on the integrity and diagnostic skill of the examiner. The aspect of integrity, however, will not be discussed here.

Diagnostic skill may be considered as a function of the following aspects: (1) the clarity of diagnostic criteria — The different categories or classes of conditions should be mutually exclusive; (2) the ability of the examiner to adhere to the criteria; (3) the composition of the target population — If there are rampant caries or mostly healthy individuals in the population, then the accuracy of the examiner will be better than in a population with varying degrees of mixture; (4) The timing of the examination — At a beginning state, an error might occur; however, when the process has gone to its advanced stage, the chance of missing a large open cavity is remote.

Unfortunately, the examiner's ability changes from day to day, and from hour to hour. The problem of controlling this type of error may be reduced by (1) the examiner's taking appropriate breaks during the examination process, and (2) by having him examine a sub-set of the population twice to serve as a quality control measure. The subjects for quality control should be taken at different times of the day, and a chart should be plotted to observe if at any time the examination process is in or out of control.

We have seen the effects of the high-low and the low-high gamesmanship on clinical trial data. To detect a high-low game, all we need to do is examine the data longitudinally. If such a game has been played, we shall see a high reversal rate within a longitudinal set from the baseline to the next year. As has been shown before, the low-high game is not quite as harmful as the high-low game. If we find an unusually high rise in caries count in the first year in both control and treatment groups, but low reversal rates in subsequent years, a low-high game might very well be at play. This is usually done for the purpose of avoiding reversals. Reversals are naturally occurring phenomena; there is no shame associated with it. Any deliberate avoidance of reversals is uncalled for and could be counter-productive.

(2) *For how long should a clinical trial be held and with what kinds of subjects?* — Due to physical limitations, a clinical trial can seldom last more than three or four years at the outside. It depends on the efficacy of the substance as well as on the natural caries rate in the population. Based on the present simulation data, it appears that two to three years would be appropriate. All differences will become significant from the third year on.

(3) *What is the appropriate measure of caries activity in a clinical trial?* — From earlier analyses, we have amply demonstrated that the DMFS is unsuitable for measuring caries activities in clinical trials. It is even insufficient to serve as a basis for random assignments of treatment groups in a clinical trial — *e.g.*, two subjects may have the same DMFS count, say, 15, but one has 15 fillings and the other has 15 active lesions. There is a great deal of difference between the two subjects; however, from the DMFS point of view, they are exactly the same.

Ideally, a clinical trial should begin with an initial examination. The assignment of treatment groups should commence after the second examination, and treatment groups should be randomized within each stratum specified by the various types of state changes. However, clinical trials are extremely costly, and some compromise has to be made in this regard. It appears that an acceptable compromise may be the assignment to treatment groups on the basis of the numbers of incipient and frank lesions, which represent the current state of oral infected sites and potential sources of future decay.

The most serious defect is that DMFS records only the end results, not the changes. It is the changes in which we are interested. However, for almost half a century, we have been doing clinical trials with DMFS. Global fame and careers have been built with it. Although most of us were privately cognizant of its deficiencies, yet none have moved away from it. We became attached to it for the sake of simplicity and held on to it in the name of tradition. It is almost like a drunkard wandering around under a street lamp looking for his lost hat. Although he knew his hat was nowhere in sight, he stayed by the lamp because the light was better there than elsewhere. The time to abandon this measure has come.

Summary.

(1) Unbiased examiners' errors have little effect on the outcome of a clinical trial, provided that they are not larger than 10%. Unbiased error greater than 10% would tend to make the variance larger than expected.

(2) Demonstration schemes as discussed in this paper may show spurious positive effects, due to attempts by examiners to avoid reversals in diagnosis.

(3) Professional intervention is mostly responsible for the SU component – the direct transition from sound surface to DMF. These surfaces are lost for reasons other than caries activities *per se*. It should not be included in the measure of caries experience in clinical trials.

(4) The component UU represents the surfaces that show incipient caries at two consecutive examinations. It represents mostly past history and has little bearing on current caries activities.

(5) Since the DMFS = IU + SU + UU, and the contributions of IU as a rule are small, the use of DMFS as a measure of caries experience in clinical trials is unjustified.

(6) The NCI = SI + IU deals directly with the stage where the cariostatic agent is supposed to be working. There is a definite advantage to using this index as the caries experience measure in clinical trials.

REFERENCES

CARLOS, J.P. and GITTELSOHN, A.M. (1965): Longitudinal Studies of Natural History of Caries. II. A Life-table Study of Caries Incidence in the Permanent Teeth, *Arch Oral Biol* 10:739-751.

GRAINGER, R.M. and REID, D.B.W. (1954): Distribution of Dental Caries in Children, *J Dent Res* 33:613-623.

KATTI, S.K. (1960): Distribution of Dental Caries in Children, *J Dent Res* 39:501-505.

KIVIAT, P.J.; VILLANUEVA, R.; and MARKOWITZ, H.M.: **Simscript II.5 Programming Language**, E.C. Russell, Ed., Los Angeles, CA: CACI, Inc., October, 1975.

KLEIJNEN, J.P.C. (1974): **Statistical Techniques in Simulation**, Part I. New York: Marcel Dekker, Inc.

KLEIJNEN, J.P.C. (1975): **Statistical Techniques in Simulation**, Part II. New York: Marcel Dekker, Inc.

LEHMAN, R.S. (1977): **Computer Simulation and Modeling.** Hillsdale, NJ: Lawrence Erlbaum Assoc.

LU, K.H. (1966): An Analysis of the Caries Process by Finite Absorbing Markov Chains, *J Dent Res* 45:998-1015.

LU, K.H. (1966): A Path-probability Approach to Irreversible Markov Chains with an Application in Studying the Dental Caries Process, *Biometrics* 22:791-809.

LU, K.H. (1966): An Experimentally Supported "Law of Tooth Decay" and Its Application, *Arch Oral Biol* 11:757-768.

NAUS, J.I. (1975): **Data Quality Control and Editing.** New York: Marcel Dekker, Inc.

PRITSKER, A.A.B. and PEGDEN, C.D. (1979): **Introduction to Simulation and SLAM.** New York: John Wiley & Sons.

ROBERTS, F.S. (1976): **Discrete Mathematical Models.** Englewood Cliffs, NJ: Prentice-Hall, Inc.

SMITH, J.M. (1977): **Mathematical Modeling and Digital Simulation for Engineers and Scientists.** New York: John Wiley & Sons.