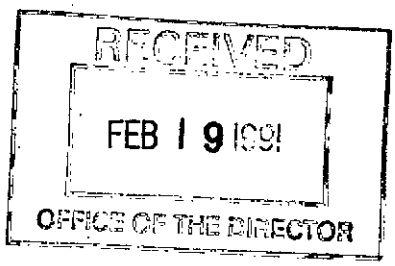
MAR 4 1991

Filcon Corporation

1186 ST. CLAIR
ST. PAUL, MN 55105
(612) 698-9661



February 12, 1991

Dr. Alfred Amendola
Centers for Disease Control
National Institute for Occupational
Safety and Health - ALOSH
944 Chestnut Ridge Road
Morgantown, WV 26505-2888

Dear Dr. Amendola:

I would like to thank you for the opportunity to make a presentation at the NIOSH Technical Conference held early last month.

On the whole, I found the conference valuable and learned a lot. I particularly welcomed the opportunity to meet several NIOSH personnel with whom I had talked on the phone on many occasions but had never met personally.

Enclosed is a revised draft of my presentation. The revision clarifies several points in my original presentation and includes some additional references. I would appreciate your substituting this draft for the original text and as Filcon's official comments for the record.

Thank you for the opportunity to participate. I look forward to doing it again.

Sincerely,

Trenton A. Niemeyer

Trenton A. Niemeyer

TAN/jb

Reinstallation and Justification For Using Laboratory Facefit Testing As The Primary Means For Respirator Selection

Three years ago NIOSH held public hearings to obtain comment on its proposed rule making 42CR84. At the time of the hearing, conversations with different manufacturers and other interested parties revealed that a great deal of confusion existed regarding why the Agency was making its proposal. Speculation abounded, and the majority of comments voiced were mainly objections on procedural rather than substantive issues.

As the weeks passed following the hearings, the story unfolded. Workplace testing by Myers, Nelson, Dixon, Campbell and others clearly showed that protection factors generated by laboratory facefit testing were artificially high and in many instances did not correlate with workplace fit testing. Laboratory fit testing was, consequently, to be discarded.

NIOSH went further and proposed that manufacturers conduct workplace testing to determine an acceptable protection factor which could be assigned to each individual respirator style. This protection factor would then be a guideline which respirator users could use in selecting an appropriate respirator for a particular application.

Filcon analyzed this scenario not only from the standpoint of what was the best way to design and present its respirators, but also from the viewpoint of product liability. We, as a manufacturer, do not only have to comply with NIOSH test procedures but must also safeguard against liability lawsuits which is a very ambiguous area. Any statements regarding an assigned protection factor which might appear on our labeling must, therefore, have a solid empirical basis and be fully supported.

After a lengthy analysis, Filcon came to several conclusions which, along with some recommendations, we would like to share with you now, and then discuss the issues on which these conclusions and recommendations are based. Collectively they define a direction we believe the respirator industry should pursue.

1) The state of the art in facefit testing is still in its infancy.
2) Any attempt by NIOSH to assign a generalized protection factor to a specific respirator is inherently dangerous, given the state of the art. This action could create a false sense of confidence in the user and foster misuse. To a mild degree, this seems to echo previous comments that facefit testing, at least for this purpose, does not belong in the certification process at this time.
3) The only current way to insure that respirator wearers receive adequate facefit protection would be through individual testing.
4) The principal means of testing should be laboratory facefit testing because, with potential modifications in protocol, it will represent a completely controlled environment. This recommendation assumes a final reconciliation between laboratory and workplace facefit testing and elimination of

the polarization in opinions that currently exists.

5) Laboratory facefit testing will "correlate" positively with workplace facefit testing if a margin of safety of 25X between the fit factor and the assigned protection factor is incorporated into the laboratory facefit test results. The 25X level is an approximation and would be subject to further adjustments. We can see possible scenarios where the figure could be adjusted downward to a 17 or 18 area and other scenarios where the figure could be adjusted upward in the vicinity of 30 to 32.

In making these conclusions, we realize that there are several major issues with which to contend and would like to addresses these presently.

Probably the most important issue is a reconciliation between workplace and laboratory facefit testing. In attempting to sort out this predicament, one fact hits us almost immediately. No one appears to have attempted a formal reconciliation of the two methodologies. At least, we could not find any published information to the contrary, and in view of the certainty and clarity which NIOSH is seeking, some sort of definitive declaration would be in order.

Laboratory testing appears to have been discarded without proof that the methodology was inherently defective. The only justification appears to be a lack of correlation between the two methods. This leaves open the equal possibility that workplace testing could also be defective. Only one particular form from among a whole series of possible protocols was shown to be deficient. No consideration seems to have been given to possibly repairing or modifying key parameters in laboratory facefit testing to make the method a better evaluator of actual worker facefit. Workplace testing seems to have been given a preference solely because it yielded more conservative data, and more conservative data was a safe haven in the face of uncertainty.

Furthermore, there has been no proof that workplace testing is the definitive methodology. If workplace testing is to be validated, how is this to be accomplished if alternative test methods are withdrawn from consideration? Clearly, credibility in test results can only be maximized if multiple test methods co-exist and are practiced.

There already is evidence for retaining laboratory facefit testing. For example, we have heard several times during this conference a concern about whether humidity in exhaled air will bias an in-mask test sample. The problem would be almost impossible to address at the work site, but it could easily be tackled in a test chamber. Healthy adults usually generate water vapor pressure at 47 mm of Hg. at STP. A similar level could be established and controlled in an ambient atmosphere inside a test booth. Reliability could at least be improved because sampling of both the ambient and in-mask air levels would be subject to the same bias. Further experimentation may also show humidity to be negligible once an equilibrium between two environments is established, which would affect accuracy.

Ultimately, we need to know exactly what, if anything, makes

workplace testing better than laboratory testing. The best way to do this is to conduct a side-by-side comparison of every parameter that exists in either method. Experiment design strategy should include provisions to make redundant all parameters common to bother methods. This would reduce the number of parameters that need to be analyzed and help pinpoint key factors necessary to resolve the two test methods. We believe an effort of this sort will, first, demonstrate the need for continuance of multiple test methods, and, second, show how the current laboratory facefit test protocol can be modified to improve test accuracy.

Filcon has already begun this process and has already located two sources of error in the current laboratory facefit protocol that illustrate in particular the second point.

We looked at Warren Meyers' work on the use of PAPRs. There was a correlation to degree that the workplace tests were consistently lower than those generated with laboratory testing. The correlation was low, but at least it was positive. This gave NIOSH a basis for downgrading the assigned protection factor from 1000 to 25.

In the case of negative pressure respirators, the situation had completely gone askew. A study by Lenhart and Campbell showed absolutely no correlation between the two methods when the sample was viewed as a whole. Variance in individual test results showed an apparent negative correlation in some cases and a positive, though very weak, correlation in others. Workplace testing levels ranged from 10 to over 40,000. One of the assumptions in the study showed that laboratory testing was used to prescreen test subjects. Minimum acceptance criterion for inclusion in the study was 250. The only way to approach such divergence in this information was to break down the results further and analyze them separately. We arbitrarily separated the results into two cases and decided to deal first with the "negative" correlation as a worst case scenario. This case showed a 25 to 1 ratio between laboratory and workplace testing. The only significance we assigned to this figure was that it was a point for comparison.

What could cause such a wide gap? To answer this question we started comparing individual parameters common to both methods and, without going into a lot of detail, discovered two possible errors. One was the absence of any means to measure and/or control the user inhalation level. This would be important because the level of air inhaled is directly proportional to the concentration of a challenge agent in the air when the agent is fully dispersed. How much air, for example, was inhaled in the deep breathing exercise in the NIOSH quantitative protocol? Did the deep breathing truly represent the maximum amount of air that a test subject might inhale under any circumstances?

There was no way to tell, and there was a real likelihood that the answer would be no. Needham et al in "Normal Standards For Lung Volumes, Intrapulmonary Gas-Mixing, and Maximum Breathing Capacity" showed a wide range of lung volumes and no central tendencies toward an average maximum inhalation rate for the entire population. There was also a marked dependence on a

subset of physiological parameters that affected maximum inhalation levels. Equally important were the assumptions on which this and other studies were based. West in his Essentials of Respiratory Physiology reports the existence of two independently functional compensatory drives which can trigger hyperventilatory responses, but which are both dormant when a subject is at rest. Foster et al in The Lung indicate that maximum expansion of the diaphragm is inhibited if a test subject is sitting. Dr. Karlman Wasserman, an internationally known expert in respiratory exercise testing, has reported increases in vital capacity to levels in the 85% to 90% range of total lung capacity by well conditioned athletes. This compares with 70% for the population as a whole and which many physiologists accept as a normal value.

All of this information forced us to consider the upper boundary of total inhaled volume as the only way to begin quantifying importance of an inhalation rate. Specifically, we began looking at the ratio of the mean tidal volume at anaerobic threshold to the mean tidal volume at rest. Values of mean tidal volume at anaerobic threshold, which is the same as vital capacity, ranged from 4,800 to 6,000 ml. A few extreme cases even showed values as high as 8,000 ml. Major variables contributing to this high upper boundary included age and body surface area. Males in the 20 to 30 year old age bracket, for example, showed a significantly higher maximum intake capacity according to Needham. Vital capacity, however, decreased by approximately 1/3 with age.

The wide range of values of this variable, therefore, could be a major contributor to intersubject variability of generated individual protection factors and have a pronounced effect on these values.

We then began hypothesizing. If we use a mean tidal volume of 400 ml for a subject at rest and a mean tidal volume of 6,000 ml at anaerobic threshold as measures of air intake , the ratio between these values is 15 to 1 and could explain in part the 25 to 1 ratio between laboratory and workplace facefit testing in our worst case scenario. If we choose the upper boundary of the tidal volume distribution at anaerobic threshold to be certain that all test subjects are included, the ratio grows to 20 to 1 (8,000 ml divided by 400 ml).

Another error contributing to the difference in the 25 to 1 ratio is the present composition of the quantitative facefit formula used to compute a protection factor based on the five exercises in the NIOSH protocol. Four of the five exercises are performed at rest. This creates a weighted average on the downside which favors a higher protection factor.

If the formula were restructured to provide a weighted average evenly distributed between a protection factor recorded at rest and a protection factor recorded at "maximum deep breathing", the increase in weighting would cut previously-generated laboratory protection factors in the neighborhood of 50%. This possibility alone could help to explain the remaining difference in the 25 to 1 ratio. Of course, accurate restructuring of this formula will require

proper assessment of work and inhalation rates that are commonly found in the workplace. In this respect, therefore, solution of the second error will depend on solution of the first.

Our hypotheses continued to grow, but did not explain why there was such a wide difference in workplace protection factors. Judging from the data we reviewed, however, we believe that the wide range of workplace protection factors shown in the Lenhart and Campbell study is in large part a reflection of the wide range of possible inhalation rates in humans.

Filcon is the first to admit that a formal study is necessary to prove this belief. All that we are doing here is exploring some possibilities. This we believe is one of the purposes of this conference. One of these possibilities is that the work rate is a prominent set of variables that has in fact been previously excluded from facefit testing and could be a key factor in reconciling the two test methodologies. Our effort to date has been limited to a literature review and did not preclude searching for other inhibiting causes. However, there is enough information in the literature to form a valid hypothesis and give direction for future research. Filcon recommends that NIOSH re-examine this dimension and include inhalation rates into calculating laboratory facefit protection factors.

Moreover, we recommend that NIOSH consider collecting physiological data during workplace testing. Until this is done, collected data that excludes quantification of physiological variables and especially inhalation levels will be applicable only to the specific test subject in his/her specific work environment for a specific job function. Any generalization of this information beyond this level would be dangerous. The fact remains: the same problem that adversely affected laboratory-generated data also affects workplace-generated data. There are no means to measure and/or control the level at which air is inhaled into a face mask. This is important because comparisons between the two methods will be meaningless until inhalation levels in both procedures are measured and equilibrated.

From a practical standpoint, there is a question of how. We recently saw an ad for the COSMED K2 portable exercise tester which straps to a user's chest and which is made by an Italian firm, whose name and address I don't have presently. One way or other, I will locate and forward this information to NIOSH. The tester contains a sample pump, micro-analyzer and transmitter which broadcasts continuously to a receiver and data storage holder located nearby. The tester is intended to monitor physiological parameters of athletes. A quick review of the brochure suggests that it would be easy to modify the unit for facefit testing purposes.

Also while on the subject of equipment, we suggest that NIOSH take a look at the totally automated cardiopulmonary analyzer manufactured by Medical Graphics Corporation of Vadnais Heights, Minnesota. Different models exist, but most provide real-time, breath by breath analysis of all pulmonary parameters. One unit provides real-time computation and feedback of 144 parameters.

Two final thoughts on laboratory facefit testing. One is

that even if laboratory and workplace are eventually correlated, laboratory testing should be given preference as the primary facefitting procedure. The preference is purely practical. We believe a reconciliation will require stress testing at anaerobic threshold to yield comparative results. Stress testing of average plant workers could cause unexpected health risks for the subjects and open the door to potential liability. Modification of the current NIOSH protocol to include a 25X margin of safety may be a practical way of dealing with this problem.

The other thought has to do with obtaining acceptance by employers of the idea of doing individual facefit testing. In conversations with Nelson Leidel, he expressed a concern that employers would simply ignore the requirement. After reflecting on this, I submit that the issue is one partly of our own attitudes. If we as respirator professionals don't take ourselves seriously about individual testing, who will? A quick comparison between facefit testing and eye examinations indicates that the issue may be one of expectation. No one thinks twice about conducting individual eye exams. Why shouldn't respirator facefit testing be given the same status?

The next issue we would like to address is the limitation of the U.S. Air Force Test Panel grid as a basis for selecting test subjects for workplace protection factor study. The adoption of this system was based on the recognition that no singular facepiece could fit everyone and that the grid system was an adequate means to classify test subjects. Implicit in this recommendation was the assumption that if a wearer whose face and lip lengths matched the same face and lip lengths which defined a particular stratum and who achieved a "satisfactory" facefit during a WPF study, then everyone whose facial measurements fit into the same category would automatically achieve the same level of facefit.

Filcon disagrees with this assumption. the anthropometric panel grid is not basically wrong. It's just incomplete. Analysis of various data sets of protection-factors shows that the range of protection factors within a particular stratum was considerably larger than the range of mean protection factors between adjacent stratums. Clearly, something was missing, and this was the third dimension.

Definition of this dimension proved to be challenging. It became clear that no single value could adequately serve as a universal measure. This forced us to look at the third dimension as a complex set of protrusion points extending from and perpendicular to the ridge of an enclosed, endless loop. Further analysis showed that a set of troughs had to be added to the set of protrusion points if a complete set of potential leakage sites were to be defined. Because each trough was located between adjacent protrusion points, the number of protrusion points equals the number of troughs.

This information lead to a model of the human face. In the case of a half-face mask, there were a minimum of 12 potential leakage sites. These are: the nose, the cheekbones, the jawbones (at least one tangent point on each side of the face), the chin, and the troughs in between. In the case of a full-face mask,

there are a minimum of 16 potential sites. These are: the middle of the forehead, the temporal crest ridges, the cheekbones, the jawbones, the chin, and the troughs in between.

Facefit could then be explained if we considered a facepiece as a similar rim with the same number of potential leakage sites. Mathematically, perfect facefit could be expressed as a 1 to 1 correspondence or mapping between all points on both rims when the rims are in contact with each other. Face mask leakage, consequently, is directly proportional to the number of contact points disengaged. Since leakage could occur at a singular site or at a combination of leakage sites, the number of possible combinations is extremely large.

Implicit in these concepts is that both rims have been rotated relative to each other so that one shape perfectly complements the other. The mapping itself is independent of the face and lip length dimensions. A very important point here is that each mapping is idiosyncratic to the particular shapes involved. For example, if a perfect mapping is assumed initially and then the nose protrusion is reduced 5MM on the face rim, it would create a gap in the mapping. The only way to restore the perfect mapping would be for the trough in the nasal area on the facepiece rim to extend 5MM toward the face rim in order to fill in the gap. This would establish a different mapping than the one initially assumed, and one which would be just as valid as the original. The argument could be extended by hypothesizing various singular or combined changes from an initially assumed perfect mapping. The result would be a large series of different mappings, which are mainly distinguished by the relative height of each protrusion point and trough to each other.

This multiplicity of acceptable facial patterns excludes the possibility of designating one singular face mask shape as the ideal universal shape. Multiplicity of acceptable facial patterns makes stratification of the third dimension necessary and, hopefully, a future goal. The difficulty presently is that there is no way to measure adequately the third dimension.

The absence of a stratified third dimension makes application of the two dimensional anthropometric grid impractical. Based on its current form, no individual could be categorized in a particular stratum because the stratum itself is not sufficiently categorized. If the two dimensional grid were used, any assignment of a general protection factor to a particular respirator would be potentially harmful to the respirator user. The only alternative under the present circumstances, consequently, is individual testing.

A third issue is the effects of dead space that occur from extended use of respirators throughout the workday. Dead space deals with CO_2 build-up inside the mask. Dead space causes hyperventilation and hypoxic pulmonary vasoconstriction (i.e. cardiac stress). The effects of dead space accumulate as the day progresses and demand some sort of relief. We have heard comments that workers periodically stop and take off their respirators to "catch their breath". It has been proposed that this response must be controlled if workplace protection factor studies are to yield accurate exposure estimates. To do this,

workers must be monitored continuously during the test to make sure that the face mask is not removed. While it is possible to do this, we submit that this action will force the test subject to reduce his/her work rate in order to conserve available O_2. Air intake will be reduced, and this will create a distortion in the test results. Extended test durations will not provide accurate results until such time as some sort of dead space compensation is introduced. Until then, shorter test durations are recommended and this suggestion will push workplace testing closer to laboratory facefit testing.

We also believe that individual testing should be looked at from a macro level. The use of statistics to make generalizations based on sample test results leaves open the possibility that at least part of the time, wrong decisions will be made. If, for example, a 95% probability is assumed with 95% confidence limit, does this mean that 5% of the time a mistake will occur? If so, what does this mean for the industrial worker? If we consider the population of respirator wearers to be conservatively 1,000,000, potentially 50,000 workers could be affected adversely by the use of statistical decision making. That's a lot! Philosophically, the use of statistics should have some limits. In the absence of an alternative basis for making decisions, statistics seem the best way to handle uncertainty. In the case of facefitting, individual testing can yield a higher rate of certainty. We believe individual testing should be given a preference on this basis.

I wanted to discuss two other issues, but I ran out of time needed to prepare a discussion because Filcon unexpectedly became involved with two projects for Operation Desert Shield. Briefly these issues are the mathematical inconsistencies of using log-normal distributions for quality assurance and the elimination of all qualitative facefitting for respirator selection. Perhaps we can do this another time.

Thank you for your attention.