



**HAL**  
open science

## Assessing the surrogacy in Raynaud's phenomenon

Olivier Gaget

► **To cite this version:**

Olivier Gaget. Assessing the surrogacy in Raynaud's phenomenon. Médecine humaine et pathologie. 2017. dumas-01958228

**HAL Id: dumas-01958228**

**<https://dumas.ccsd.cnrs.fr/dumas-01958228>**

Submitted on 3 Jan 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



## AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il n'a pas été réévalué depuis la date de soutenance.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact au SID de Grenoble :  
[bump-theses@univ-grenoble-alpes.fr](mailto:bump-theses@univ-grenoble-alpes.fr)

## LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

<http://www.cfcopies.com/juridique/droit-auteur>

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>



UNIVERSITÉ GRENOBLE ALPES  
UFR DE MÉDECINE DE GRENOBLE

Année : 2017

ASSESSING THE SURROGACY IN RAYNAUD'S PHENOMENON

THÈSE  
PRÉSENTÉE POUR L'OBTENTION DU TITRE DE DOCTEUR EN MÉDECINE  
DIPLOME D'ÉTAT

Olivier GAGET

[Données à caractère personnel]

SPÉCIALITÉ : SANTÉ PUBLIQUE

THÈSE SOUTENUE PUBLIQUEMENT À LA FACULTÉ DE MÉDECINE DE GRENOBLE

Le : 25/10/2017

DEVANT LE JURY COMPOSÉ DE

Président du jury :

M. le Professeur Jean-Luc CRACOWSKI

Membres :

M. le Docteur Matthieu ROUSTIT (directeur de thèse)

M. le Professeur José LABARÈRE

M. le Docteur Joris GIAI

*L'UFR de Médecine de Grenoble n'entend donner aucune approbation ni improbation aux opinions émises dans les thèses ; ces opinions sont considérées comme propres à leurs auteurs.*

Doyen de la Faculté : Pr. Jean Paul ROMANET

Année 2017-2018

ENSEIGNANTS A L'UFR DE MEDECINE

CORPS	NOM-PRENOM	Discipline universitaire
PU-PH	ALBALADEJO Pierre	Anesthésiologie réanimation
PU-PH	APEL Florent	Ophthalmologie
PU-PH	ARVIEUX-BARTHELEMY Catherine	Chirurgie générale
PU-PH	BAILLET Athan	Rhumatologie
PU-PH	BARONE-ROCHETTE Gilles	Cardiologie
PU-PH	BAYAT Sam	Physiologie
PU-PH	BENHAMOU Pierre Yves	Endocrinologie, diabète et maladies métaboliques
PU-PH	BERGER François	Biologie cellulaire
MCU-PH	BIDART-COUTTON Marie	Biologie cellulaire
MCU-PH	BOISSET Sandrine	Agents infectieux
PU-PH	BONAZ Bruno	Gastro-entérologie, hépatologie, addictologie
PU-PH	BONNETIERRE Vincent	Médecine et santé au travail
PU-PH	BOREL Anne-Laure	Endocrinologie, diabète et maladies métaboliques
PU-PH	BOSSON Jean-Luc	Biostatistiques, informatique médicale et technologies de communication
MCU-PH	BOTTARI Serge	Biologie cellulaire
PU-PH	BOUGEROL Thierry	Psychiatrie d'adultes
PU-PH	BOUILLET Laurence	Médecine interne
PU-PH	BOUZAT Pierre	Réanimation
MCU-PH	BRENIER-PINCHART Marie Pierre	Parasitologie et mycologie
PU-PH	BRICAULT Ivan	Radiologie et imagerie médicale
PU-PH	BRICHON Pierre-Yves	Chirurgie thoracique et cardio-vasculaire
MCU-PH	BRIOT Raphaël	Thérapeutique, médecine d'urgence
MCU-PH	BROUILLET Sophie	Biologie et médecine du développement et de la reproduction
PU-PH	CAHN Jean-Yves	Hématologie
PU-PH	CARPENTIER Françoise	Thérapeutique, médecine d'urgence
PU-PH	CARPENTIER Patrick	Chirurgie vasculaire, médecine vasculaire
PU-PH	CESBRON Jean-Yves	Immunologie
PU-PH	CHABARDES Stephan	Neurochirurgie
PU-PH	CHABRE Olivier	Endocrinologie, diabète et maladies métaboliques
PU-PH	CHAFFANJON Philippe	Anatomie
PU-PH	CHARLES Julie	Dermatologie
PU-PH	CHAVANON Olivier	Chirurgie thoracique et cardio-vasculaire
PU-PH	CHIQUET Christophe	Ophthalmologie

PU-PH	CHIRICA Mircea	Chirurgie générale
PU-PH	CINQUIN Philippe	Biostatistiques, informatique médicale et technologies de communication
MCU-PH	CLAVARINO Giovanna	Immunologie
PU-PH	COHEN Olivier	Biostatistiques, informatique médicale et technologies de communication
PU-PH	COURVOISIER Aurélien	Chirurgie infantile
PU-PH	COUTURIER Pascal	Gériatrie et biologie du vieillissement
PU-PH	CRACOWSKI Jean-Luc	Pharmacologie fondamentale, pharmacologie clinique
PU-PH	CURE Hervé	Oncologie
PU-PH	DEBILLON Thierry	Pédiatrie
PU-PH	DECAENS Thomas	Gastro-entérologie, Hépatologie
PU-PH	DEMATTEIS Maurice	Addictologie
MCU-PH	DERANSART Colin	Physiologie
PU-PH	DESCOTES Jean-Luc	Urologie
MCU-PH	DETANTE Olivier	Neurologie
MCU-PH	DIETERICH Klaus	Génétique et procréation
MCU-PH	DOUTRELEAU Stéphane	Physiologie
MCU-PH	DUMESTRE-PERARD Chantal	Immunologie
PU-PH	EPAULARD Olivier	Maladies Infectieuses et Tropicales
PU-PH	ESTEVE François	Biophysique et médecine nucléaire
MCU-PH	EYSSERIC Hélène	Médecine légale et droit de la santé
PU-PH	FAGRET Daniel	Biophysique et médecine nucléaire
PU-PH	FAUCHERON Jean-Luc	Chirurgie générale
MCU-PH	FAURE Julien	Biochimie et biologie moléculaire
PU-PH	FERRETTI Gilbert	Radiologie et imagerie médicale
PU-PH	FEUERSTEIN Claude	Physiologie
PU-PH	FONTAINE Éric	Nutrition
PU-PH	FRANCOIS Patrice	Epidémiologie, économie de la santé et prévention
MCU-MG	GABOREAU Yoann	Médecine Générale
PU-PH	GARBAN Frédéric	Hématologie, transfusion
PU-PH	GAUDIN Philippe	Rhumatologie
PU-PH	GAVAZZI Gaëtan	Gériatrie et biologie du vieillissement
PU-PH	GAY Emmanuel	Neurochirurgie
MCU-PH	GILLOIS Pierre	Biostatistiques, informatique médicale et technologies de communication
MCU-PH	GRAND Sylvie	Radiologie et imagerie médicale
PU-PH	GRIFFET Jacques	Chirurgie infantile
PU-PH	GUEBRE-EGZIABHER Fitsum	Néphrologie
MCU-PH	GUZUN Rita	Endocrinologie, diabétologie, nutrition, éducation thérapeutique
PU-PH	HAINAUT Pierre	Biochimie, biologie moléculaire
PU-PH	HENNEBICQ Sylviane	Génétique et procréation
PU-PH	HOFFMANN Pascale	Gynécologie obstétrique
PU-PH	HOMMEL Marc	Neurologie
PU-MG	IMBERT Patrick	Médecine Générale
PU-PH	JOUK Pierre-Simon	Génétique
PU-PH	JUVIN Robert	Rhumatologie

PU-PH	KAHANE Philippe	Physiologie
MCU-PH	KASTLER Adrian	Radiologie et imagerie médicale
PU-PH	KRACK Paul	Neurologie
PU-PH	KRAINIK Alexandre	Radiologie et imagerie médicale
PU-PH	LABARERE José	Epidémiologie ; Eco. de la Santé
MCU-PH	LABLANCHE Sandrine	Endocrinologie, diabète et maladies métaboliques
MCU-PH	LANDELLE Caroline	Bactériologie - virologie
MCU-PH	LAPORTE François	Biochimie et biologie moléculaire
MCU-PH	LARDY Bernard	Biochimie et biologie moléculaire
MCU-PH	LARRAT Sylvie	Bactériologie, virologie
MCU - PH	LE PISSART Audrey	Biochimie et biologie moléculaire
PU-PH	LECCIA Marie-Thérèse	Dermato-vénéréologie
PU-PH	LEROUX Dominique	Génétique
PU-PH	LEROY Vincent	Gastro-entérologie, hépatologie, addictologie
PU-PH	LEVY Patrick	Physiologie
PU-PH	LONG Jean-Alexandre	Urologie
PU-PH	MAGNE Jean-Luc	Chirurgie vasculaire
MCU-PH	MAIGNAN Maxime	Thérapeutique, médecine d'urgence
PU-PH	MAITRE Anne	Médecine et santé au travail
MCU-PH	MALLARET Marie-Reine	Epidémiologie, économie de la santé et prévention
MCU-PH	MARLU Raphaël	Hématologie, transfusion
MCU-PH	MAUBON Danièle	Parasitologie et mycologie
PU-PH	MAURIN Max	Bactériologie - virologie
MCU-PH	MC LEER Anne	Cytologie et histologie
PU-PH	MERLOZ Philippe	Chirurgie orthopédique et traumatologie
PU-PH	MORAND Patrice	Bactériologie - virologie
PU-PH	MOREAU-GAUDRY Alexandre	Biostatistiques, informatique médicale et technologies de communication
PU-PH	MORO Elena	Neurologie
PU-PH	MORO-SIBILOT Denis	Pneumologie
PU-PH	MOUSSEAU Mireille	Cancérologie
PU-PH	MOUTET François	Chirurgie plastique, reconstructrice et esthétique ; brûlologie
MCU-PH	PACLET Marie-Hélène	Biochimie et biologie moléculaire
PU-PH	PALOMBI Olivier	Anatomie
PU-PH	PARK Sophie	Hémo - transfusion
PU-PH	PASSAGGIA Jean-Guy	Anatomie
PU-PH	PAYEN DE LA GARANDERIE Jean-François	Anesthésiologie réanimation
MCU-PH	PAYSANT François	Médecine légale et droit de la santé
MCU-PH	PELLETIER Laurent	Biologie cellulaire
PU-PH	PELLOUX Hervé	Parasitologie et mycologie
PU-PH	PEPIN Jean-Louis	Physiologie
PU-PH	PERENNOU Dominique	Médecine physique et de réadaptation
PU-PH	PERNOD Gilles	Médecine vasculaire
PU-PH	PIOLAT Christian	Chirurgie infantile
PU-PH	PISON Christophe	Pneumologie

PU-PH	<b>PLANTAZ Dominique</b>	<b>Pédiatrie</b>
PU-PH	<b>POIGNARD Pascal</b>	<b>Virologie</b>
PU-PH	<b>POLACK Benoît</b>	<b>Hématologie</b>
PU-PH	<b>POLOSAN Mircea</b>	<b>Psychiatrie d'adultes</b>
PU-PH	<b>PONS Jean-Claude</b>	<b>Gynécologie obstétrique</b>
PU-PH	<b>RAMBEAUD Jacques</b>	<b>Urologie</b>
PU-PH	<b>RAY Pierre</b>	<b>Biologie et médecine du développement et de la reproduction</b>
PU-PH	<b>REYT Émile</b>	<b>Oto-rhino-laryngologie</b>
PU-PH	<b>RIGHINI Christian</b>	<b>Oto-rhino-laryngologie</b>
PU-PH	<b>ROMANET Jean Paul</b>	<b>Ophthalmologie</b>
PU-PH	<b>ROSTAING Lionel</b>	<b>Néphrologie</b>
MCU-PH	<b>ROUSTIT Matthieu</b>	<b>Pharmacologie fondamentale, pharmaco clinique, addictologie</b>
MCU-PH	<b>ROUX-BUISSON Nathalie</b>	<b>Biochimie, toxicologie et pharmacologie</b>
MCU-PH	<b>RUBIO Amandine</b>	<b>Pédiatrie</b>
PU-PH	<b>SARAGAGLIA Dominique</b>	<b>Chirurgie orthopédique et traumatologie</b>
MCU-PH	<b>SATRE Véronique</b>	<b>Génétique</b>
PU-PH	<b>SAUDOU Frédéric</b>	<b>Biologie Cellulaire</b>
PU-PH	<b>SCHMERBER Sébastien</b>	<b>Oto-rhino-laryngologie</b>
PU-PH	<b>SCHWEBEL-CANALI Carole</b>	<b>Réanimation médicale</b>
PU-PH	<b>SCOLAN Virginie</b>	<b>Médecine légale et droit de la santé</b>
MCU-PH	<b>SEIGNEURIN Arnaud</b>	<b>Epidémiologie, économie de la santé et prévention</b>
PU-PH	<b>STAHL Jean-Paul</b>	<b>Maladies infectieuses, maladies tropicales</b>
PU-PH	<b>STANKE Françoise</b>	<b>Pharmacologie fondamentale</b>
MCU-PH	<b>STASIA Marie-José</b>	<b>Biochimie et biologie moléculaire</b>
PU-PH	<b>STURM Nathalie</b>	<b>Anatomie et cytologie pathologiques</b>
PU-PH	<b>TAMISIER Renaud</b>	<b>Physiologie</b>
PU-PH	<b>TERZI Nicolas</b>	<b>Réanimation</b>
MCU-PH	<b>TOFFART Anne-Claire</b>	<b>Pneumologie</b>
PU-PH	<b>TONETTI Jérôme</b>	<b>Chirurgie orthopédique et traumatologie</b>
PU-PH	<b>TOUSSAINT Bertrand</b>	<b>Biochimie et biologie moléculaire</b>
PU-PH	<b>VANZETTO Gérald</b>	<b>Cardiologie</b>
PU-PH	<b>VUILLEZ Jean-Philippe</b>	<b>Biophysique et médecine nucléaire</b>
PU-PH	<b>WEIL Georges</b>	<b>Epidémiologie, économie de la santé et prévention</b>
PU-PH	<b>ZAOUTI Philippe</b>	<b>Néphrologie</b>
PU-PH	<b>ZARSKI Jean-Pierre</b>	<b>Gastro-entérologie, hépatologie, addictologie</b>

**PU-PH** : Professeur des Universités et Praticiens Hospitaliers  
**MCU-PH** : Maître de Conférences des Universités et Praticiens Hospitaliers  
**PU-MG** : Professeur des Universités de Médecine Générale  
**MCU-MG** : Maître de Conférences des Universités de Médecine Générale

# Remerciements

Au Docteur Matthieu Roustit, mon directeur de thèse, pour m'avoir coaché durant la rédaction. Pour m'avoir proposé une ribambelle de sujets de travail tous plus intéressants les uns que les autres. Pour m'avoir, quand c'était nécessaire, mis un coup ... de boost pour me remettre sur les rails. Pour avoir, en un mot comme en cent, accepté de diriger ce travail de thèse.

Au Professeur Jean-Luc Cracowski, pour avoir accepté de présider mon jury de thèse, mais aussi et surtout pour m'avoir fourni l'environnement de travail le plus stimulant de mon DES de santé publique. Pour avoir non seulement écouté, mais même encouragé mes élucubrations statistiques. Pour m'avoir mis en contact avec une multitude de personnes toutes aussi intéressantes que les sujets de travail évoqués au paragraphe précédent. Pour s'être battu bec et ongles, enfin, pour ce poste d'assistant.

Au Professeur José Labarère, pour avoir accepté de juger ce travail, mais aussi pour les discussions en staff méthodo, toujours formatrices. Pour la formation générale en santé publique. Pour s'être battu avec la même vigueur que Jean-Luc Cracowski.

Au Docteur Joris Giai, pour avoir accepté de juger ce travail, et aussi pour m'avoir mis sur la voie des biostats en général et du master B3S en particulier. Pour m'avoir donné envie de découvrir le bayésien. Pour les TD de Gibbs-sampling. Pour les discussions tout à fait sérieuses et en même temps pas du tout, sur mille et un sujets méthodologiques... ou autres ;)

À Céline, ma co-titulaire (comprenez qui pourra), pour avoir supporté mes nuits blanches, mes divagations statistiques, médicales, politiques, autres... Si j'en écrivais plus, je tomberais dans le mièvre, donc merci. Je t'aime.

À la Twingo rouge.

À Charles Khouri, sans qui ce travail n'existerait tout simplement pas.

À toute l'équipe du CIC, un service génial qui accueillera, je l'espère, un grand nombre d'internes de santé publique qui sauront en profiter !

À la R core team, qui ne lira jamais ces remerciements, mais qui fait un logiciel incroyable. À la communauté R, qui ne fait qu'y ajouter de la valeur.

À André, sans qui le master aurait été bien plus difficile ;)

À Elsa Bey, Catherine Arvieux, Sandrine Barbois, Flavie Colin, Olivier Epaulard... Il peut paraître étrange de remercier ceux qui vous donnent du travail, mais ce travail m'a toujours permis d'apprendre et de progresser. Et pour vous donner du travail, il faut déjà vous faire confiance. J'espère l'avoir méritée.

À Vincent Canuel, la toute première personne à m'avoir dit que la santé publique était une spécialité. À Matt Leacock : c'est en jouant à un jeu de société, à l'été 2013, que je me suis fait la réflexion que la médecine m'intéressait plus à l'échelle collective qu'individuelle. Ce jeu s'appelait... Pandémie.

À David, Aurélien et Pierre, pour avoir suivi avec moi le feuilleton « poste ou pas poste ? ».



À Claudine Roland, pour ne PAS avoir malencontreusement planté une fourchette dans l'œil de quelqu'un.

À ma mère pour avoir essayé de faire de moi quelqu'un de sérieux. Jugez sur pièces.

À mon père, pour la boîte de chimie et le microscope.

(Ici, la logique voudrait que je remercie l'école Polytechnique pour avoir fait de moi ce que je suis. Je ne suis pas sûr de la pertinence de ce remerciement. Surtout qu'il serait étrange de remercier une école où je ne suis PAS passé, tout en passant sous silence l'université Grenoble-Alpes, l'université Pierre-et-Marie-Curie, l'université Paris-XI et l'université Claude-Bernard-Lyon 1)

À Luc, pour Pink Floyd, Led Zeppelin et Iron Maiden.

Au Docteur Roland Gitard, qui « savait bien que je n'étais pas fait pour la médecine clinique ».

*« Les grandes personnes aiment les chiffres »*

*Antoine de Saint-Exupéry*

## Introduction

Raynaud's phenomenon (RP) is a clinical condition associated with exaggerated vasoconstriction of the microvasculature in the extremities, in response to cold or emotional stress (1). Primary RP is idiopathic and affects 4 to 6% of the general population with significant geographic variations (2). It is usually a benign, seasonal condition that mostly requires conservative measures such as avoiding exposure to cold and using protective clothing. By contrast, secondary causes of RP include connective tissue diseases (mainly systemic sclerosis, SSc), hand-arm-vibration syndrome, vascular compression, and drugs (1,3). SSc-related RP can lead to severe complications such as digital ulceration or gangrene (1).

Most randomized controlled trials (RCTs) assessing the effect of therapies on RP have used one, or a combination, of the following outcomes: frequency of attacks, total duration of attacks over a given period of time, and scores such as the Raynaud Condition Score (RCS). The latter is a daily self-assessment of RP activity using a 0–10 ordinal scale, with 0 being no impact of the disease, and 10 the highest imaginable impairment (4). They involve filling a diary every day to rate these 3 outcomes. However, these outcomes suffer from drawbacks: the RCS is a subjective measure, the duration of a crisis can be inaccurately measured by the patient and the patient may forget to fill the diary on a day without any attack, biasing the results. Furthermore, assessing these outcomes requires following the patients for a long period of time in order to get a large amount of diaries per patient.

For these reasons, a number of surrogate outcomes have been used in RCTs on RP, in an attempt to address the issue of subjectiveness and to reduce follow-up. These surrogates include: finger arterial pressure, finger skin temperature, and skin blood flow assessed with laser Doppler or plethysmography. All these tests can be performed at rest or in response to a cold stress, also known as cold challenge. However, the surrogacy of these outcomes has not been evaluated.

Surrogacy is the ability of a proposed surrogate outcome to reflect the effect of a treatment on a true outcome. In a founding paper, Prentice identified 4 criteria for perfect surrogacy (5). Let S be the surrogate outcome, T the true outcome, and Z the treatment group. Let  $F(X)$  be the distribution of the variable X, and  $F(X|Y)$  the distribution of X, supposing Y is known. Then perfect surrogacy is achieved if:

$$F(T|Z) \neq F(T) \quad [1]$$

$$F(S|Z) \neq F(S) \quad [2]$$

$$F(T|S) \neq F(T) \quad [3]$$

$$F(T|S,Z) = F(T|S) \quad [4]$$

[1] means the treatment should have an effect on the true outcome and it is necessary for [2], [3] and [4] to be interpretable. [2] means that the treatment should also have an effect on the surrogate outcome. [3] means that knowing the value of the surrogate should bring some information on the value of the true outcome. Finally, [4] means that, if one knows the value of the surrogate outcome, knowing the treatment group should not bring any further information, allowing the surrogate outcome to fully "capture" the treatment effect.

In practice, these criteria suffer from several major drawbacks. First, they are only valid for perfect surrogacy, but imperfect surrogacy can be useful too. Secondly, they require having an effective treatment in order to assess surrogacy. Once a treatment is proven to be effective on a given outcome, it may seem unnecessary and unethical to study the surrogacy of a surrogate outcome. Finally, these criteria are not directly usable.

Buyse et al. refined these criteria, proposing usable criteria (6). First, the correlation coefficient between the treatment effect on the true outcome, on one hand, and the treatment effect on the

surrogate outcome on the other hand should be high (close to 1, although Buyse et al don't suggest any particular threshold), and the correlation coefficient between the true outcome and the surrogate outcome should be high. The latter should be assessed at an individual level, from single or multiple crossover trials. In contrast, trial-level surrogacy can be estimated using a meta-analytic approach. Buyse suggests a good surrogate outcome is an outcome which performs well on both criteria. Good individual-level surrogacy is an indication that the surrogate and the clinical outcomes are causally linked, while good trial-level surrogacy indicates that the surrogate captures all the treatment effect.

The aim of this study is to assess the surrogacy of the various outcomes in trials assessing the effect of drugs on RP, combining trial-level and individual-level surrogacy.

## **Methods**

### **Meta-analytic approach**

Trial-level surrogacy was assessed as an ancillary objective of an umbrella systematic review with a network meta-analysis, assessing pharmacological interventions for the treatment of secondary RP. This systematic review complies with the PRISMA statement guideline, and the full protocol is available online (PROSPERO registry, CRD42017057518). Eligibility criteria included parallel or crossover RCTs. A crossover trial was eligible if a washout of more than one week was present. RCTs were deemed eligible if included patients had secondary RP. If the trial included primary and secondary RP patients, it was eligible if subgroup analysis was available for secondary RP patients, or if more than 50% of patients had secondary RP.

The following surrogate outcomes were extracted: finger skin temperature; skin blood flow (whether it was assessed with laser Doppler imaging, laser Doppler flowmetry, or plethysmography); and digital arterial pressure; at rest or in response to a cold challenge. The

following clinical outcomes were used as references: frequency of RP attacks, cumulative daily duration of RP attacks, and RCS or any other numeric scale exploring RP severity. Among all RCTs selected in the meta-analysis, we excluded those where no surrogate had been used. One trial was excluded because it had evaluated only surrogates, but no clinical outcome. Each RCT was reviewed, and means and standard deviations were extracted for the outcomes in each treatment group. When subgroups were reported (for example, primary and secondary RP phenomenon), each subgroup was extracted as a different RCT. In one RCT, laser Doppler flowmetry measurements were reported before and after the cold challenge. In that RCT, the difference between before and after the cold challenge was used as the cold challenge outcome, and laser Doppler flowmetry measurements before the cold challenge were used as ambient-temperature measurements.

When the standard deviation was not reported, it was computed using either the standard deviation of the mean or the confidence interval width. In 3 RCTs, the mean and standard deviation were not reported for the surrogates, but the first quartile (Q1), the median (Q2) and the third quartile (Q3) were reported instead. In this case, the mean and standard deviation were computed with the following formulae:  $\text{mean} = (Q1 + Q2 + Q3)/3$  and  $\text{SD} = (Q3-Q1)/1.35$ .

Finally, we computed the standardized mean difference for each outcome in each RCT. When data was reported before and after treatment, the standard deviation of the treatment-related variation was calculated assuming a within-patient  $R^2 = 0.7$ . This value was obtained from the trial used for individual-level surrogacy. The SMD was then calculated using this standard deviation.

When data was reported only after treatment, the SMD was calculated between the treatment data and the placebo data.

If the treatment effect on the surrogate were to predict the treatment effect on the true criteria, then the standardized mean difference (SMD) of the surrogate in a given trial should predict the SMD

for the clinical outcomes in this trial. In order to test this, we calculated Spearman's correlation coefficients between surrogate SMDs and clinical outcomes SMDs.

The correlation coefficient to a cold challenge response was calculated by pooling all types of surrogate outcomes, provided they were coupled with a cold challenge. Digital arterial pressure, skin blood flow and skin temperature correlation coefficients were calculated by pooling data for each test, whether coupled with a cold challenge or not.

### **Individual-level approach**

Individual-level surrogacy was based on a series of randomized, controlled, N-of-1 trials (PROFIL trials, ClinicalTrials.gov Identifier NCT02050360) aiming to estimate the efficacy of on-demand sildenafil as a treatment of moderate to severe primary or secondary RP. Three treatments were tested: on-demand sildenafil at a 40mg dose, on-demand sildenafil at an 80-mg dose, and placebo. N-of-1 trials consist in repeated cross-over trials during which treatments are administered in repeated blocks. In the present case, each block consisted in 3 one-week periods, with a different treatment each week, in a random order. This way, a given patient would get 40-mg sildenafil, 80-mg sildenafil and placebo over each 3-week period. Patients could complete two to five blocks.

A cold provocation test was performed on the first day of each treatment period of the first block. During this cold provocation test, the right hand was placed in a fenestrated box with air at  $9\pm 1^{\circ}\text{C}$ . Microvascular reactivity to hand cooling was assessed by measuring cutaneous blood flow with laser Speckle contrast imaging (LSCI), a non-invasive and reproducible technique. Skin blood flow was measured before, during and after a 30-min hand cooling. Data were expressed as cutaneous vascular conductance (CVC), i.e. blood flow divided by mean arterial pressure, recorded over 10 regions of interest on the back of three fingers and the back of the hand.

Using the PROFIL data, we computed a treatment effect for every treatment for every patient for the following outcomes: RCS, RP frequency, RP duration, and CVC during cooling (at 5 min and 30 min), and after cooling (at 5 min and 60 min). The individual-level surrogacy was assessed by calculating Spearman's correlation coefficient between each of the treatment effects for the clinical outcomes on one hand, and the treatment effect on CVC on the other hand.

All correlation coefficient confidence intervals were computed through bootstrapping.

## Results

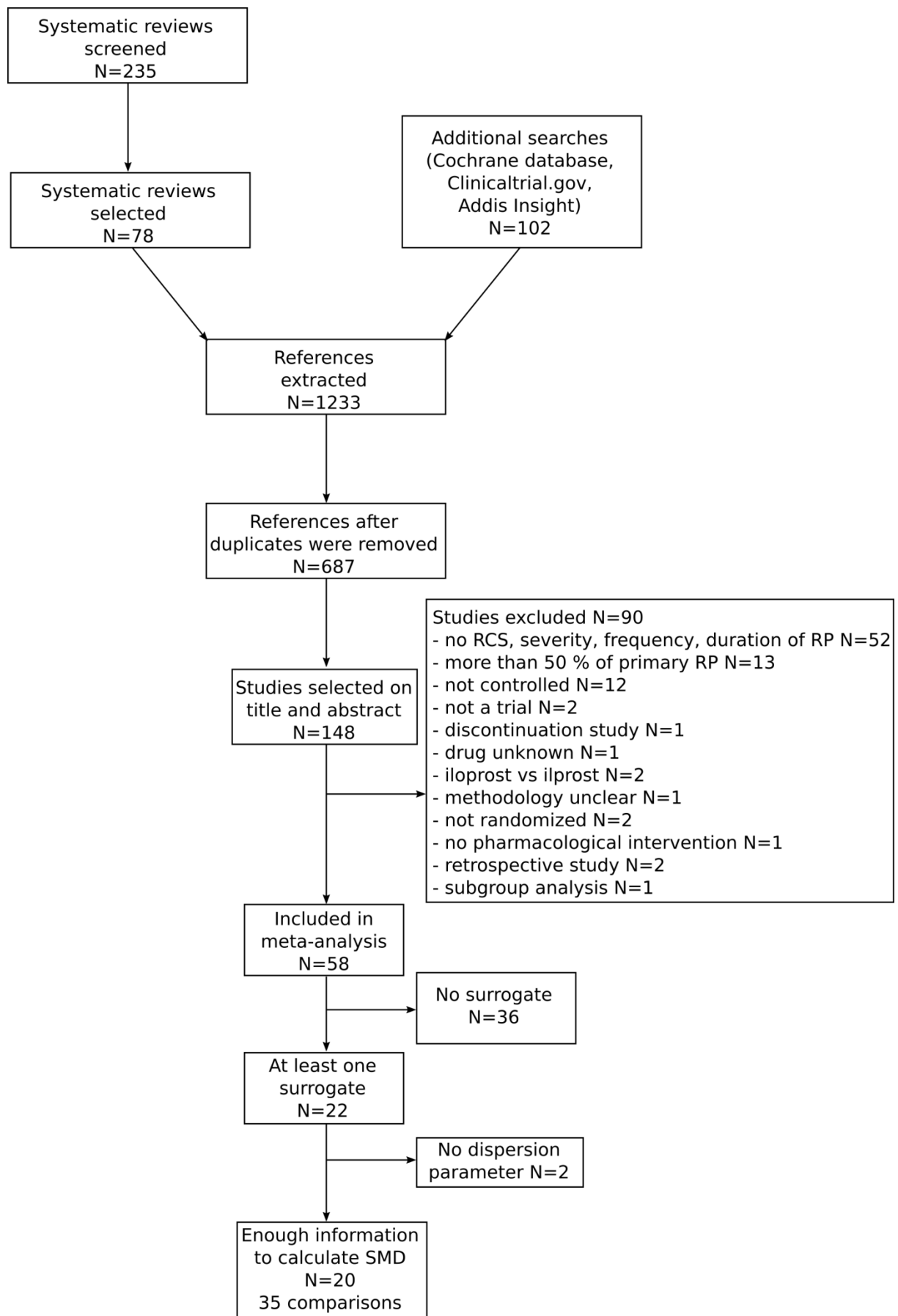
### Meta-analytic approach

Final analysis included 35 comparisons from 20 RCTs. The flow chart of included RCTs is represented in Figure 1. Among those comparisons, 18 used a cold challenge to as a surrogate outcome, coupled with different techniques to assess skin blood flow (Table 1).

**Table 1. Methods used to assess skin perfusion as a surrogate.**

Type of outcome	Cold challenge	No cold challenge
Digital arterial pressure	3	1
Skin blood flow	11	11
Skin temperature	4	5
Total	18	17





**Figure 1: flow chart of included RCTs.**

There was a variety of methods to achieve cooling, and heterogeneity in temperatures used for reactivity tests. A description of those methods is provided in Table 2.

There was no significant correlation between any of the surrogate outcomes and clinical outcomes, either during a cold challenge or at rest (Table 3).

**Table 2. Description of the methods used for cold challenges.**

Cooling system (temperatures in Celsius)	Number of comparisons
Water (15°C and 16°C)	10
Warm room and cold room (20°C, 28.3°C)	2
Ice (temperature unreported)	2
Unknown (-10°C)	2
Air (4°C)	1
Cooling blanket (15°C)	1

**Table 3. Correlation coefficients between the surrogate outcomes and the clinical outcomes.**

Outcomes	RCS	Duration of attacks	Frequency of attacks
Cold challenge	-0.07 (N=15, p=0.79) 95% CI : [-0.65 - 0.59]	0.30 (N=11, p=0.37) 95% CI : [-0.44 - 0.83]	0.42 (N=17, p=0.09) 95% CI : [-0.1 - 0.80]
All trials			
Skin blood flow*	-0.12 (N=8, p=0.64) 95% CI : [-0.96 - 0.86]	-0.47 (N=8, p=0.24) 95% CI : [-1 - 1]	-0.02 (N=9, p=0.92) 95% CI : [-0.93 - 0.78]
Skin temperature	0.15 (N=5, p=0.8) 95% CI : [-1 - 1]	0.21 (N=4, p=0.79) 95% CI : [-1 - 1]	-0.05 (N=5, p=0.93) 95% CI : [-1 - 1]

*Data are expressed as Spearman's correlation coefficients, with their p-values and 95% confidence intervals (95%CI). Due to the lack of data, the surrogacy of digital arterial pressure could not be assessed. \*The following methods to estimate skin blood flow were pooled: laser Doppler flowmetry, laser Doppler imaging, and plethysmography.*

## Individual-level approach

Thirty-eight patients were enrolled in the PROFIL trials. Among them, 33 completed a full cold challenge for all treatment periods. There was no significant correlation between skin blood flow during or immediately after the cold challenge, and clinical outcomes (Table 4).

**Table 4. Correlation between skin blood flow at various times of the cold challenge and estimated treatment effect on the RCS, frequency of attacks and total duration of attacks over the day.**

Time	RCS	Frequency of attacks	Duration of attacks
Cooling (5 min)	-0.03 [-0.29 - 0.22] (0.80)	-0.11 [-0.34 – 0.13] (0.37)	-0.13 [-0.36 – 0.13] (0.37)
Cooling (30 min)	-0.07 [-0.32 - 0.18] (0.37)	-0.01 [-0.38 – 0.09] (0.93)	-0.23 [-0.45 – 0.02] (0.07)
Post-cooling (5 min)	-0.06 [-0.31 - 0.19] (0.61)	-0.03 [-0.40 – 0.07] (0.79)	-0.23 [-0.46 - 0.02] (0.06)
Post-cooling (60 min)	0.02 [-0.23 – 0.29] (0.83)	-0.14 [-0.37 – 0.11] (0.26)	-0.20 [-0.44 - 0.05] (0.11)

*Data are expressed as Spearman's correlation coefficients, with their 95% confidence intervals in square brackets and p-values in brackets.*

## Discussion

Raynaud's phenomenon is a condition associated with impaired cutaneous microvascular reactivity to local cooling (7). As a consequence, most treatments used for RP aim at maintaining adequate perfusion in the fingers; and skin blood flow has thus been used in many clinical trials, often coupled with a cold challenge. Yet, its surrogacy has never been properly established.

On the individual level, there was no significant correlation between treatment effect on any clinical outcome and the response to cold challenge. This holds true for every cooling and post-cooling time. Individual results, while not significant, show a trend: all correlation coefficients are negative or close to zero, suggesting that a treatment which improves the cold challenge response tends to decrease the RP duration over a day, which is the expected behaviour. However, this is of

course not sufficient to consider skin perfusion during a cold challenge as an acceptable surrogate outcome.

Our results do not show any significant relationship between skin microvascular reactivity to cold challenge (or skin blood flow at rest) and clinical outcomes at the trial level, and therefore do not support acceptable surrogacy. Although general guidelines do not exist, thresholds for trial-level correlation coefficients to conclude to surrogacy have been proposed. In oncology for example, a surrogate outcome may be considered acceptable if the lower boundary of the 95% confidence interval for Pearson's correlation coefficient is above 85%; in contrast, surrogacy will be rejected if the upper boundary of the 95% confidence interval for Pearson's correlation coefficient is below 0.7 (8). While this applies to Pearson's correlation coefficient, it makes sense to be at least as demanding if Spearman's correlation coefficient is used, since Spearman's correlation coefficient is more robust in the case of non-linear relationships. With these criteria, surrogacy of a cold challenge may even be rejected for RCS. The absence of even a trend to a relationship between skin blood flow in response to cold and clinical outcomes is surprising. It is worth noting that the small number of available RCTs did not allow for a precise estimation of correlation coefficients at the trial level. For this reason, we had to pool different surrogate outcomes, which introduced significant heterogeneity. Indeed, different indicators of skin perfusion were used (i.e. temperature, skin blood flow), and among them techniques were also different. In addition, cold challenges were not standardized, using various cooling methods and temperatures. Such heterogeneity is in itself interesting information, highlighting the lack of consensus concerning what a good cold challenge is, and at what temperature it should be performed.

Trials in this field remain particularly challenging: they are characterized by a strong placebo effect and a large variability in treatment effects (9). Clinical outcomes are subject to memory bias, and possibly Heisenberg effect: the sole fact of observing an outcome may have an effect on this outcome. All these effects have led to a large number of negative trials. Developing good surrogates and assessing their surrogacy at the individual level and at the trial level is a necessity

to reduce the variability of observed outcomes. However, developing good surrogates is a long process and assessing their surrogacy is even longer (10). A first step is to develop reproducible, standardized outcomes. The second step is to test the individual-level correlation between surrogate and clinical outcomes. This test is simple and can be achieved on a single trial. Finally, the trial-level correlation is the longest to test, as it requires having multiple RCTs in which the proposed surrogate has been reproducibly gathered. Unlike blood markers, this step presents a supplementary difficulty in the case of cold challenges, which involve precise measurement conditions and specific devices.

In the meantime, RP trials should focus on large sample sizes in order to increase statistical power. On an individual level, N-of-1 trials should be explored, as they provide the best estimation of treatment effect on a single patient (11). This is achieved by gathering a large number of observations for every patient, allowing to further reduce the variability of outcomes, thus improving the estimation of treatment effect.

## REFERENCES

1. Herrick AL. The pathogenesis, diagnosis and treatment of Raynaud phenomenon. *Nat Rev Rheumatol*. 2012 août;8(8):469–79.
2. Maricq HR, Carpentier PH, Weinrich MC, Keil JE, Palesch Y, Biro C, et al. Geographic variation in the prevalence of Raynaud's phenomenon: a 5 region comparison. *J Rheumatol*. 1997 May;24(5):879–89.
3. Khouri C, Blaise S, Carpentier P, Villier C, Cracowski J-L, Roustit M. Drug-induced Raynaud's phenomenon: beyond  $\beta$ -adrenoceptor blockers. *Br J Clin Pharmacol*. 2016 Jul;82(1):6–16.
4. Merkel PA, Herlyn K, Martin RW, Anderson JJ, Mayes MD, Bell P, et al. Measuring disease activity and functional status in patients with scleroderma and Raynaud's phenomenon. *Arthritis Rheum*. 2002;46(9):2410–20.
5. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. *Stat Med*. 1989 Apr;8(4):431–40.
6. Buyse M, Molenberghs G, Burzykowski T, Renard D, Geys H. The validation of surrogate endpoints in meta-analyses of randomized experiments. *Biostatistics*. 2000;1(1):49–67.
7. Roustit M, Blaise S, Millet C, Cracowski JL. Impaired transient vasodilation and increased vasoconstriction to digital local cooling in primary Raynaud's phenomenon. *Am J Physiol Heart Circ Physiol*. 2011;301(2):H324-30.
8. Validity of surrogate endpoints in oncology Executive summary of rapid report A10-05, Version 1.1. In: Institute for Quality and Efficiency in Health Care: Executive Summaries [Internet]. Cologne, Germany: Institute for Quality and Efficiency in Health Care (IQWiG); 2005. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK198799/>
9. Seibold JR, Wigley FM. Clinical Trials in Raynaud Phenomenon: A spoonful of sugar (pill) makes the medicine go down (in flames). *Arthritis Rheumatol* [Internet]. 2017 Aug 31 [cited 2017 Sep 12]; Available from: <http://doi.wiley.com/10.1002/art.40307>
10. Roustit M, Maggi F, Isnard S, Hellmann M, Bakken B, Cracowski JL. Reproducibility of a local cooling test to assess microvascular function in human skin. *Microvasc Res*. 2010;79(1):34–9.
11. Vohra S. N-of-1 trials to enhance patient outcomes: identifying effective therapies and reducing harms, one patient at a time. *J Clin Epidemiol*. 2016 Aug 1;76:6–8.

THESE SOUTENUE PAR : OLIVIER GAGET

TITRE : ASSESSING THE SURROGACY IN RAYNAUD'S PHENOMENON

## CONCLUSION

Le phénomène de Raynaud est une affection caractérisée par des épisodes douloureux de vasoconstriction des extrémités en réponse au froid ou à un stress émotionnel. Bien que généralement bénin, il peut se compliquer d'ulcérations cutanées lorsqu'il est secondaire à une autre pathologie, notamment la sclérodermie systémique. Les critères de jugement généralement utilisés dans les essais sur le phénomène de Raynaud sont des critères cliniques : fréquence et durée des crises, gêne ressentie par le patient. Ces critères sont sujets à des biais de mémorisation et requièrent des temps de suivi longs, et présentent par ailleurs une importante variabilité inter-individuelle. Différents critères de jugement intermédiaires objectifs ont donc été proposés : pression artérielle digitale, flux sanguin cutané ou température digitale, en réponse ou non à un test au froid. Toutefois, la capacité de ces critères intermédiaires à prédire fidèlement l'effet d'un traitement sur le phénomène de Raynaud n'a jamais été testée.

Pour qu'un critère de jugement intermédiaire puisse être utilisé comme critère de substitution, il doit être validé à deux niveaux : au niveau méta-analytique et au niveau individuel. La validation méta-analytique a été réalisée en étude ancillaire d'une revue systématique avec méta-analyse réseau sur les traitements pharmacologiques du phénomène de Raynaud secondaire (PROSPERO : CRD42017057518). Nous avons extrait les effets des traitements sur les critères de jugement intermédiaires que sont la pression sanguine digitale, la température digitale et le flux sanguin digital. Pour chacun de ces critères, nous avons calculé le coefficient de corrélation de Spearman avec chaque critère de jugement clinique. Au niveau individuel, nous avons utilisé les données d'une série d'essais cliniques contrôlés randomisés de taille 1 (N-of-1 trials) évaluant l'effet du sildénafil à la demande dans le Raynaud sévère (ClinicalTrials.gov : NCT02050360). La méthodologie N-of-1 permet de calculer un effet traitement pour chaque patient. Nous avons calculé les coefficients de corrélation de Spearman entre, d'une part, l'effet traitement individuel sur chaque critère de jugement clinique et, d'autre part, sur le flux sanguin cutané en réponse à un refroidissement loco-régional.

Au niveau méta-analytique, 20 essais ont été retenus dans l'analyse, présentant un total de 35 comparaisons: certains essais présentaient plus d'une comparaison. Aucun critère de jugement intermédiaire ne s'est avéré significativement corrélé à un critère de jugement clinique. De plus, les coefficients de corrélation étaient de signes différents, interdisant d'analyser ne serait-ce qu'une tendance. Au niveau individuel, 33 patients ont été inclus dans l'analyse. Tous les coefficients de corrélation étaient négatifs, ce qui pourrait correspondre aux attentes d'un bon critère de jugement intermédiaire. Toutefois, aucun coefficient de corrélation n'était significativement différent de zéro.

En conclusion, notre étude montre que les critères couramment utilisés dans les essais thérapeutiques ne permettent pas de prédire l'effet thérapeutique en conditions de vie réelles. Ceci pourrait être dû à une limitation méthodologique : le nombre d'essais inclus était petit, les méthodes de mesure employées étaient variables. Cependant, il ne se dessine même pas de tendance concernant la corrélation entre critères intermédiaires et critères cliniques. Notre étude montre donc la nécessité de développer des critères intermédiaires dans le phénomène de Raynaud et de tester leur capacité à prédire les critères cliniques, qui sont ceux qui ont réellement un impact du point de vue des patients.

VU ET PERMIS D'IMPRIMER

Grenoble, le 12/10/2017

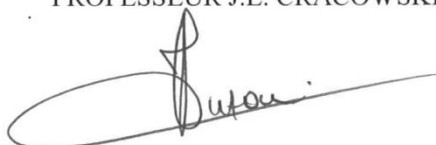
LE DOYEN

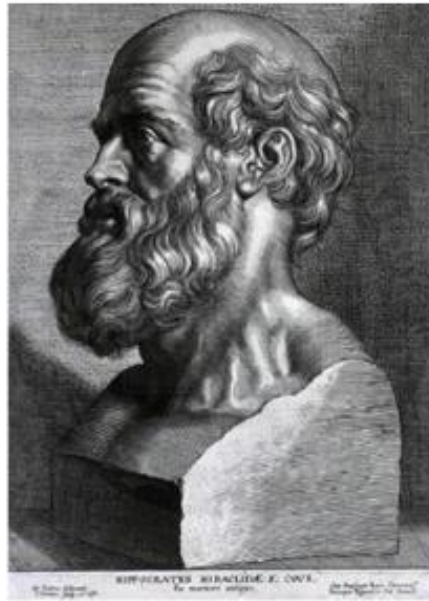
LE PRESIDENT DE LA THESE

J.P. ROMANET



PROFESSEUR J.L. CRACOWSKI





## **SERMENT D'HIPPOCRATE**

*En présence des Maîtres de cette Faculté, de mes chers condisciples et devant l'effigie d'HIPPOCRATE,*

*Je promets et je jure d'être fidèle aux lois de l'honneur et de la probité dans l'exercice de la Médecine.*

*Je donnerais mes soins gratuitement à l'indigent et n'exigerai jamais un salaire au dessus de mon travail. Je ne participerai à aucun partage clandestin d'honoraires.*

*Admis dans l'intimité des maisons, mes yeux n'y verront pas ce qui s'y passe ; ma langue taira les secrets qui me seront confiés et mon état ne servira pas à corrompre les mœurs, ni à favoriser le crime.*

*Je ne permettrai pas que des considérations de religion, de nation, de race, de parti ou de classe sociale viennent s'interposer entre mon devoir et mon patient.*

*Je garderai le respect absolu de la vie humaine.*

*Même sous la menace, je n'admettrai pas de faire usage de mes connaissances médicales contre les lois de l'humanité.*

*Respectueux et reconnaissant envers mes Maîtres, je rendrai à leurs enfants l'instruction que j'ai reçue de leurs pères.*

*Que les hommes m'accordent leur estime si je suis fidèle à mes promesses.*

*Que je sois couvert d'opprobre et méprisé de mes confrères si j'y manque.*